

Imperial College London
Department of Mathematics

Heterogeneously Coupled Maps. From High to Low Dimensional Systems through Ergodic Theory

Matteo Tanzi

Supervised by:
Prof. Tiago Pereira and Prof. Sebastian van Strien

Report submitted in partial fulfilment of the requirements of the PhD degree at
Imperial College London.

To my parents.

Abstract

In this thesis we study ergodic theoretical properties of high-dimensional systems coupled on graphs. The local dynamics at each node is hyperbolic and coupled with other nodes according to the edges of the graph. We focus our attention on the case of graphs with heterogeneous degrees meaning that most of the nodes make a small number of interactions, while a few hub nodes have very high degree. For such high-dimensional systems there is a regime of the interaction strength for which the coupling is small for poorly connected systems, and large for the hub nodes. In particular, global hyperbolicity might be lost. We show that, under certain hypotheses, the dynamics of hub nodes can be very well approximated by a low-dimensional system for exponentially long time in the size of the network and that the system exhibit hyperbolic behaviour in this time window. Even if this describes only a long transient, we argue that this is the behaviour that one expects to observe in experiments. Such a description 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). The HCM we study provide a new paradigm to explain why and how the dynamics of a network dynamical system can change across layers.

Declaration of Originality

I hereby declare that the work presented in this thesis is my own work, unless otherwise specified, and that I have rightfully acknowledged and referenced the work of others. The material presented has not previously been submitted for assessment at Imperial College London or elsewhere. Chapter 3 and part of the introduction will appear as a joint paper with Tiago Pereira and Sebastian van Strien.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work

Acknowledgments

The author would like to thank his supervisors Prof. Tiago Pereira and Prof. Sebastian van Strien for their invaluable guidance and kind support they have provided him during his doctoral studies. The author is also grateful to Prof. Gerhard Keller for fruitful conversations and precious insights into the topics of coupled map lattices and ergodic theory.

Funding from the Imperial College Scholarship Scheme, BREUDS and the ERC AdG grant no 339523 RGDD is acknowledged.

Notation

$[n]$	compact notation for the set $\{i \in \mathbb{N} : 1 \leq i \leq n\}$, $n \in \mathbb{N}$.
$[m, n]$	compact notation for the set $\{i \in \mathbb{N} : m \leq i \leq n\}$, $m, n \in \mathbb{N}$.
$\mathbb{R}^+$	strictly positive real numbers.
$B_\varepsilon(x)$	the set $\{y \in X : d(y, x) < \varepsilon\}$ where (X, d) is a given metric space.
$B_\varepsilon(\Lambda)$	the set $\{y \in X : d(y, \Lambda) < \varepsilon\}$ where (X, d) is a given metric space.
m_n	Lebesgue measure on $\mathbb{T}^n$, $n \in \mathbb{N}$.
m_W	Lebesgue measure induced on the submanifold $W \subset \mathbb{T}^n$.
$\ \mathbf{v}\ _p$	p -norm of $\mathbf{v} \in \mathbb{R}^n$, for $1 \leq p \leq +\infty$, unless specified otherwise.
$\mathcal{M}(n, m)$	$n \times m$ matrices with real coefficients.
$ \mathbf{M} $	modulus of the determinant of the square matrix $\mathbf{M}$.
$\mathbf{M}_i$	i -th row vector of the matrix $\mathbf{M}$.
$\mathbf{M}^i$	i -th column vector of the matrix $\mathbf{M}$.
$\hat{\mathbf{x}}_{\mathbf{i}}$	given $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{T}^N$, $\hat{\mathbf{x}}_{\mathbf{i}} := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N) \in \mathbb{T}^{N-1}$.
$\varphi(x_i; \hat{\mathbf{x}}_{\mathbf{i}})$	given $\varphi : \mathbb{T}^N \rightarrow X$, $\varphi(x_i; \hat{\mathbf{x}}_{\mathbf{i}}) = \varphi(x_1, \dots, x_N)$.
$\mathcal{O}(N), \mathcal{O}(\varepsilon)$	$\mathcal{O}(N)/N$, $\mathcal{O}(\varepsilon)/\varepsilon$ are bounded for $N \rightarrow \infty$ and $\varepsilon \rightarrow 0^+$ respectively.

Contents

1	Introduction	1
1.1	Emergent Dynamics on Heterogeneously Coupled Maps (HCM).	2
1.2	Low Dimensional Reduction and Hub Synchronisation	3
1.3	Dynamics Across Connectivity Scales: Predictions and Experiments	6
1.3.1	Predicted Impact of the Network Structure	6
1.3.2	Numerical Simulations of Large-Scale Layered Random Networks	7
1.4	Impact of the Network Structure: Homogeneous versus Heterogeneous Networks	8
1.4.1	Informal Statement of Theorem B on Coherence of Hub Dynamics	9
1.4.2	Informal Statement of Theorem C Comparing Dynamics on Homogeneous and Heterogeneous Networks	9
1.5	Coupled Maps with Nonzero Distortion	11
1.6	Guide to the Thesis	11
2	Network Dynamical Systems	13
2.1	Graphs	14
2.1.1	Homogeneous and Heterogeneous Graphs	15
2.1.2	Random Graphs	16
2.2	Dynamics on Networks	18
2.2.1	Network Maps and Network Vector Fields	18
2.2.2	Important Features of Dynamics on Networks	19
2.2.3	Maps with Additive Coupling Structure	20
2.3	Literature Review	20
2.3.1	Globally Coupled Systems and the Kuramoto Model	21
2.3.2	Coupled Maps Lattices and Networks	24
2.3.3	Bifurcation and Symmetries in Network Dynamical Systems	25
2.3.4	Necessity of a New Approach to Deal With HCM	26
2.4	High-Dimensional Dynamical Systems	27
2.4.1	Statistical Description of Uncoupled Product Maps	28
2.4.2	Exponential Growth of Nonlinearities	28
2.4.3	Special Classes of Observables	30
2.4.4	Large Deviations of Spatially Averaged Observables for Product Maps	31
3	Heterogeneously Coupled Maps	33
3.1	Setting and Statement of the Main Theorems	34
3.2	Sketch of the Proof and the Use of a Truncated System	38
3.2.1	A Trivial Example Exhibiting Main Features of HCM	38

3.2.2	Truncated System	40
3.2.3	Steps of the Proof and Challenges	43
3.3	Proof of Theorem A when all Reduced Maps are Uniformly Expanding	44
3.3.1	Global Expansion of $\mathbf{F}_\varepsilon$	45
3.3.2	Distortion of $\mathbf{F}_\varepsilon$	47
3.3.3	Invariant Cones of Densities	50
3.3.4	Proof of Theorem A in the Expanding Case	51
3.4	Proof of Theorem A when some Reduced Maps have Hyperbolic Attractors	52
3.4.1	Strategy of the proof of Theorem A in the Presence of Hyperbolic Attractors	54
3.4.2	Invariant Cone Fields for $\mathbf{F}_\varepsilon$	55
3.4.3	Admissible Manifolds for $\mathbf{F}_\varepsilon$	58
3.4.4	Evolution of Densities on the Admissible Manifolds for $\mathbf{F}_\varepsilon$	60
3.4.5	Invariant Cone of Densities on Admissible Manifolds for $\mathbf{F}_\varepsilon$	62
3.4.6	Jacobian of the Holonomy Map along Stable Leaves of $\mathbf{F}_\varepsilon$	64
3.4.7	Proof of Proposition 3.5	67
3.4.8	Proof of Theorem A	70
3.4.9	Mather's Trick and proof of Theorem A when $n \neq 1$	73
3.4.10	Persistence of the Result Under Perturbations	74
3.5	An Excursion into One-Dimensional Dynamics: Proof of Theorem B	75
3.6	Homogeneous vs Heterogeneous Networks: Proof of Theorem C	78
4	Coupled Maps with Nonzero Distortion	85
4.1	Sparsely Coupled High-Dimensional Systems with Nonzero Distortion	87
4.1.1	Condition for Uniform Expansion of $\mathbf{F}$	88
4.1.2	Estimates on the Entries of $\mathbf{D}_x \mathbf{F}^{-1}$	89
4.1.3	Invariance of Cones of Functions Regular along each Coordinate	91
4.1.4	Decoupling of Nodes from the Network	94
4.1.5	One-Dimensional Marginals along Coordinate Directions	97
4.1.6	One-Dimensional Disintegration along Coordinate Directions	99
4.1.7	Spatial Large Deviations	103
4.2	HCM with Local Dynamics having Nonzero Distortion	103
4.2.1	Global Expansion of $\mathbf{F}_\varepsilon$	104
4.2.2	Distortion of $\mathbf{F}_\varepsilon$	104
4.2.3	Invariant Cones of Functions Regular along Each Coordinate	109
4.2.4	One-Dimensional Disintegration Along the Coordinates of the Low Degree Nodes	112
5	Conclusions and Further Developments	117
5.1	Conclusions	117
5.2	Open problems and new research directions	118

A	Appendix To Chapter 3	121
A.1	Estimates on the Truncated System	121
A.2	Estimate on Ratios of Determinants	124
A.3	Graph Transform: Some Explicit Estimates	126
A.4	Random Graphs	130
B	Appendix To Chapter 4	135
B.1	Estimates on Entries of $(\mathbf{D}_{\mathbf{x}}\mathbf{F}^{-1})_{ij}$	135
B.2	Lipschitz Constants of the Entries of $\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon} - \mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}$	140
B.3	Distances of Preimages of Nearby Functions	142
B.4	Derivatives Manipulations	143
C	Transfer Operator	145
D	Disintegration of Measures	147
	Bibliography	149

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 1.1 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 synchronise while other neurons remain out of synchrony. This particular behaviour shapes the network dynamics towards a healthy state [BGP⁺09]. Surprisingly, disrupting synchronisation 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 [MMAN13], 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 synchronisation is observed to be unstable [NMLH03]. However, in certain situations hubs can undergo a transition to collective dynamics [GGMA07, Per10, BRS⁺12]. Despite the large amount of recent work, a mathematical understanding of dynamical properties of such networks remains elusive.

In this thesis, we are going to discuss ergodic properties of high dimensional system and introduce a new way of describing Heterogeneously Coupled Maps (referred to as HCM in 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. The description we provide reduces the number of the degrees of freedom necessary to describe the evolution of the hubs' dynamics, at the expense of an arbitrary 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, we can describe the hubs with 1% accuracy in networks with 10^6 nodes for a time up to roughly $T = e^{40}$ for a set of initial conditions of measure roughly $1 - e^{-40}$. 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 synchronisation naturally emerges between hub nodes. This synchronisation is not due to a direct mutual interaction between hubs (as in the usual “Huygens” synchronisation) but results from the common environment that the hub nodes experience.

1.1 Emergent Dynamics on Heterogeneously Coupled Maps (HCM).

Figure 1.1 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 these three types constitute 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 long time in the size of the network, and it is robust under small perturbations.

Synchronisation due to the 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 synchronisation due to mutual interaction, i.e. different from “Huygens” synchronisation. In HCM, hubs display highly correlated behaviour even in the absence of direct connections between themselves. The poorly connected layer consisting of a huge number of weakly connected nodes plays the role of a “*heat bath*” providing a common forcing to the hubs which is responsible for the emergence of

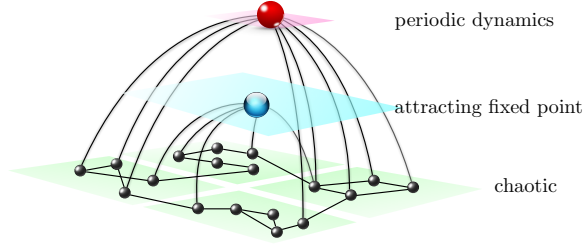


Figure 1.1: The dynamics across connectivity layers change depending on the connectivity of the hubs. We will exhibit an example where the hubs with the highest number of connections (in red, at the top) have periodic dynamics. In the second connectivity layer, where hubs have half of the number of connection (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).

coherence.

1.2 Low Dimensional Reduction and Hub Synchronisation

The Model. A network of coupled dynamical systems can be obtained from the datum (G, f, h, α) , where G is a labelled graph of the set of nodes $\mathcal{N} = [N]$, $f : \mathbb{T} \rightarrow \mathbb{T}$ is the local dynamics at each node of the graph, $h : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}$ 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 $\sigma > 1$. Nonlinear dynamics is found in some applications, [Izh07, WAH88, SSTC01], and Bernoulli maps are a sufficiently easy case that already produces very interesting behaviour. The graph G can be represented by its adjacency matrix $\mathbf{A} = (A_{in})_{1 \leq i, n \leq N}$ 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 and pointing at i . $A_{in} = 0$ otherwise. The degree $d_i := \sum_{n=1}^N A_{in}$ is the number of incoming edges at i . For sake of simplicity, in this introductory section we consider undirected graphs (A is symmetric), unless otherwise specified, but our results hold in greater generality (see Section 3.1).

The dynamics on the network is described by

$$z_i(t+1) = f(z_i(t)) + \frac{\alpha}{\Delta} \sum_{n=1}^N A_{in} h(z_i(t), z_n(t)) \bmod 1, \quad \text{for } i \in [N]. \quad (1.1)$$

In the above equations, Δ is a structural parameter of the network equal to the maximum degree. Rescaling the coupling strength in (1.1) dividing by Δ allows

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 we have that the degree d_i of the nodes $1, \dots, L$ are much smaller than the incoming degrees of nodes $L + 1, \dots, M$. A prototypical sequence of heterogeneous degrees is

$$\mathbf{d}(N) = (\underbrace{d, \dots, d}_L, \underbrace{\kappa_m \Delta, \dots, \kappa_m \Delta}_{M_m}, \dots, \underbrace{\kappa_2 \Delta, \dots, \kappa_2 \Delta}_{M_2}, \underbrace{\Delta, \dots, \Delta}_{M_1}). \quad (1.2)$$

with $\kappa_m < \dots < \kappa_2 < 1$ fixed and d/Δ small when N is large, then we will refer to blocks of nodes corresponding to $(\kappa_i \Delta, \dots, \kappa_i \Delta)$ as the i -th *connectivity layer* of the network, and to a graph G having sequence of degrees prescribed by Eq. (1.2) as a *layered heterogeneous graph*. (We will make all this more precise below.)

It is a consequence of structural and statistical stability of uniformly expanding maps, that for very small coupling strengths, the network dynamics will remain chaotic, and topologically and statistically similar to the uncoupled maps. 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.

Low-Dimensional Approximation for the Hubs. Given a hub $i_j \in \mathcal{N}$ in the i -th connectivity layer, our result give 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 $z_1, \dots, z_N \in \mathbb{T}$ be the state of each node, and assume that this collection of 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 (1.1) is a *mean field* (Monte-Carlo) approximation of the corresponding integral:

$$\frac{\alpha}{\Delta} \sum_{n=1}^N A_{i_j n} h(z_{i_j}, z_n) \approx \alpha \kappa_i \int h(z_{i_j}, y) dm(y) \quad (1.3)$$

where d_{i_j} is the incoming degree at i_j and $\kappa_i := d_{i_j}/\Delta$ is its normalized incoming degree¹. The parameter κ_i determines the effective coupling strength. Hence, the right hand side of expression (1.1) at the node i_j is approximately equal to the

¹The factor κ_i takes into account the fact that only a fraction κ_i of the terms in the sum is nonzero.

reduced map

$$g_{i_j}(z_{i_j}) := f(z_{i_j}) + \alpha \kappa_i \int h(z_{i_j}, y) dm(y), \quad (1.4)$$

Equations (1.3) and (1.4) 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 3.1. 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, e.g. only a logarithmic function of N . For these networks, our theorem implies

Theorem A (Informal Statement in Special Case). *Consider the dynamics (1.1) on a layered heterogeneous graph. If the degrees of the hubs are sufficiently large, i.e. $\Delta = O(N^{1/2+\varepsilon})$, 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 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 Synchronisation 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 with respect to the C^1 norm. Moreover fluctuations are correlated, and with probability one, $\xi(t)$ will be large for some $t > T$.

Idea of the Proof. The proof 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 $f(z) = 2z \bmod 1$ and the following simple coupling function:

$$h(z_i, z_n) = -\sin 2\pi z_i + \sin 2\pi z_n. \quad (1.5)$$

Since $\int_0^1 \sin(2\pi y) dy = 0$, the reduced equation, see Eq. (1.4), becomes

$$g_j(z_j) = T_{\alpha\kappa_j}(z_j) \text{ where } T_\beta(z) = 2z - \beta \sin(2\pi z) \bmod 1. \quad (1.6)$$

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 $\beta \in I_p := (3/2\pi, 4/2\pi]$ it has an attracting periodic orbit of period two. In fact, it follows from a recent result in [RGvS15] that the set of parameters β for which T_β is hyperbolic, as specified by Definition 3.1 below, is open and dense. (See Proposition 3.12 and 3.13 in the Appendix for a rigorous treatment). Figure 1.2 shows the graphs and bifurcation diagram of T_β 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 $\kappa_2 = 1/2$. 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,

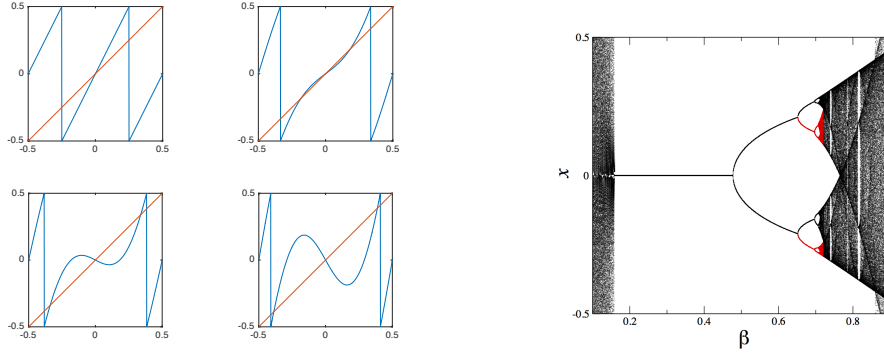


Figure 1.2: On the left the graphs of T_β for $\beta = 0, 0.2, 0.4, 0.6$. On the right the bifurcation diagram for the reduced dynamics of hubs. We considered the identification $\mathbb{T} = [-1/2, 1/2]/\sim$. We obtained the diagram numerically. To build the bifurcation diagram we reported a segment of a typical orbit of length 10^3 , for a collection of values of the parameter β .

and so for the three levels $j = 1, 2, 3$ we predict different types of dynamics looking at the bifurcation diagram. The predictions are summarised in Table 1.1.

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

Table 1.1: Dynamics across connectivity scales

1.3.2 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 equation (1.2) above, with $N = 10^5$, $\Delta = 500$, $w = 20$, $m = 2$, $M_1 = M_2 = 20$, $\kappa = 1$ and $\kappa_2 = 1/2$. The layer with highest connectivity is made of 20 hubs connected to 500 nodes, and the second layer is made of 20 hubs connected to 250 nodes. The local dynamics is again given by $f(z) = 2z \mod 1$, the coupling as in Eq. (1.5). We fixed the coupling strength at $\alpha = 0.6$ as in Section 1.3.1 so that Table 1.1 summarises the theoretically predicted dynamical behaviour for the two layers. We choose initial conditions for each of the N nodes independently and according to the Lebesgue measure. Then we evolve 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 1.3. 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 1.3 is in excellent agreement with the predictions of Table 1.1.

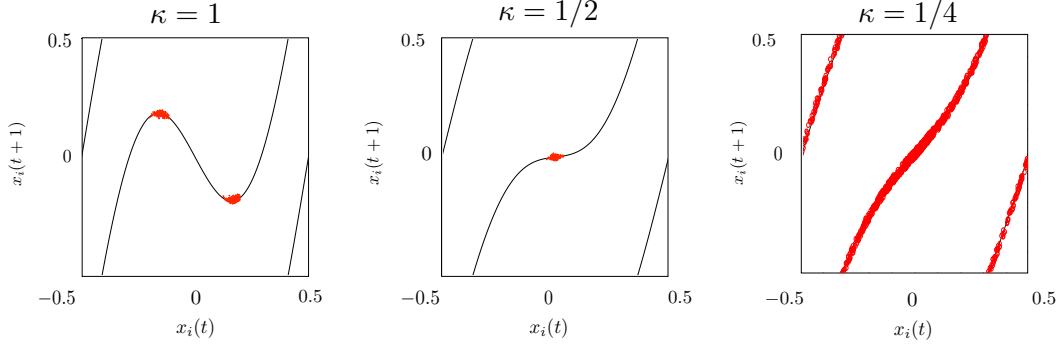


Figure 1.3: Simulation results of the dynamics of a layered graph with two layers of hubs. We plot the return maps $z_i(t) \times z_i(t+1)$. The solid line is the low dimensional approximation of the hub dynamics given by Eq. (1.6). The red circles are points taken from the hub time-series. In the first 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 stay near a fixed point, and in the third layer ($\kappa = 1/4$) the dynamics is still uniformly expanding.

1.4 Impact of the Network Structure: Homogeneous versus Heterogeneous Networks

The importance of network structure in shaping the dynamics has been highlighted by many studies [GS06, AADF11, NRS16] where it was shown that network topology and its symmetries shape bifurcations patterns and synchronisation spaces. Here we continue with this philosophy and exhibit the dynamical features that are to be expected in HCM. In particular, fixing the local dynamics and the coupling, the network structure dictates the resulting dynamics. In fact we show that

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

In Subsection 3.1 the content of 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 Eq. (1.2) with $M := \sum_{k=1}^m M_k$, each M_i being the number of nodes in the i -th connectivity layer. Assume

$$\Delta = O(N^{1/2+\varepsilon}), M = O(\log N) \text{ and } d = O(\log N) \quad (1.7)$$

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 Eq. (1.5).

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 so that for any $\alpha \in I$, the reduced dynamics $T_{\alpha\kappa_i}$ (Eq. (1.6)) has a periodic attractor. In fact it has at most two, $\{\bar{z}(t)\}_{t=1}^p$ and $\{-\bar{z}(t)\}_{t=1}^p$, and there is $s \in \{\pm 1\}$*

$$\text{dist}(z_{i_j}(t), s\bar{z}(t \bmod p)) \leq \xi$$

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 $T_0 \ll T_1$ 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 homogenous ones. 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 Section 2.1.2 for more details). We choose $p > \log N/N$ so that in the limit that $N \rightarrow \infty$ almost every random graph is connected (see [Bol01]).

Diffusive Coupling Functions The coupling functions satisfying

$$h(z_i, z_j) = -h(z_j, z_i)$$

are called *diffusive*. For each network G , we consider the corresponding system of

coupled maps defined by (1.1). In this case the subset

$$\mathcal{S} := \{(z_1, \dots, z_N) \in \mathbb{T}^{\mathbb{N}} : z_1 = z_2 = \dots = z_N\} \quad (1.8)$$

is invariant. $\mathcal{S}$ is called the *synchronisation 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 synchronisation depending on whether the graph is homogeneous or heterogeneous.

Theorem C (Informal Statement).

- a) Take any 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 (1.2) and (1.7), $\mathcal{S}$ is unstable.

Example 1.1. Take $f(z) = 2z \mod 1$ 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 synchronisation manifold for some values of the coupling strength ($\alpha \sim 0.3$) while any sufficiently large layered heterogeneous graph do not have any stable synchronised orbit. However, in a layered graph G the reduced dynamics for a hub node in the i -th layer is

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

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 orbit of the hubs in the layer follow this orbit (modulo small fluctuations) exhibiting coherent behaviour.

1.5 Coupled Maps with Nonzero Distortion

The above results consider distortionless local maps and the proof of Theorem A heavily relies on this property. Under this hypothesis it is possible to find uniform global bounds on the invariant densities and use these estimates to upper bound the measure of the set where fluctuations of the mean fields exceed a fixed threshold. Whenever the local maps have nonzero distortion, already in the uncoupled case, the invariant density has supremum exponentially large in the number of nodes of the network. A bound on this supremum alone is not enough to obtain significant information on the fluctuations of the mean field.

To treat the case with nonzero distortion we look for a more precise description of the invariant density and how the coupling influences the product structure. In fact, for uncoupled maps, the coordinates evolve independently and the invariant density is the product of factors each depending on a single coordinate. All the information on the mean field can be retrieved from knowledge of the single factors and large deviation results for independent random variables. In the presence of coupling, the product structure is lost and coordinates become correlated. To show that a large deviation estimate still holds, we study the disintegration of the invariant measure for the coupled system (or a modification of it) along one-dimensional leaves from the foliation obtained fixing all the coordinates but one corresponding to a low degree node. We show that the regularity of the disintegrated measure is controlled and, with the help of a decoupling argument already used in [KL06], we show that such disintegrations are close to the a.c.i.p. measure for the local uncoupled map. This allows to recover the claim in Theorem A when local maps have nonzero distortion in the case where all the low-dimensional approximations of the hub nodes are uniformly expanding maps.

1.6 Guide to the Thesis

In Chapter 2 we give some basic definitions, we review relevant literature on dynamical systems coupled on networks, and we start investigating high dimensional systems. In particular, we highlight in a very simple case the difficulties that arise when dealing with dimension dependent perturbations. In Chapter 3, we study HCM with distortionless local dynamics. Sections 3.3 and 3.4 contain the proof of the main results, while in Section 3.6 we compare dynamics on heterogeneous and homogeneous networks. In Chapter 4 we consider further developments to the case of local maps with nonzero distortion. We achieve this giving a thorough description of the product structure of the invariant measure of coupled systems. We

completely analyse the case of sparsely coupled maps, and we consider a particular case of HCM. Finally in Chapter 5 we remark the conclusions and list some open questions. The appendices are devised to support chapters 3 and 4 with the computations involved, and to give a basic background on some of the standard tools and techniques used throughout the thesis.

A final disclaimer is due before the main content of the thesis. A problem faced when dealing with the rigorous description of complex systems is the degree of generality in the assumptions. For example, in our case, one can consider different classes of local maps, couplings, and heterogeneous networks. For smaller classes one can obtain more precise results (in terms of quantitative estimates), but of course these are less general. The class we have chosen for the heterogeneous networks assumes very little on their structure. In fact, there are only 4 parameters describing it: number of low degree nodes, L , and their maximum degree, δ , number of hub nodes, M , and their maximum degree, Δ . With this choice we have decided to keep in our estimates only the explicit dependence on these parameters and drop the others (coming from local maps f and coupling h) which are indicated as factors uniform on L , δ , M and Δ . However, we also tried to keep the proofs as general as possible to provide a general strategy easy to adapt to more specific cases and potentially able to give explicit estimates also in other parameters of the system.

Network Dynamical Systems

Contents

2.1	Graphs	14
2.1.1	Homogeneous and Heterogeneous Graphs	15
2.1.2	Random Graphs	16
2.2	Dynamics on Networks	18
2.2.1	Network Maps and Network Vector Fields	18
2.2.2	Important Features of Dynamics on Networks	19
2.2.3	Maps with Additive Coupling Structure	20
2.3	Literature Review	20
2.3.1	Globally Coupled Systems and the Kuramoto Model	21
2.3.2	Coupled Maps Lattices and Networks	24
2.3.3	Bifurcation and Symmetries in Network Dynamical Systems	25
2.3.4	Necessity of a New Approach to Deal With HCM	26
2.4	High-Dimensional Dynamical Systems	27
2.4.1	Statistical Description of Uncoupled Product Maps	28
2.4.2	Exponential Growth of Nonlinearities	28
2.4.3	Special Classes of Observables	30
2.4.4	Large Deviations of Spatially Averaged Observables for Product Maps	31

In this chapter we define the main object under study: dynamical systems coupled on networks. We first revise notions from graph theory and random graphs theory that will be useful in the following. We then define what we mean by dynamical systems coupled on networks, and review the relevant literature on the topic.

2.1 Graphs

Definition 2.1 (Graph). A *directed graph* of size N is a couple $(\mathcal{N}, \mathcal{E})$ where $\mathcal{N}$ is a set of labeled nodes and $\mathcal{E} \subset \mathcal{N}^2 \setminus \{(i, i) : i \in \mathcal{N}\}$ ¹ is the set of edges connecting the nodes. One can also define an infinite graph in the same way, by choosing as set of nodes $\mathbb{N}$ or any countable² set $\mathcal{N}$.

From now on, for graphs of size $N \in \mathbb{N}$, we are going to consider $\mathcal{N} = [N] = \{1, \dots, N\}$, and call $\mathcal{G}_N$ the *set of all possible directed graphs on N vertices*.

A graph is *undirected* if whenever $(i, j) \in \mathcal{E}$ then also $(j, i) \in \mathcal{E}$ ³. The *adjacency matrix* of a graph is the matrix whose entries encode the presence of an edge of the network. In particular for a graph $([N], \mathcal{E})$ of size N , this is the $N \times N$ matrix $\mathbf{A} = (A_{ij})_{i,j \in [N]}$ satisfying

$$A_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases} \quad (2.1)$$

Notice that the adjacency matrix of an undirected graph is symmetric.

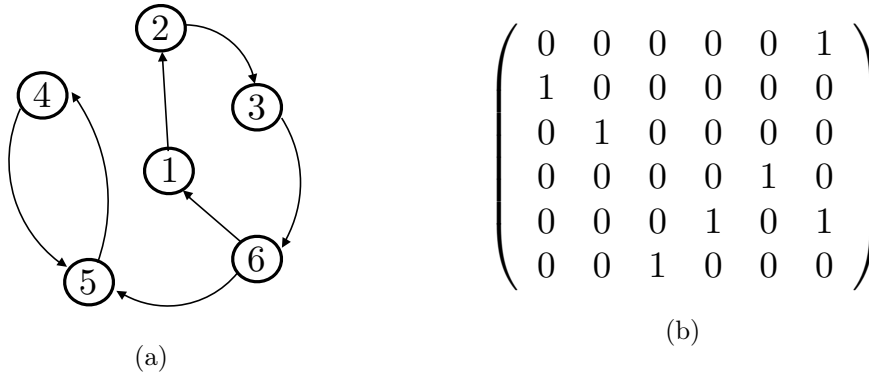


Figure 2.1: An example of a directed graph (a), and its associated adjacency matrix (b). The graph has $N = 6$ nodes and set of edges $\mathcal{E} = \{(2, 1), (1, 6), (3, 2), (6, 3), (5, 6), (5, 4), (4, 5)\}$.

In a graph, the degree of a node is defined as the number of edges concurring at that node. For a directed graph one can distinguish outgoing and ingoing degree

¹We eliminate the self-loops from the set of edges for convenience. As it will be clear later, this does not change much to our ends since from the dynamical perspective the contribution of a self-loop to the equations can always be included in the local dynamics.

²One can also consider graphs on “uncountable” sets of nodes. Loosely speaking, these are defined as functions from the unit square $[0, 1]^2$ to $\{0, 1\}$, and are sometimes called *graph limits* because they arise as continuous limits of finite graphs with size going to infinity. See for example [Lov12].

³Alternatively an undirected graph of size N is a couple $([N], \mathcal{E})$ where $\mathcal{E} \subset \{\{i, j\} : i \neq j \in [N]\}$.

depending whether the number of outgoing edges or incoming edges is counted. In formulae

$$d_{I,i} := \sum_{j=1}^N A_{ij}, \quad d_{O,i} := \sum_{j=1}^N A_{ji}.$$

The nodes of a graph can then be classified according to their degrees. A graph is often loosely called *heterogeneous* whenever the degrees of the nodes vary greatly across the network. In contrast, it is called *homogeneous* whenever the nodes in the graph have all approximately the same degree.

2.1.1 Homogeneous and Heterogeneous Graphs

In general we can qualify graphs according to the degrees of their nodes. In particular, we are interested in the distinction between *homogeneous* and *heterogeneous* ones. The former have all nodes with approximately the same degree, and fluctuations of the number of connections are small compared to the size of the graph (or some other relevant parameter). The latter exhibit very different connectivity⁴ across the graph.

Homogeneous Graphs Among the homogeneous graphs, we list finite and infinite *lattices* (see Fig. 2.2). In these graphs the nodes are bijectively associated to a subset of $\mathbb{Z}^d$, $d \in \mathbb{N}$, and the edges (directed or undirected) are between neighbouring nodes (first neighbours, second neighbours, ...). The lattice could be also periodic, see for example the *ring* graph in Fig. 2.2c.

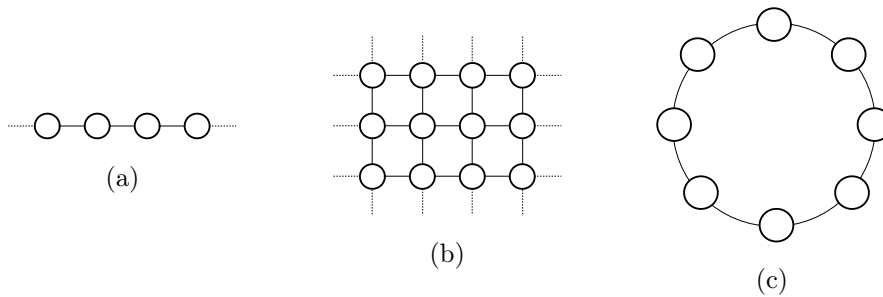


Figure 2.2: Linear lattice (a), 2D lattice (b), and ring (c). A ring is a finite one-dimensional lattice with periodic boundaries.

Another type of homogeneous graphs are the *all-to-all* graphs where all possible edges are present. See Fig. 2.3.

⁴number of connections.

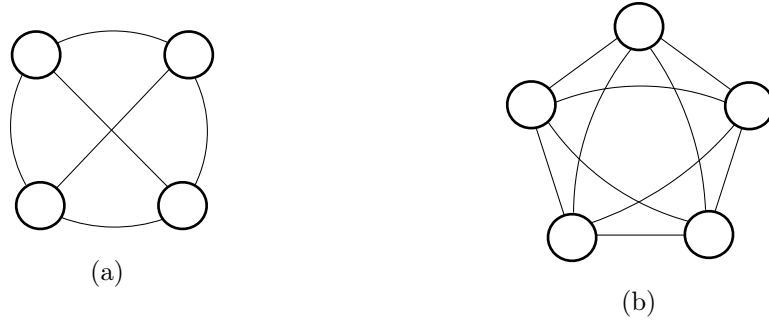


Figure 2.3: All-to-all graphs with 4 (a), and 5 (b) nodes.

Heterogeneous Graphs The simplest kind of heterogeneous graph is the *star network* where the only present edges are between a unique hub node and every other node in the network. A generalisation of the star network, are *layered graphs* where nodes can be divided into subsets of nodes sharing the same degree, which are also called *connectivity layers* and most of the nodes belong to the layer with the smallest degree.

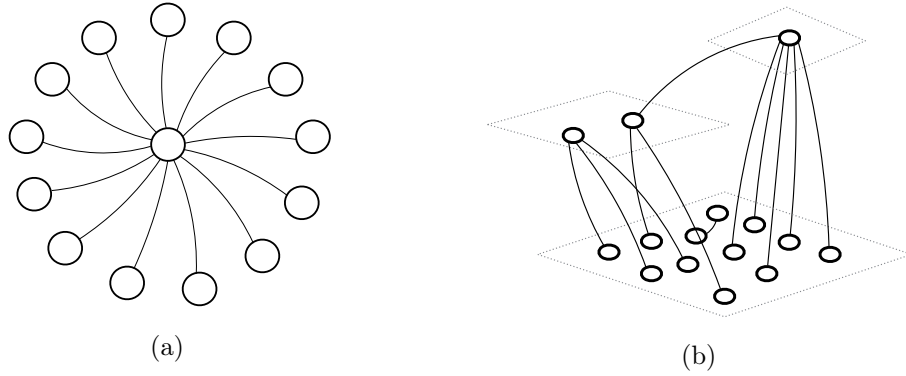


Figure 2.4: Star network (a), and layered graph (b). In the layered graph one can distinguish three sets of nodes with three different connectivity scales.

2.1.2 Random Graphs

Definition 2.2 ([Bol01]). A random graph of size $N \in \mathbb{N}$ is a probability measure $\mathbb{P}$ on $\mathcal{G}_N$ with the algebra of subsets.

There are various random variables of interest associated to a random graph. Namely, every quantity or object associated to a graph becomes a random variable for a random graph. For example, the adjacency matrix of a graph corresponds to a random matrix $A : \mathcal{G}_N \rightarrow \mathcal{M}(N, N)$ for a random graph $(\mathcal{G}_N, \mathbb{P})$. Notice that giving a random matrix supported on the N by N square matrices with entries either zero

or one corresponds to giving a random graph. Random graphs are usually defined prescribing a procedure (also called a *model*) on how to assign a probability to a given graph. Very often such procedures assign probabilities to the presence or absence of edges.

Definition 2.3. [Asymptotic Property] A sequence $\{Q_N\}_{N \in \mathbb{N}}$, with $Q_N \subset \mathcal{G}_N$, is an asymptotic property of the sequence of random graphs $\{(\mathcal{G}_N, \mathbb{P}_N)\}_{N \in \mathbb{N}}$ if

$$\lim_{N \rightarrow 1} \mathbb{P}_N(Q_N) = 1.$$

Erdős-Renyi Random Graphs One of the best known model for random graphs is the Erdős-Renyi model. This assigns an independent and identical probability $p \in [0, 1]$ to the presence of every edge from $[N]^2 \setminus \{(i, i) : i \in \mathcal{N}\}$. We call $(\mathcal{G}_N, \mathbb{P}_p^{ER})$ such random graph parametrised by $p \in [0, 1]$. The entries of the adjacency matrix of $(\mathcal{G}_N, \mathbb{P}_p^{ER})$ are independent identically distributed random variables taking value 1 with probability p and 0 with probability $1 - p$. Erdős-Renyi random graphs are prototypes of homogeneous random graphs, in fact for every $i \in [N]$ the degree⁵

$$d_i = \sum_{j \in [N]} A_{ij}$$

of every node $i \in [N]$ is identically and independently distributed with mean value pN .

Chung-Lu model [CL06] This model generalise the above case to allow for inhomogeneous distribution of the degrees across the graph. Take a vector

$$\mathbf{w} = (w_1, \dots, w_n) \in (\mathbb{R}^+)^n, \quad w_n \geq w_{n-1} \geq \dots \geq w_1, \quad (2.2)$$

which has been ordered just for later convenience. Suppose $\mathbf{w}$ satisfies

$$w_n^2 \rho \leq 1 \quad \text{with} \quad \rho := \frac{1}{\sum_{k=1}^n w_k}. \quad (2.3)$$

Then $(\mathcal{G}_N, \mathbb{P}_{\mathbf{w}}^{CL})$ is the associated random graph on N vertices whose adjacency matrix A has independent random entries such that

$$\mathbb{P}(A_{k\ell} = 1) = p_{k\ell} := w_k w_\ell \rho. \quad (2.4)$$

Condition (2.3) must be added in order to ensure that $p_{k\ell} \leq 1$.

⁵Here we consider the in-degree, but the same holds for the out-degree.

2.2 Dynamics on Networks

Graphs and their generalisations⁶ are ubiquitously used to describe complex systems. These are often successfully modelled as a number of interacting “units”. By *unit* we mean a small portion of the overall complex system to which is recognised existence independent of the global system to which it belongs⁷. Example of units could be atoms, or cells, or elements of a power grid, and represent, to the ends of the model, an elementary building block of the overall system. The interactions can be often described by a graph whose nodes are bijectively associated to the units of the system and the presence of an edge among two nodes prescribe the presence of an interaction between the corresponding units. The *reductionist* approach to the study of complex systems then consist to obtain information on the global dynamics from the knowledge of the microscopic⁸ behaviour and local interactions of elementary constituents.

Sometimes the word *network* is used as synonym for graph, but often presumes the existence of a dynamical process (for the units or their connections) and carries a connotation of “interaction” among the nodes. At this point, it is important to stress the difference between dynamics *on* networks and dynamics *of* networks. The first one is the object of study of this thesis, and it means that to each of the node of the network is associated a coordinate that describe the state of the unit on that node and that evolves with time interacting with adjacent nodes on the graph. Dynamics of networks, instead, refers to the evolution of the graph itself, namely how the sets $[N]$ and $\mathcal{E}$ change with time. Even if the dynamics of network has profound relations with the dynamics on networks we will not consider it here⁹.

2.2.1 Network Maps and Network Vector Fields

Definition 2.4. Given $\{M_i\}_{i \in [N]}$ manifolds, let $\Omega = \prod_{i \in [N]} M_i$ be their cartesian product. A map $\mathbf{F} : \Omega \rightarrow \Omega$ is said to be a network map on the graph $G = ([N], \mathcal{E})$, if for every $\mathbf{x} = (x_1, \dots, x_N) \in \Omega$, $F(\mathbf{x}) = (F^{(1)}(x), \dots, F^{(N)}(x))$ satisfies

$$F^{(i)}(x) = F^{(i)}(x_i; x_{j_1}, \dots, x_{j_{d_{I,i}}}) \text{ with } (i, j_k) \in \mathcal{E}, \forall k \in [d_{I,i}]. \quad (2.5)$$

The above definition means that we can think to associate to each node $i \in [N]$ a

⁶Like coupled cell networks [RS15] or multilayer graphs [KAB⁺14].

⁷Sometimes units can be rather arbitrary and result from discretisation or coarse-graining approximations. This is the case in some FEM models, notably for fluids.

⁸A complex systems is very often thought as being made of an enormous number of units. Said that, “microscopic” refers to the dimension of the units with respect to the whole.

⁹Notable recent advances on how the state of the nodes influence the network structure can be found, for example, in [BF17].

coordinate $x_i \in M_i$. A network map then prescribe the evolution of each coordinate so that the evolution of x_i depends on x_i itself and on the coordinates associated to nodes having an outgoing edge pointing to i . One can define analogously a network vector field. For example, a vector field $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ and the associated ODE, is *admissible for the graph* $G = ([N], \mathcal{E})$ if

$$\dot{x}_i = f^{(i)}(x_i; x_{j_1}, \dots, x_{j_{d_{I,i}}}) \quad (2.6)$$

for all $i \in [N]$, where $k = d_{I,i}$ and $f^{(i)} : \mathbb{R} \times \mathbb{R}^{d_{I,i}} \rightarrow \mathbb{R}$ are the components of $\mathbf{f}$ and depend only on the coordinates $(x_{j_1}, \dots, x_{j_{d_{I,i}}}) \in \mathbb{R}^{d_{I,i}}$ where $(i, j_k) \in \mathcal{E}$ for all $k \in [d_{I,i}]$. One can also say that the ODE (2.6) is *coupled on the graph* G .

2.2.2 Important Features of Dynamics on Networks

Many experimental observations of real world systems and numerical simulations have drawn attention on some dynamical features of the dynamics on networks which appear to play a crucial role in the processes they are involved with.

Synchronisation A dynamical system on a network exhibit synchronisation if two or more coordinates associated to the nodes behave identically. With reference to Eq. (2.5), manifolds like

$$\mathcal{S} := \left\{ \mathbf{x} \in \prod_i M_i : x_{i_1} = x_{i_2} = \dots = x_{i_k} \right\}$$

for $k \in [2, N]$ and $\{i_1, \dots, i_k\} \subset [N]$ are called *synchronisation spaces*. Their invariance and stability is the object of study of many works in the literature (see next section for a literature review). Stability of the synchronisation manifolds is important in applications because it implies asymptotic convergence of orbits of the system to synchronised states.

Synchronisation is a feature which complex systems exhibit in many natural processes, for example in neuroscience. Synchronised states have been observed in neuronal networks during epileptic seizures [DWG⁺05], [US06], [JDCJ⁺13], and, even if their role in the pathology has not been clarified, the hypothesis that it aids the recovery of a healthy state has been put forward. In [BGP⁺09], synchronisation of the hub nodes in neuronal networks of the hippocampus has been associated to the correct development of the network itself.

Phase-Locking Two or more nodes of a network dynamical system are *phase-locked* (or *phase-related*) if their coordinate follow the same periodic orbit apart

from a phase shift. This notion is very close to the notion of synchronisation.

Incoherent Behaviour The term “*incoherence*” is very often used in contrast with the previously described features. Loosely speaking, two or more nodes are said to behave incoherently if their orbits do not exhibit any relation and appear to evolve “independently”. This concept can be made precise in various ways. For example, two nodes can be said to behave incoherently if the linear (temporal) correlation of their orbits is, in absolute value, below some fixed threshold. In case the global dynamical system, as defined in Eq. (2.5), admits a physical measure μ , one can alternatively say that two or more nodes behave incoherently if the marginal of μ on the coordinates of those nodes is approximately a product measure (where approximately means that the distance in a suitable metric of the marginal to some product measure is below a fixed threshold). Even if one could expect that incoherence is the prevalent behaviour exhibited by chaotic high-dimensional dynamical systems coupled on a network, proving it rigorously in the sense of one of the definitions above can be far from easy.

2.2.3 Maps with Additive Coupling Structure

Suppose that $\{M_i\}_{i \in [N]}$ is a collection of manifolds, $G = ([N], \mathcal{E})$ is a graph with adjacency matrix $\mathbf{A}$, $\{f_i\}_{i \in [N]}$ a collection of maps $f_i : M_i \rightarrow M_i$, and $\{H_{ij}\}_{i,j \in [N]}$ a collection of maps $H_{ij} : M_i \times M_j \rightarrow M_i$. If each manifold M_i has an additive structure, one can define the map $\mathbf{F} : \prod_{i=1}^N M_i \rightarrow \prod_{i=1}^N M_i$ as

$$F^{(i)}(x_1, \dots, x_N) := f_i(x_i) + \sum_{j=1}^N A_{ij} H_{ij}(x_i, x_j). \quad (2.7)$$

By construction, the above map is a network map on the graph G . The coupling maps H_{ij} are usually chosen among the diffusive coupling functions as defined in the following

Definition 2.5. A function $H : M_1 \times M_2 \rightarrow \mathbb{R}$ is diffusive, if

$$H(x, y) = -H(y, x).$$

2.3 Literature Review

The mathematical analysis of the dynamics on networks counts hundreds of papers encompassing a variety of approaches to the field, from the purely theoretical ones

to solely numerical investigations. Here we offer only a glimpse of this vast literature paying closer attention to those works which present rigorous results and are related to what is contained in this thesis, either for the questions investigated or the techniques involved. We thus recall the main lines of research on: the emergence of behaviour on networks of large size, the influence of the network structure and its symmetries on the dynamical behaviour, and the statistical description of hyperbolic dynamics coupled on networks. More complete surveys can be found in [PG14, Fer14].

2.3.1 Globally Coupled Systems and the Kuramoto Model

Globally weakly coupled maps and vector fields of oscillators have received great attention in the literature. *Weakly coupled* refers to the scaling of the coupling strength that decreases as the inverse of the system size. *Globally coupled* means that the dynamics is coupled on an all-to-all graph.

Kuramoto Model A notable series of papers stems from the work of Kuramoto [Kur84] on chemical oscillations. The literature on the topic is huge, and for a complete list of references see [Str00]. In this work the object of study is a system of ODE describing globally weakly coupled *phase*¹⁰ oscillators

$$\dot{\theta}_i = \omega_i + \frac{\alpha}{N} \sum_{j=1}^N H_i(\theta_j - \theta_i), \quad \forall i \in [N], \quad (2.8)$$

where H_i are odd functions which make the coupling diffusive (see Definition 2.5). The system above also comes up as reduced equations in networks of weakly coupled neurons ([Izh07]) where the unperturbed vector field describing the dynamics of an isolated neuron is supposed to have an attracting limit cycle and the interactions are weak and have additive structure. Most approaches to study systems as in Eq. (2.8) are of numerical nature. In some instances results are rigorous and scope the existence and stability of synchronised states. One of these instances, already analysed in [Kur84], is

$$\dot{\theta}_i = \omega_i + \frac{\alpha}{N} \sum_{j=1}^N \sin(2\pi\theta_j - 2\pi\theta_i), \quad \forall i \in [N], \quad (2.9)$$

when the local frequencies ω_i have a symmetric unimodal distribution (often considered Gaussian or Cauchy) around their mean value. One possible approach is

¹⁰Each coordinate belongs to $\mathbb{T}$.

to consider a complex order parameter defined by $re^{i2\pi\psi} := \frac{1}{N} \sum_{i=1}^N e^{i2\pi\theta_i}$. The nonnegative modulus r measures the phase coherence of the oscillators, for example, whenever $r \sim 0$, it means that the oscillators are approximately uniformly distributed over $\mathbb{T}$, while when $r = 1$ they are completely synchronised. One can rewrite Eq. (2.8) in terms of r and ψ to obtain

$$\dot{\theta}_i = \omega_i + \alpha r \sin(2\pi\psi - 2\pi\theta_i) \quad (2.10)$$

which already convey the idea that when the coupling strength α is sufficiently large θ_i evolves toward the average phase ψ . As also detailed in [Str00], Eq. (2.10) is then studied in the continuum limit when N tends to infinity. In this way, the system of N coupled ODEs is substituted by only one nonlinear partial integro-differential equation where the unknown is a time-dependent density function $\rho(\theta, t)$ that describes the continuous distribution of the phase oscillators on $\mathbb{T}$ at time t . This reduced equation is amenable to mathematical analysis and one can determine the stability of its solutions. As a result, one is able to find a critical value for the coupling strength, α_c , for which $r(t)$ tends to one when t tends to infinity.

The above rigorous description is not suitable to analyse most generalisations of Eq. (2.9). In fact, already adding a higher harmonic to the coupling function jeopardise the approach. Also the case of interactions with different coupling strengths, or with network structure which is not all-to-all has been investigated mainly with numerics.

Globally Coupled Chaotic Maps Globally coupled maps (GCM) have been studied rigorously and numerically in many papers in the literature, among others: [Kan90], [Kan95], [Jus95], [NK98], [BKZ09], [Kel00], [Fer14], [SB16], and [S  16]. The prototype map considered is $\mathbf{F} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ with $F = \Phi_\alpha \circ \mathbf{f}$, where $\mathbf{f}(x_1, \dots, x_N) = (f(x_1), \dots, f(x_N))$ is the cartesian product of N copies of a uniformly expanding map $f : \mathbb{T} \rightarrow \mathbb{T}$, $\Phi_\alpha : \mathbb{T}^N \rightarrow \mathbb{T}^N$ is a coupling map and $\alpha \in \mathbb{R}$ is a coupling strength parameter¹¹. The prototype Φ_α for the mean-field interaction is

$$x_i \mapsto (1 - \alpha)x_i + \frac{\alpha}{N} \sum_{j=1}^N h(x_j - x_i) \mod 1 \quad (2.11)$$

This type of dynamics first appeared in works by Kaneko ([Kan90] and [Kan95]), where the author observed a peculiar phenomenon which has not yet been completely understood. He observed that taking the local map f from the logistic family and globally coupling it on a finite network, even for small values of the parameter α ,

¹¹When $\alpha = 0$, Φ_α is the identity.

the statistical behaviour of the coupled coordinated looked quite correlated. In particular, when $\alpha \rightarrow 0$ one would expect correlations among coordinates to go to zero, and the meanfield

$$\frac{1}{N} \sum_{j=1}^N x_j - x_i \quad (2.12)$$

to satisfy the Law of Large Numbers and the Central Limit Theorem. Surprisingly, numerical simulation showed that even if the CLT was approximately satisfied (apart from the tail behaviour), for most parameters of the logistic family, the LLN did not hold in the following sense. If the coordinates $\{x_j\}_{j=1}^N$ are uncorrelated, one would expect the variance of the average in (2.12) to decay with N as $\sim N^{-1}$. However, sampling the mean field along an orbit, one observes that its mean square error¹² decay as $\sim N^{-\alpha}$ with $\alpha \in (0, 1)$. Kaneko called this phenomenon *violation of the law of large numbers*. Even more interestingly, adding some independent random noise to the equations improves the law of large numbers (increases α) suggesting that some dynamical *deterministic* phenomena is responsible for the observed behaviour. It was recognised that this behaviour is related to the different characteristics of maps drawn from the logistic family under change of parameter. Later results ([Jär97],[Kel00]) proved that the violation of the law of large number is not observed in systems where f is a hyperbolic, and that non-hyperbolic points are responsible for the emergence of different phases. However, coupled tent maps with slope greater than $\sqrt{2}$ exhibit the violation of the law of large numbers, and the singularities of the map have been recognised to be responsible for the violation ([NK98]).

In [Fer14] and [SB16], $f(x) = 2x \bmod 1$ and Φ_α is piecewise linear with $h : \mathbb{T} \rightarrow \mathbb{R}$ a map whose “lift” to the interval $[-1/2, 1/2]$, $H : [-1/2, 1/2] \rightarrow [-1/2, 1/2]$, is the identity on this interval. This choice of coupling makes Φ_α discontinuous at points $(x_1, \dots, x_n) \in \mathbb{T}^n$ with $|x_i - x_j| = 1/2$ for some $i, j \in [N]$. When $\alpha = 0$, the system has a unique a.c.i.p. measure. Some questions asked in this situation concern: (1) stochastic stability of the unique a.c.i.p. measure when $\alpha \rightarrow 0$; (2) number of ergodic components when $\alpha \neq 0$, but the system still presents expansion; (3) existence of a parameter $\alpha_c \in \mathbb{R}$ such that for $\alpha > \alpha_c$ the global synchronisation manifold is stable.

Question (2) has been completely answered for N up to 4 ([Sél16]). It has been established that, when α is sufficiently large, the presence of the singularities in the coupling function are responsible for the emergence of different ergodic components that support a.c.i.p. measures. The components have been classified for N up to 4. The procedure relies on the precise geometric understanding of the sets of

¹²Which is a statistical estimator of the variance.

singularities which makes it hard to generalise to higher dimensions.

Question (1) and (3) have been again considered in the continuum limit for $N \rightarrow \infty$ using a *self-consistent transfer operator*. This approach was already used by Kaneko in [Kan90]. (See [Bla11] for a general discussion). The main idea is the following. One can see the sum in Eq. (2.11) as the evaluation of the function $\bar{h}(y) = \alpha h(x_i - y)$ against the discrete probability distribution $\frac{1}{N} \sum_{j=1}^N \delta_{x_j}$. Generalising the above expression to a probability measure μ_0 for the distribution of the states, under suitable hypotheses, one obtains the well defined function

$$\Phi_{\alpha, \mu_0}(x) = (\text{Id} + \mu_0(\bar{h}))(x) = x + \int \alpha h(x - y) d\mu_0(y).$$

In turn, the push-forward of the measure μ_0 under the function $\Phi_{\alpha, \mu_0} \circ \mathbf{f}$ gives the evolution of the states to the new probability measure

$$\mu_1 := (\Phi_{\alpha, \mu_0} \circ f)_* \mu_0$$

and one can obtain the orbit $\{\mu_n\}_{n \in \mathbb{N}}$. In [SB16] the authors show that if α is sufficiently small and μ_0 is a probability measure with a C^1 density, then the densities of the $\{\mu_n\}_{n \in \mathbb{N}}$ converge exponentially fast to the constant uniform density. However, for larger α and any initial measure μ_0 with density supported on an interval of diameter less than $1/2$, μ_n converges to a point mass.

In [BKZ09] the authors consider local dynamics f sampled from a family of two branched uniformly expanding rational maps on an interval which are globally coupled via a meanfield interaction. A surprising difference between the finite system and its continuous limit is found. In the parameter range considered for the coupling strength, the finite system always exhibit a unique SRB measure and exponential decay of correlations. In the continuous limit, the self-consistent transfer operator undergoes a pitchfork bifurcation from a single stable fixed point to two stable fixed points and an unstable fixed point. It is conjectured that this result might be related to the violation of the law of large numbers observed by Kaneko.

2.3.2 Coupled Maps Lattices and Networks

Coupled Map Lattices (CML), in contrast with GCM, consider maps which are coupled to their nearest neighbours on a lattice $(\mathbb{Z}^d, d \geq 1)$. The most studied example is that of maps coupled to first nearest neighbours on the one-dimensional lattice $\mathbb{Z}$

$$x_i \mapsto (1 - \alpha)f(x_i) + \alpha(f(x_{i-1}) + f(x_{i+1})), \quad \forall i \in \mathbb{Z}, \quad \varepsilon \in [0, 1].$$

Again, among the first people to investigate these systems was Kaneko in [Kan84]. The study of CLM arises as a discrete alternative to PDEs for the study of spatially extended systems. The questions addressed were inspired by statistical mechanics and concerned with the existence of phase transitions for nonzero values of the coupling parameter α . Numerical studies on the emergence of coherent structures were the first to appear (see the book [KT11] and references therein).

Rigorous theoretical investigation started with [BS88] where symbolic dynamics¹³ and thermodynamic formalism were used to study uniformly expanding maps coupled on $\mathbb{Z}$ via some interaction function preserving the Markov structure. The study proved that for small values of the parameter α , no phase transition happens and there is only one equilibrium measure as in the uncoupled case.

Another early work is [BK95] where a cluster expansion approach is used to determine existence and uniqueness of an equilibrium measure, and properties of its finite-dimensional marginals for a CML of expanding analytic maps. Other works that use the same kind of technique comprise [BDEI⁺98], [FR00], [BR01], and [Rug02]. This approach is able to give a thorough description of the systems (in terms of the spectrum of the associated transfer operator) with the inconvenience that it requires analyticity of the maps.

In the series of works [KK92], [Sch04], [KL04], [KL05], [KL06] different functional spaces are considered for the study of the transfer operator of CML of maps satisfying Lakota-Yorke type of conditions relaxing a lot the strong regularity assumptions that were otherwise essential in other approaches.

Coupled Map Networks In [KY10], the authors consider local dynamics and coupling analogous to those described above, but on more general network structures. In particular, they find conditions on spectral properties of the graph (expressed in terms of the eigenvalues of the Laplacian matrix associated to a graph) and the local dynamics in order to have an hyperbolic splitting of the tangent space. This in turn rigorously shows the presence of bifurcations when the coupling strength is varied: increasing the magnitude of the coupling strength, unstable directions tend to become stable.

2.3.3 Bifurcation and Symmetries in Network Dynamical Systems

The last line of research we review is not directly implicated in the work of this thesis, but highlights the importance of network structure in shaping the dynamics of coupled maps and vector fields. In particular it studies how the symmetries (evident

¹³Symbolic dynamics will be also used in [Jus95] to study GCM.

and hidden) of a graph constraint the dynamics of a network dynamical system on it. For example, it relates stability of synchronisation spaces and bifurcation patterns to the existence of equivariant actions of symmetry groups.

The main object of studies are ODEs coupled on networks as in Eq. (2.6). The first results arise when looking at network dynamics under the light of theorems on symmetry of dynamical systems (see for example [Fie07] or [GS03]). This approach is summarised for example in [GS06] or [AADF11]. A symmetry of a graph $G = ([N], \mathcal{E})$ is usually defined as a couple of permutations

$$\sigma_1 : [N] \rightarrow [N] \quad \text{and} \quad \sigma_2 : \mathcal{E} \rightarrow \mathcal{E}$$

of nodes and edges respectively that preserve incidence, namely $\sigma_2(i, j) = (\sigma_1(i), \sigma_1(j)) \in \mathcal{E}$ for all $(i, j) \in \mathcal{E}$ ¹⁴. One can show that if $\mathbf{x}(t)$, $t \in I \subset \mathbb{R}$, is a solution of an ODE admissible for G , then the action of σ_1 on the coordinates of $\mathbb{R}^N$ sends $\mathbf{x}(t)$ to another solution of the ODE, $\sigma_1 \mathbf{x}(t)$. Results in [GS03] and [BG01] give necessary and sufficient conditions on the symmetry group of a network and a synchronisation space $\mathcal{S}$, for the existence of network vector fields such that $\mathcal{S}$ is invariant and stable.

2.3.4 Necessity of a New Approach to Deal With HCM

- The study of **Global Synchronization** [Kur84, BP02, EM14, PERV14] deals with 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 Eq. (1.8) 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 of a non uniform mean field across connectivity layers.
- **Bifurcation Theory** [GSBC98, GS06, KY10, AADF11, 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 center 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 Fig-

¹⁴ σ_2 is determined by σ_1 , but of course σ_2 might not exist in general. It always exists if either there are no edges ($\mathcal{E} = \emptyset$) or all of them are present ($\mathcal{E} = [N]^2 \setminus \{(i, i) : i \in [N]\}$).

ure 3.2). For these reasons it is difficult to frame HCM in this setting or use perturbative arguments.

- The statistical description of **Coupled Map Lattices** [Kan92, BS88, BR01, BDEI⁺98, KL05, KL06, KL04, CF05] 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, however, 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.

2.4 High-Dimensional Dynamical Systems

In this thesis we are going to treat dynamical systems coupled on networks of arbitrary, but large, size. Such systems are intrinsically high-dimensional. With this we mean that the phase space is given by a manifold with large dimension that cannot be reduced by any coordinate transformation. It is well known that the number of degrees of freedom influences drastically the dynamics. This is already evident in low dimensional examples. Given an endomorphism $f : M \rightarrow M$ of the n -dimensional manifold M , there is a huge difference in what one expects to be able to say about f whether the dimension is $n = 1$, $n = 2$ or higher. A well known example is density of hyperbolicity which has been recently proved, [KSvS07], whenever $M = \mathbb{S}^1$ or $M = [a, b]$, with $a, b \in \mathbb{R}$, and f is a C^k endomorphism, but which is generally not true when $\dim M \geq 2$.

Here we are not going to deal with such subtle questions, but rather refer to high-dimensional systems as systems where *the dimension of the phase space is a free parameter* and, although finite, can be arbitrarily large. In particular we are

going to deal with instances where the ergodic properties of systems (mixing rates, stochastic stability,...) depend on its size. Our focus is going to be on the case of uniformly expanding maps coupled on networks with additive interactions.

We consider uncoupled product maps and discuss the complications that arise when dealing with the statistical description of high-dimensional systems.

2.4.1 Statistical Description of Uncoupled Product Maps

Consider N uncoupled systems evolving independently one of the other. This is a ready and easy example to see which problems arise when studying high-dimensional systems.

Definition 2.6. Given a finite collection of maps $\{f_i\}_{i \in [N]}$, $f_i : M_i \rightarrow M_i$, the associated product map $\mathbf{F} : \prod_{i \in [N]} M_i \rightarrow \prod_{i \in [N]} M_i$ is the cartesian product of the maps from the collection so that for every $\mathbf{x} = (x_1, \dots, x_N) \in \prod_{i \in [N]} M_i$, $F^{(i)}(\mathbf{x}) = f_i(x_i)$.

The system is uncoupled in the sense that every coordinate evolve independently and can be treated separately. In the case that the f_i are uniformly hyperbolic maps, ergodic properties of $\mathbf{F}$ are easy to establish from knowledge of the properties of the single maps f_i . However, including the dependence on N and treating it as a parameter raises some non-trivial questions. As usual, we will focus our attention on uniformly expanding maps of the circle.

2.4.2 Exponential Growth of Nonlinearities

Consider a uniformly expanding map $f \in C^{1+\nu}(\mathbb{S}^1, \mathbb{S}^1)$. Then the product map on the N -torus, $\mathbb{T}^N$, of N identical copies of f ,

$$\mathbf{f} = \prod_{i \in [N]} f = f \times \dots \times f$$

is itself a uniformly expanding map of $C^{1+\nu}(\mathbb{T}^N, \mathbb{T}^N)$. It is well known (see e.g. [Via97] and references therein), that

Theorem 2.1 ([Via97]). *Any $f \in C^{1+\nu}(M, M)$ uniformly expanding¹⁵ self-map of the Riemannian manifold M*

- *has an absolutely continuous invariant measure $\mu \ll m_M$,*

¹⁵A self-map f of the Riemannian manifold M is uniformly expanding if there is $\sigma > 1$ such that $\forall(\mathbf{x}, \mathbf{v}) \in TM$, $\|D_{\mathbf{x}}f\mathbf{v}\|_{f(\mathbf{x})} \geq \sigma\|\mathbf{v}\|_{\mathbf{x}}$.

- the measure is mixing and there are $C > 0$ and $\lambda \in (0, 1)$ such that

$$\left| \int \varphi \cdot \psi \circ f^n d\mu - \int \varphi d\mu \cdot \int \psi d\mu \right| \leq C\lambda^n, \quad \forall n \geq 0 \quad (2.13)$$

and $\forall \varphi \in C^\nu(M, \mathbb{R})$ and $\psi \in L^\infty(M, \mathbb{R})$ with $\|\varphi\|_{C^\nu} = \|\psi\|_\infty = 1$,

- μ is stochastically stable under small perturbations of f in $C^{1+\nu}$.

The results above hold in particular for the map $\mathbf{f}$. However, the constant C in Eq. (2.13) will depend on N and become arbitrarily large for $N \rightarrow \infty$. The reason behind the behaviour above is intrinsic and unavoidable in high-dimensional systems as it is shown by the following observations.

Asymptotic Singularity of Product Measures Suppose that μ and ν are two equivalent probability measures, on some σ -algebra (Ω, Σ) , meaning that $\mu \ll \nu$ and $\nu \ll \mu$. Comparing the infinite powers of these measures, $\mu^{\mathbb{N}}$ and $\nu^{\mathbb{N}}$, one observe that if $\mu \neq \nu$, then $\mu^{\mathbb{N}}$ and $\nu^{\mathbb{N}}$ are singular with respect to each other. This has been remarked for example in [KL06], and can be easily seen from the following argument already presented in that paper. Suppose that $\psi \in L^\infty(\mathbb{T}, \mathbb{R})$, then it follows from the law of large numbers that there is a set $A \subset \mathbb{T}^{\mathbb{N}}$ of measure $\mu^{\otimes \mathbb{N}}(A) = 1$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(x_i) = \mathbb{E}_\mu[\psi], \quad \forall (x_i)_i \in A,$$

and there is $B \subset \mathbb{T}^{\mathbb{N}}$ with $\nu^{\otimes \mathbb{N}}(B) = 1$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(x_i) = \mathbb{E}_\nu[\psi], \quad \forall (x_i)_i \in B.$$

If $\mu \neq \nu$ then there is $\psi \in L^\infty(\mathbb{T}, \mathbb{R})$ for which $\mathbb{E}_\mu[\psi] \neq \mathbb{E}_\nu[\psi]$ implying that A and B are disjoint, and thus $\mu^{\otimes \mathbb{N}}$ and $\nu^{\otimes \mathbb{N}}$ are singular with respect to each other. This means that when the dimension is large, already fixing a reference measure with respect to which determine the physical measure(s) of the system can be a subtle issue.

Asymptotic Divergence of Distortion From the equality $\mathbb{T}^N = (\mathbb{S}^1)^N$ it seems natural to take on $\mathbb{T}^N$ the product Riemannian structure induced by the standard metric on $\mathbb{S}^1$. This gives the Lebesgue measure on $\mathbb{T}^N$ as reference measure, and the Euclidean distance as Riemannian distance on $\mathbb{T}^N$. The Jacobian at $(x_1, \dots, x_N)$

of the map $\mathbf{f}$ with respect to the Riemannian volume is then

$$J_{\mathbf{f}}(x_1, \dots, x_N) = \prod_{i \in [N]} J_f(x_i) = \prod_{i \in [N]} |D_{x_i} f|.$$

If the local map f has nonzero distortion, its Jacobian J_f is non constant, and the volume distortion under $\mathbf{f}$ at some points in the phase space can be as large as $\mathcal{O}(\sqrt{N})$. In fact,

$$\frac{J_{\mathbf{f}}(x_1, \dots, x_N)}{J_{\mathbf{f}}(\bar{x}_1, \dots, \bar{x}_N)} = \prod_{i \in [N]} \left| \frac{D_{x_i} f}{D_{\bar{x}_i} f} \right| = \prod_{i \in [N]} \left| 1 - \frac{D_{x_i} f - D_{\bar{x}_i} f}{D_{\bar{x}_i} f} \right| \geq \prod_{i \in [N]} |1 + cd(x_i, \bar{x}_i)|$$

For example, choosing for simplicity $f \in C^2(\mathbb{S}^1, \mathbb{S}^2)$, and $y \in \mathbb{S}^1$ such that $D_x^2 f > c > 0$, for $\bar{y} \in U$ a neighbourhood of x_0 then for $x_i = y$ and $\bar{x}_i = \bar{y}$:

$$\log \frac{J_{\mathbf{f}}(x_1, \dots, x_N)}{J_{\mathbf{f}}(\bar{x}_1, \dots, \bar{x}_N)} \geq cN^{1/2} d((x_1, \dots, x_N), (\bar{x}_1, \dots, \bar{x}_N)).$$

Stochastic Stability and Dimension Very much related to the previous observations is the interplay of statistical stability and the dimension of the system. This is already evident in the case of product maps. Take uniformly expanding $C^{1+\nu}$ maps f and f_ε with a.c.i.p. measures μ and μ_ε different from μ satisfying $\|f - f_\varepsilon\|_{C^1} \leq \varepsilon$. The N -fold products of these maps $\prod_{i \in [N]} f$ and $\prod_{i \in [N]} f_\varepsilon$ will have a.c.i.p. measures $\mu^{\otimes N}$ and $\mu_\varepsilon^{\otimes N}$. Fixing N and letting ε go to zero one has convergence of $\mu_\varepsilon^{\otimes N} \rightarrow \mu^{\otimes N}$ ¹⁶, however, we saw above that fixing ε and letting N tend to infinity, these measures become singular with respect to each other. Supposing, now that ε depends on the dimension, the order of ε with N determines the statistical behaviour of $\prod_{i \in [N]} f_{\varepsilon(N)}$ in relation to that of $\prod_{i \in [N]} f$ when $N \rightarrow \infty$.

2.4.3 Special Classes of Observables

Given a product map, or more generally a network dynamical systems, on the phase space $\prod_{i=1}^N M_i$, one can imagine it being the result of a modelling procedure where a real system is simplified as a set of distinct units, each of which having in principle independent identity. To every unit is associated one of N coordinates $x_i \in M_i$.

¹⁶In the sup norm, $\forall N \in \mathbb{N}$,

$$\lim_{\varepsilon \rightarrow 0} \sup_{\mathbf{x} \in \mathbb{T}^N} \left| \frac{d\mu^{\otimes N}}{dm_N}(\mathbf{x}) - \frac{d\mu_\varepsilon^{\otimes N}}{dm_N}(\mathbf{x}) \right| = 0.$$

See for example [Via97] for details.

There is than some reason in focusing one's attention to observables of the kind

$$\Psi_+(x_1, \dots, x_N) = \sum_{i=1}^N \psi_i(x_i) \quad (2.14)$$

for observable functions $\psi_i \in L^1(M_i, \mathbb{R})$. This is often the case for large systems, as in statistical physics, where one suppose not to have access to the microscopic components of the system and the object of study are averages of observables defined on single coordinates (like internal energy, temperature,...) over a large number of components. For example, given $\psi \in L^1(\mathbb{T}, \mathbb{R})$ the study of the *spatial average* of a given observable on the coordinates of the system

$$\bar{\psi} := \frac{1}{N} \sum_{j=1}^N \psi(x_j) \quad (2.15)$$

comes up very often in the description of mean field interactions, where the sum spans either all coordinates, as in Eq. (2.15), or a subset of them.

Another type of observables one can look at are of the kind

$$\Psi_\times(x_1, \dots, x_N) = \prod_{i=1}^N \psi_i(x_i). \quad (2.16)$$

This instead arise, for example, when considering the invariant measure of product systems. Both in Eq. (2.14) and Eq. (2.16), coordinates are uncoupled and this makes the analysis of the evolution of such observables easier than the more general case, plus some of the complications shown in the previous section do not arise in this case. This types of observable are going to play a prominent role in the rest of the thesis.

2.4.4 Large Deviations of Spatially Averaged Observables for Product Maps

Consider the uniformly expanding map $f \in C^2(\mathbb{S}^1, \mathbb{S}^1)$, and $\mathbf{f}$ the N -fold cartesian product of identical copies of f . Suppose that μ is the unique a.c.i.p. measure for f given by Theorem 2.1. Let $\psi \in L^\infty(\mathbb{T}, \mathbb{R})$. We want to find hitting time statistics to sets of the kind

$$\mathcal{B}_{\xi, N} := \left\{ \left| \frac{1}{N} \sum_{i=1}^N \psi(x_i) - \mathbb{E}_\mu[\psi] \right| \geq \xi \right\}. \quad (2.17)$$

for $\xi > 0$ and $N \in \mathbb{N}$ as N becomes large. We know that if μ_0 is the a.c.i.p. measure of f then $\mu_0^{\otimes N}$ is the a.c.i.p. for $\mathbf{f}_N$. Using the Hoeffding inequality on concentration of empirical averages of bounded independent random variables (see Appendix A), one can show that there exists a constant $C > 0$ such that

$$\mu_0^{\otimes N}(\mathcal{B}_{\xi,N}) \leq 2 \exp[-C\xi^2 N], \quad \forall N \in \mathbb{N}.$$

This means that for $N \rightarrow \infty$, fixed any precision ξ , the measure of the set for which the empirical average of ψ fall outside of the interval $(\mathbb{E}_\mu[\psi] - \xi, \mathbb{E}_\mu[\psi] + \xi)$ decreases exponentially fast with N .

From the invariance and ergodicity of the measure $\mu_0^{\otimes N}$ with respect to $\mathbf{f}$ follows that orbits will rarely visit $\mathcal{B}_{\xi,N}$, but, if $\xi < \text{Essup } \psi - \text{Essinf } \psi$, they will eventually do.

Measure of $\mathcal{B}_{\xi,N}$ after a small perturbation Suppose that $f_\varepsilon \in C^2(\mathbb{T}, \mathbb{T})$ is a perturbation of the map f with $d_{C^1}(f, f_\varepsilon) < \varepsilon$, and μ_ε its invariant measure. Then, statistical stability of uniformly expanding maps implies that for every $\varepsilon_1 > 0$ there is a $\varepsilon_2 > 0$ such that $\|\varphi_\varepsilon - \varphi_0\|_{C^0} < \varepsilon_1$ when $\varepsilon < \varepsilon_2$, where φ_ε and φ_0 are the densities of μ and μ_0 with respect to m . One can obtain an estimate of $\mu^{\otimes N}(\mathcal{B}_{\xi,N})$ in the following way.

$$\begin{aligned} \mu^{\otimes N}(\mathcal{B}_{\xi,N}) &= \int_{\mathcal{B}_{\xi,N}} \prod_{j=1}^N \varphi_\varepsilon(x_j) dm_N(\mathbf{x}) = \int_{\mathcal{B}_{\xi,N}} \prod_{j=1}^N \varphi_0(x_j) \frac{\varphi_\varepsilon(x_j)}{\varphi_0(x_j)} dm_N(\mathbf{x}) \\ &\leq (1 + \varepsilon_1)^N \int_{\mathcal{B}_{\xi,N}} \prod_{j=1}^N \varphi_0(x_j) dm_N(\mathbf{x}) \\ &\leq (1 + \varepsilon_1)^N \mu_0^{\otimes N}(\mathcal{B}_{\xi,N}) \\ &\leq 2 \exp \{N[\log(1 + \varepsilon_1) - C\xi^2]\}. \end{aligned}$$

After fixing ξ , for $\varepsilon_1 > 0$ sufficiently small, the measure of $\mathcal{B}_{\xi,N}$ with respect to the a.c.i.p. of the perturbed map still decreases exponentially fast with the dimension N . This means that bounds on large deviations from $\mathbb{E}_{\mu_0}[\psi]$ of the spatial average continue to hold even if in the limit for N that tends to infinity $\mu_0^{\otimes N}$ and $\mu_\varepsilon^{\otimes N}$ become singular with respect to each other.

Heterogeneously Coupled Maps

Contents

3.1	Setting and Statement of the Main Theorems	34
3.2	Sketch of the Proof and the Use of a Truncated System . .	38
3.2.1	A Trivial Example Exhibiting Main Features of HCM	38
3.2.2	Truncated System	40
3.2.3	Steps of the Proof and Challenges	43
3.3	Proof of Theorem A when all Reduced Maps are Uniformly Expanding	44
3.3.1	Global Expansion of $\mathbf{F}_\varepsilon$	45
3.3.2	Distortion of $\mathbf{F}_\varepsilon$	47
3.3.3	Invariant Cones of Densities	50
3.3.4	Proof of Theorem A in the Expanding Case	51
3.4	Proof of Theorem A when some Reduced Maps have Hyperbolic Attractors	52
3.4.1	Strategy of the proof of Theorem A in the Presence of Hyperbolic Attractors	54
3.4.2	Invariant Cone Fields for $\mathbf{F}_\varepsilon$	55
3.4.3	Admissible Manifolds for $\mathbf{F}_\varepsilon$	58
3.4.4	Evolution of Densities on the Admissible Manifolds for F_ε . .	60
3.4.5	Invariant Cone of Densities on Admissible Manifolds for $\mathbf{F}_\varepsilon$.	62
3.4.6	Jacobian of the Holonomy Map along Stable Leaves of $\mathbf{F}_\varepsilon$. .	64
3.4.7	Proof of Proposition 3.5	67
3.4.8	Proof of Theorem A	70
3.4.9	Mather's Trick and proof of Theorem A when $n \neq 1$	73
3.4.10	Persistence of the Result Under Perturbations	74
3.5	An Excursion into One-Dimensional Dynamics: Proof of Theorem B	75
3.6	Homogeneous vs Heterogeneous Networks: Proof of Theorem C	78

In this chapter we are going to precisely state and prove the main results concerning HCM that we have already extensively presented in the introduction to the thesis. In Section 3.1 we give the setting and the statements of the main theorems. In Section 3.2 we give a heuristical justification of the results and provide the main ideas of the proofs. Section 3.3 is devoted to proving one of the main theorems in a particular case which is both instructive and instrumental for the general statement. In sections number 3.4, 3.5 and 3.6 we complete the proofs of the three main theorems. The content of this chapter is the subject of the paper [PvST17].

3.1 Setting and Statement of the Main Theorems

Take a N -dimensional heterogeneous graph $G = ([N], \mathcal{E})$ and let $\mathbf{A}$ be its adjacency matrix. In this section we will be only concerned with the in-degree of a node which counts the contributions to the interaction felt by that node. Suppose that the nodes are labeled so that their degrees are ordered from the smallest to the largest. The graph is heterogeneous in a sense that we later specify, but one should think that the in-degrees $d_1, \dots, d_L$ of the nodes $\{1, \dots, L\}$ are low compared to the size of the network while the in-degrees $d_{L+1}, \dots, d_N$ of the nodes $\{L+1, \dots, N\}$ are much larger (for example comparable to the size of the network). For this reason, the first L nodes are called *low degree nodes* and the remaining $M = N - L$ nodes will be called *hubs*. The important *structural parameters* of the network, which we will also call *global parameters*, 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 the hubs;
- $\delta := \max_{1 \leq i \leq L} d_i$, the maximum in-degree of the low degree nodes.

The building blocks of the dynamics, which we will call *local parameters* are:

- the *local map*, $f : \mathbb{T} \rightarrow \mathbb{T}$, $f(x) = \sigma x \mod 1$, for some integer $\sigma \geq 2$;
- the *coupling function*, $h : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}$ which we assume¹ is C^{10} ;
- the *coupling strength*, $\alpha \in \mathbb{R}$.

¹This is to ensure sufficiently fast decay of Fourier coefficients in the series of h .

Expressing the coordinates as $\mathbf{z} = (z_1, \dots, z_N) \in \mathbb{T}^N$, the discrete-time evolution is given by the network map $\mathbf{F} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ on G defined by $\mathbf{z}' := F(\mathbf{z})$ with

$$z'_i = f(z_i) + \frac{\alpha}{\Delta} \sum_{n=1}^N A_{in} h(z_i, z_n) \mod 1, \quad i \in [N]. \quad (3.1)$$

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 and to deal best with indices in later calculations we write $\mathbf{z} = (\mathbf{x}, \mathbf{y})$, with $\mathbf{x} = (x_1, \dots, x_L) := (z_1, \dots, z_L) \in \mathbb{T}^L$ and $\mathbf{y} = (y_1, \dots, y_M) := (z_{L+1}, \dots, z_N) \in \mathbb{T}^M$. Moreover, decompose

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

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

$$x'_i = f(x_i) + \frac{\alpha}{\Delta} \sum_{n=1}^N A_{in} h(x_i, z_n) \mod 1 \quad i \in [L] \quad (3.2)$$

$$y'_j = g_j(y_j) + \xi_j(z) \mod 1 \quad j \in [M] \quad (3.3)$$

where

$$g_j(y) := f(y) + \alpha \kappa_j \int h(y, x) dm_1(x) \mod 1, \quad \kappa_j := \frac{d_{j+L}}{\Delta}, \quad (3.4)$$

and

$$\xi_j(\mathbf{z}) := \alpha \left[\frac{1}{\Delta} \sum_{n=1}^N A_{jn}^h h(y_j, z_n) - \kappa_j \int_{\mathbb{T}} h(y_j, x) dm_1(x) \right] \quad (3.5)$$

where we indicate with m_n the n -dimensional Lebesgue measure on $\mathbb{T}^n$.

In the following, we let $B_r(\Lambda)$ be the r -neighborhood of a set Λ and we define one-dimensional maps $g_j : \mathbb{T} \rightarrow \mathbb{T}$, $j = 1, \dots, M$ to be hyperbolic in a uniform sense.

Definition 3.1 (A Hyperbolic Collection of 1-Dimensional Map, see e.g. [dMvS93]). Given $\lambda \in (0, 1)$, $r > 0$ and $m, n \in \mathbb{N}$, we say that $g : \mathbb{T} \rightarrow \mathbb{T}$ is (n, m, λ, r) -hyperbolic if there exist an attracting set $\Lambda \subset \mathbb{T}$, with

1. $g(\Lambda) = \Lambda$,
2. $|Dg^n(x)| < \lambda$ for all $x \in B_r(\Lambda)$,
3. $|Dg^n(x)| > \lambda^{-1}$ for all $x \in B_r(\Upsilon)$ where $\Upsilon := \mathbb{T} \setminus W^s(\Lambda)$,

4. for each $x \notin B_r(\Upsilon)$, we have $f^k(x) \in B_r(\Lambda)$ for all $k \geq m$,

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

$$W^s(\Lambda) := \{x \in \mathbb{T} \text{ s.t. } \lim_{k \rightarrow \infty} d(g^k(x), \Lambda) = 0\}.$$

It is well known, see e.g. [dMvS93, Theorem IV.B] that for each C^2 map $g: \mathbb{T} \rightarrow \mathbb{T}$ (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 3.2. We say that a network with parameters L, M, Δ, δ is η -heterogeneous with $\eta > 0$ if there is $p, q \in [1, \infty)$ with $1 = 1/p + 1/q$, such that the following conditions are met:

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

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

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

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

Remark 3.1. (H1)-(H4) arise as sufficient conditions for requiring that the coupled system $\mathbf{F}$ is “close” to the product system $f \times \dots \times f \times g_1 \times \dots \times g_M : \mathbb{T}^{L+M} \rightarrow \mathbb{T}^{L+M}$ and preserves good hyperbolic properties on most of the phase space. They are verified in many common settings, as is shown in Appendix 3.6. An easy example to have in mind where those conditions are asymptotically satisfied as $N \rightarrow \infty$ for every $\eta > 0$, is the case where M is constant (so $L \sim N$), $\delta \sim L^\tau$, $\Delta \sim L^\gamma$ with $0 \leq \tau < 1/2$ and $(\tau + 1)/2 < \gamma < 1$. In particular the layered heterogeneous graphs satisfying (1.7) in the introduction to the paper have these properties.

Theorem A ([PvST17]). 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 , $j = 1, \dots, M$ is (n, m, λ, r) -hyperbolic. Then there exist $\xi_0, \eta, C > 0$ such that if the network is η -heterogeneous, for every $0 < \xi < \xi_0$ and for every $1 \leq T \leq T_1$ with

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

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

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

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

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

Remark 3.2. The result hold 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 viceversa 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 3.3), and when at least one of the g_j have an attracting point (Section 3.4).

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 ([PvST17]). *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 $h \in \Gamma$, $g_j \in C^{10}(\mathbb{T}, \mathbb{T})$, defined by Eq. (3.4), is hyperbolic (as in Definition 3.1).*

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 $I \subset \mathbb{R}$ for which if $\alpha\kappa_j \in I$ then g_j has a nonempty 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 that for the hub $j \in [M]$, $\alpha\kappa_j \in I$. Then there exists $C > 0$ and $\chi \in (0, 1)$ so that the following holds. Let $T_1 := \exp[C\Delta\xi^2]$. There is a set of initial conditions $\Omega_T \subset \mathbb{T}^N$ with

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

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

$$\text{dist}(z_j(t), \bar{z}(t \bmod p)) \leq \xi$$

for each $1/\xi \leq t \leq T \leq T_1$.

Proof. See Section 3.5. □

Remark 3.3. In the setting of the theorem above, consider the case where two hubs $j_1, j_2 \in [M]$ have the same connectivity κ , and their reduced dynamics g_{j_i} have 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 one to another resulting in highly coherent behaviour:

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

under the same 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 hubs's coordinates in the same connected component of the basin of attraction of one of the periodic orbits yield 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.

Theorem C ([PvST17]).

- a) Take a diffusive coupling function $h(x, y) = \varphi(y - x)$ for some $\varphi : \mathbb{T} \rightarrow \mathbb{R}$ 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 $p \in \left(\frac{\log N}{N}, 1\right]$ there exists a subset of undirected homogeneous graphs $\mathcal{G}_{Hom}(N) \subset \mathcal{G}(N)$, with $\mathbb{P}_p(\mathcal{G}_{Hom}(N)) \rightarrow 1$ as $N \rightarrow \infty$ so that for any $\alpha \in \mathcal{I}$ the synchronization manifold $\mathcal{S}$, defined in Eq. (1.8), is locally exponentially stable (normally attracting) for each network coupled on $G \in \mathcal{G}_{Hom}(N)$.
- b) Take any sequence of graphs $\{G(N)\}_{N \in \mathbb{N}}$ where $G(N)$ has N nodes and non-decreasing sequence of degrees $\mathbf{d}(N) = (d_{1,N}, \dots, d_{N,N})$. Then, if $d_{N,N}/d_{1,N} \rightarrow \infty$ for $N \rightarrow \infty$, 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 Section 3.6. □

3.2 Sketch of the Proof and the Use of a Truncated System

3.2.1 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 from the reduced dynamics,
- control on the *hitting time* to the bad set,
- *finite time* exponentially large on 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 $\{1, \dots, L + 1\}$ and set of directed edges $\mathcal{E} = \{(L + 1, i) : 1 \leq i \leq L\}$ (see Figure 3.1). The hub node $\{L + 1\}$ has an incoming directed edge from every other node of the network, while the other nodes have just the outgoing edge. Take as interaction function the diffusive coupling $h(x, y) := \sin(2\pi y) - \sin(2\pi x)$. Equations (3.2) and (3.3) then become

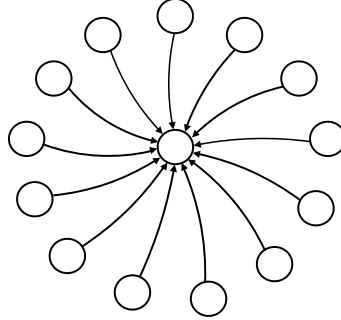


Figure 3.1: Star network with uncoupled low degree nodes.

$$x_i(t + 1) = 2x_i(t) \pmod{1} \quad i \in [L] \quad (3.6)$$

$$y(t + 1) = 2y(t) + \frac{\alpha}{L} \sum_{i=1}^L [\sin(2\pi x_i(t)) - \sin(2\pi y(t))] \pmod{1}. \quad (3.7)$$

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

$$y(t + 1) = 2y(t) - \alpha \sin(2\pi y(t)) + \frac{\alpha}{L} \sum_{i=1}^L \sin(2\pi x_i(t)) \quad (3.8)$$

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

$$\xi(t) = \frac{\alpha}{L} \sum_{i=1}^L \sin(2\pi x_i(t)) \quad (3.9)$$

depending on the low degree nodes coordinates. The Lebesgue measure on $\mathbb{T}^L$ is invariant and mixing for the dynamics restricted to first L uncoupled coordinates. The set of bad states where fluctuations (3.9) are above a fixed threshold $\varepsilon > 0$ is

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

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 on L and ε , see the Hoeffding Inequality in Appendix A.1 for details). Since m_L is invariant and ergodic for the dynamics of the low degree nodes, 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 full set of initial conditions will visit $\mathcal{B}_\varepsilon$ in finite time, making any mean-field approximation result for infinite time hopeless.
- As a consequence of Kac Lemma, the average hitting time to 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 $\Omega_T \subset \mathbb{T}^{L+1}$ with measure $m_{L+1}(\Omega_T) > 1 - T \exp(-C\varepsilon^2 N)$ such that $\forall x \in \Omega_T$ and for every $1 \leq t \leq T$

$$\left| \frac{1}{L} \sum_{i=1}^L \sin(2\pi x_i(t)) \right| \leq \varepsilon.$$

3.2.2 Truncated System

We obtain a description of the fully coupled system² by restricting our attention to a subset of phase space where the evolution prescribed by equations (3.2) and (3.3) 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 $\mathbf{F}_\varepsilon : \mathbb{T}^N \rightarrow \mathbb{T}^N$, for which the fluctuations of the mean field averages are artificially cut-off at the level $\varepsilon > 0$, resulting in a well behaved hyperbolic

²when also the low degree nodes receive an interaction from the hub.

dynamical system. In the following sections we will then determine existence and bounds on the invariant measure for the truncated system and prove that the portion of phase space where the original system, $\mathbf{F}$, and the truncated one coincide is 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

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

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

$$|c_{\mathbf{s}}| \leq \frac{\|h\|_{C^{10}}}{|s_1|^5 |s_2|^5}. \quad (3.10)$$

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

$$\xi_j(\mathbf{z}) := \alpha \sum_{\mathbf{s} \in \mathbb{Z}^2} c_{\mathbf{s}} \left[\frac{1}{\Delta} \sum_{n=1}^N A_{jn}^h \theta_{s_1}(z_n) - \kappa_j \bar{\theta}_{s_1} \right] \theta_{s_2}(y_j) + \frac{\alpha}{\Delta} \sum_{n=1}^M A_{jn}^h h(y_j, y_n) \quad (3.11)$$

For every $\varepsilon > 0$ choose a C^∞ map $\zeta_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$ with $\zeta_\varepsilon(t) = t$ for $|t| < \varepsilon$, $\zeta_\varepsilon(t) = 2\varepsilon$ for $|t| > 2\varepsilon$. So for each $\varepsilon > 0$, the function $t \mapsto |D\zeta_\varepsilon(t)|$ is uniformly bounded in t and ε . We define the evolution for the truncated dynamics $\mathbf{F}_\varepsilon : \mathbb{T}^{L+M} \rightarrow \mathbb{T}^{L+M}$ by the following modification of equations (3.2) and (3.3):

$$x'_i = f(x_i) + \frac{\alpha}{\Delta} \sum_{n=1}^N A_{in} h(x_i, z_n) \mod 1 \quad i \in [L] \quad (3.12)$$

$$y'_j = g_j(y_j) + \xi_{j,\varepsilon}(\mathbf{z}) \mod 1 \quad j \in [M] \quad (3.13)$$

where the expression of $\xi_{j,\varepsilon}(z)$ modifies that of $\xi_j(z)$ in (3.11):

$$\xi_{j,\varepsilon}(\mathbf{z}) := \alpha \sum_{\mathbf{s} \in \mathbb{Z}^2} c_{\mathbf{s}} \zeta_{\varepsilon|s_1|} \left(\frac{1}{\Delta} \sum_{i=1}^L A_{ji} \theta_{s_1}(x_i) - \kappa_j \bar{\theta}_{s_1} \right) \theta_{s_2}(y_j) + \frac{\alpha}{\Delta} \sum_{n=1}^M A_{jn}^h h(y_j, y_n). \quad (3.14)$$

So the only difference between $\mathbf{F}$ and $\mathbf{F}_\varepsilon$ are the cut-off functions $\zeta_{\varepsilon|s_1|}$ appearing in (3.14). For every $\varepsilon > 0$, $j \in [M]$ and $s_1 \in \mathbb{Z}$ define

$$\mathcal{B}_\varepsilon^{(s_1, j)} := \left\{ \mathbf{x} \in \mathbb{T}^L : \left| \frac{1}{\Delta} \sum_{i=1}^L A_{ji} \theta_{s_1}(x_i) - \kappa_j \bar{\theta}_{s_1} \right| > \varepsilon |s_1| \right\}. \quad (3.15)$$

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

$$\mathcal{Q}_\varepsilon := \bigcap_{j=1}^M \bigcap_{s_1 \in \mathbb{Z}} \mathbb{T}^L \setminus \mathcal{B}_\varepsilon^{(s_1, j)} \quad (3.16)$$

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 where the fluctuations exceed the threshold, and the systems $\mathbf{F}$ and $\mathbf{F}_\varepsilon$ are different. Furthermore we can control the perturbation introduced by the term $\xi_{j,\varepsilon}$ in equation (3.3) so that $\mathbf{F}_\varepsilon$ is close to the hyperbolic uncoupled product map $\mathbf{f} : \mathbb{T}^N \rightarrow \mathbb{T}^N$

$$\mathbf{f}(x_1, \dots, x_L, y_1, \dots, y_M) := (f(x_1), \dots, f(x_L), g_1(y_1), \dots, g_M(y_M)). \quad (3.17)$$

All the bounds on relevant norms of $\xi_{j,\varepsilon}$ are reported in Appendix A.1. To upper bound the Lebesgue measure $m_L(\mathcal{B}_\varepsilon)$ we use the Hoeffding's inequality (reported in Appendix A.1) on concentration of the average of independent bounded random variables around the expected value.

Proposition 3.1.

$$m_L(\mathcal{B}_\varepsilon) \leq \frac{\exp \left[-\frac{\Delta \varepsilon^2}{2} + \mathcal{O}(\log M) \right]}{1 - \exp \left[-\frac{\Delta \varepsilon^2}{2} \right]}. \quad (3.18)$$

Proof. See Appendix A.1. □

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 it to upper bound the measure of $\mathcal{B}_\varepsilon$ with respect to SRB measures for $\mathbf{F}_\varepsilon$, which is the measure giving statistical informations on the orbits of $\mathbf{F}_\varepsilon$.

Remark 3.4. Notice that in (3.18) 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 magnitudes with respect to the aforementioned parameters of the network (Δ, δ, L, M) .

3.2.3 Steps of the Proof and Challenges

The basic steps of the proof are the following.

- (i) First of all we restrict the attention to the case where g_j satisfy Definition 3.1 with $n = 1$ for all $j \in [M]$.
- (ii) Secondly, we establish the hyperbolicity of the map $\mathbf{F}_\varepsilon$ for an η -heterogeneous network with $\varepsilon, \eta > 0$ small. This is achieved by constructing forward and backward invariant cone-fields made of expanding and contracting directions respectively for the cocycle defined by application of $\mathbf{D}_z \mathbf{F}_\varepsilon$ (A.2).
- (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 are sometimes called *standard pairs*, [CLP16], to estimate the regularity properties of the SRB measures for the endomorphism $\mathbf{F}_\varepsilon$, and the hitting time to the set $\mathcal{B}_\varepsilon$.
- (v) Finally, we show that Mather's trick allow us to generalise the proofs to the case in which g_j satisfy Definition 3.1 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 (Section 3.3 and Section 3.4). At the end of Section 3.4 we put the results together to obtain the proof of Theorem A.

In the above points we treat $\mathbf{F}_\varepsilon$ 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 $\mathbf{F}_\varepsilon$ is close to the uncoupled product map $\mathbf{f}$. 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 propagate nonlinearities of the maps and reduces the global regularity of the invariant measures. Secondly, the situation is inherently different from 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 on the system size. For these reasons we find the geometric approach advantageous with respect to the functional analytic approach [KL99] where the explicit dependence of most constants on the dimension are hidden in the functional analytic machinery.

3.3 Proof of Theorem A when all Reduced Maps are Uniformly Expanding

In this section we assume that the collection of reduced maps g_j from equation (3.4) is uniformly expanding for all $j \in [M]$. As shown in Lemma 3.6 this means that we can assume that there exists $\lambda \in (0, 1)$, so that $|g_j(x)| \geq \lambda^{-1}$ for all $x \in \mathbb{T}$ and $j \in [M]$. Under this assumptions we prove

Theorem 3.1. *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 $\mathbf{F}_\varepsilon$. The density $\rho = d\nu/dm_N$ satisfies*

$$\frac{\rho(\mathbf{z})}{\rho(\bar{\mathbf{z}})} \leq \exp \{ad_p(\mathbf{z}, \bar{\mathbf{z}})\}, \quad a = \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M). \quad (3.19)$$

for all $\mathbf{z}, \bar{\mathbf{z}} \in \mathbb{T}^N$

In Section 3.3.1 we obtain conditions on the heterogenous structure of the network which ensure that the truncated system F_ε is sufficiently close, in the C^1 topology, to the uncoupled system $\mathbf{f}$, in Eq. (3.17), 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 $\mathbf{F}_\varepsilon$ 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 $\mathbf{F}_\varepsilon$

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

(See Appendix C for a quick review on the theory of transfer operators). For our purposes we will also require bounds on ρ which are explicit on the structural parameters of the network (for suitable ε). In Section 4.2.2 we obtain bounds on the distortion of the Jacobian of $\mathbf{F}_\varepsilon$ (Proposition 4.6), 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 3.4) and to which ρ belongs. To deduce the conclusion of Theorem A, we need that the ν -measure of the set $\mathcal{B}_\varepsilon$ is small. This will be obtained from an upper bound for the supremum of the functions in the invariant cone and is shown in Section 3.3.4 under some additional conditions on the network.

3.3.1 Global Expansion of F_ε

For $1 \leq p \leq \infty$, let $1 \leq q \leq \infty$ be so 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 $\mathbb{R}^k$. $\|\cdot\|_p$ induces the operator norm of any linear map $\mathcal{L}: \mathbb{R}^N \rightarrow \mathbb{R}^N$, namely

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

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

Proposition 3.2. *Suppose that for every $j \in [M]$ the reduced map g_j is uniformly expanding, i.e. there exists $\lambda \in (0, 1)$ so that $|Dg_j(y)| > \lambda^{-1}$ for all $y \in \mathbb{T}$. Then*

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

$$\frac{\|(\mathbf{D}_z \mathbf{F}_\varepsilon) \mathbf{w}\|_p}{\|\mathbf{w}\|_p} \geq \left[\min\{\sigma, \lambda^{-1} - \varepsilon C_\#\} - \mathcal{O}(\Delta^{-1/p} M^{1/p}) - \mathcal{O}(\Delta^{-1} N^{1/p} \delta^{1/q}) \right];$$

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

$$\varepsilon < \frac{\lambda^{-1} - 1}{C_\#} =: \varepsilon_0 \tag{3.20}$$

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

$$\frac{\|(\mathbf{D}_z \mathbf{F}_\varepsilon) \mathbf{w}\|_p}{\|\mathbf{w}\|_p} \geq \bar{\sigma} > 1, \quad \forall \mathbf{z} \in \mathbb{T}^N, \quad \forall \mathbf{w} \in \mathbb{R}^N \setminus \{0\}.$$

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

$$\begin{pmatrix} \mathbf{u}' \\ \mathbf{v}' \end{pmatrix} = \mathbf{D}_z \mathbf{F}_\varepsilon \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}, \quad \mathbf{u}' \in \mathbb{R}^L, \mathbf{v}' \in \mathbb{R}^M.$$

Using (3.12)-(3.13), or (A.2), we obtain that for every $i \in [L]$ and $j \in [M]$,

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

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

Hence

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

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

$$\|\mathbf{w}\|_{1,\mathbb{R}^k} \leq k^{1/q} \|\mathbf{w}\|_{p,\mathbb{R}^k}, \text{ with } \frac{1}{p} + \frac{1}{q} = 1 \quad (3.21)$$

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

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

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

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

Analogously using the estimates in Lemma A.1

$$\begin{aligned} \|\mathbf{v}'\|_p &\geq (\lambda^{-1} - \varepsilon C_\#) \|\mathbf{v}\|_p - \mathcal{O}(\Delta^{-1}M^{1/p}) \max_{j=1,\dots,M} \left[\sum_n A_{jn} |w_n| \right] \\ &\geq (\lambda^{-1} - \varepsilon C_\# - \mathcal{O}(\Delta^{-1}M)) \|\mathbf{v}\|_p - \mathcal{O}(\Delta^{-1/p}M^{1/p}) \|\mathbf{w}\|_p \end{aligned} \quad (3.22)$$

since in the sum $\sum_n A_{jn} |w_n|$ in (3.22), at most Δ terms are different from zero and since $\Delta^{-1}\Delta^{1/q} = \Delta^{-1/p}$. This implies

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

For the proof of (ii), notice that condition (3.20) implies that $\min\{\sigma, \lambda^{-1} - \varepsilon C_\#\} > 1$ and conditions (H1)-(H3) imply that the $\mathcal{O}$ are bounded by η and so

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

and choosing $\eta > 0$ sufficiently small one obtains the proposition. $\square$

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

3.3.2 Distortion of $\mathbf{F}_\varepsilon$

Proposition 3.3. *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 $\mathbf{z}, \bar{\mathbf{z}} \in \mathbb{T}^N$*

$$\frac{|\mathbf{D}_\mathbf{z}\mathbf{F}_\varepsilon|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon|} \leq \exp \left\{ [\mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \right\}.$$

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

$$(\mathcal{D}(\mathbf{z}))_{k\ell} := \begin{cases} 1 + \frac{\alpha}{\Delta} \sum_{n=1}^L A_{kn} \frac{h_1(x_k, z_n)}{\sigma} & k = \ell \leq L, \\ \frac{\alpha}{\Delta} A_{k\ell} \frac{h_2(x_k, x_\ell)}{\sigma} & k \neq \ell \leq L, \\ \frac{\partial_{x_\ell} \xi_{k-L, \varepsilon}}{\sigma} & k > L, \ell \leq L, \\ \frac{\alpha}{\Delta} A_{k\ell} \frac{h_2(y_{k-L}, y_{\ell-L})}{D_{y_{\ell-L}} g_{\ell-L}} & k \neq \ell > L, \\ 1 + \frac{\partial_{y_{k-L}} \xi_{k-L, \varepsilon}}{D_{y_{\ell-L}} g_{\ell-L}} & k = \ell > L. \end{cases} \quad (3.23)$$

and

$$\frac{|\mathbf{D}_\mathbf{z}\mathbf{F}_\varepsilon|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon|} = \frac{\prod_{j=1}^M D_{y_j} g_j}{\prod_{j=1}^M D_{\bar{y}_j} g_j} \cdot \frac{|\mathcal{D}(\mathbf{z})|}{|\mathcal{D}(\bar{\mathbf{z}})|}.$$

For the first ratio:

$$\prod_{j=1}^M \frac{D_{y_j} g_j}{D_{\bar{y}_j} g_j} = \prod_{j=1}^M \left(1 + \frac{D_{y_j} g_j - D_{\bar{y}_j} g_j}{D_{\bar{y}_j} g_j} \right) \leq \prod_{j=1}^M (1 + \mathcal{O}(1)|y_j - \bar{y}_j|) \leq \exp [\mathcal{O}(M) d_\infty(y, \bar{y})]. \quad (3.24)$$

To estimate the ratio $\frac{|\mathcal{D}(\mathbf{z})|}{|\mathcal{D}(\bar{\mathbf{z}})|}$ we apply Proposition A.1 in Appendix A.2. To this end define the matrix

$$\mathbf{B}(\mathbf{z}) := \mathcal{D}(\mathbf{z}) - \text{Id}.$$

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

$$\|\mathbf{B}(\mathbf{z})\|_p \leq \max\{\mathcal{O}(\Delta^{-1}M), C_\# \varepsilon\} + \mathcal{O}(\Delta^{-1/p} M^{1/p}) + \mathcal{O}(\Delta^{-1} N^{1/p} \delta^{1/q}) \quad (3.25)$$

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

$$\begin{aligned} u'_i &= \left[\frac{\alpha}{\Delta} \sum_{n=1}^L A_{in} \frac{h_1(x_i, z_n)}{\sigma} \right] u_i + \frac{\alpha}{\Delta} \sum_{\ell=1}^L A_{i\ell}^{ll} \frac{h_1(x_i, x_\ell)}{\sigma} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^M A_{im}^{lh} \frac{h_2(x_i, y_m)}{D_{y_m} g_m} v_m \\ v'_j &= \sum_{\ell=1}^L \frac{\partial_{x_\ell} \xi_{j,\varepsilon}}{\sigma} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^M A_{jm}^{hh} \frac{h_2(y_j, y_m)}{D_{y_m} g_m} v_m + \frac{\partial_{y_j} \xi_{j,\varepsilon}}{D_{y_j} g_j} v_j. \end{aligned}$$

Using estimates analogous the ones used in the proof of Proposition 3.2

$$\begin{aligned} \|\mathbf{u}'\|_p &\leq \mathcal{O}(\Delta^{-1}\delta) \|\mathbf{u}\|_p + \mathcal{O}(\Delta^{-1}) \max_i \left[\sum_{\ell=1}^L A_{i\ell}^{ll} |u_\ell| + \sum_{m=1}^M A_{im}^{lh} |v_m| \right] \\ &\leq \mathcal{O}(\Delta^{-1}\delta) \|u\|_p + \mathcal{O}(\Delta^{-1} N^{1/p} \delta^{1/q}) \|(\mathbf{u}, \mathbf{v})\|_p \\ \|\mathbf{v}'\|_p &\leq C_\# \varepsilon \|\mathbf{v}\|_p + \mathcal{O}(\Delta^{-1} N^{1/p}) \max_i \left[\sum_{\ell} A_{i\ell}^{ll} |u_\ell| + \sum_m A_{im}^{lh} |v_m| \right] \\ &\leq C_\# \varepsilon \|\mathbf{v}\|_p + \mathcal{O}(\Delta^{-1/p} M^{1/p}) \|(\mathbf{u}, \mathbf{v})\|_p \end{aligned}$$

so using conditions (H1), (H2), we obtain (3.25):

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

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

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

For $i \in [L]$, looking at the entries of $\mathcal{D}(\mathbf{z})$, (3.23), it is clear that the non-vanishing entries $[\mathbf{B}(\mathbf{z})]_{ik}$ for $k \neq i$ are Lipschitz functions with Lipschitz constants of the order $\mathcal{O}(\Delta^{-1})$:

$$|B_{ik} - \bar{B}_{ik}| \leq A_{ik} \mathcal{O}(\Delta^{-1}) d_\infty(\mathbf{z}, \bar{\mathbf{z}})$$

Instead, for $k = i$,

$$\begin{aligned} |B_{ii} - \bar{B}_{ii}| &= \frac{\alpha}{\Delta} \left| \sum_{\ell} A_{in}^{ll} (h_1(x_i, x_\ell) - h_1(\bar{x}_i, \bar{x}_\ell)) + \sum_m A_{im}^{lh} (h_1(x_i, y_m) - h_1(\bar{x}_i, \bar{y}_m)) \right| \\ &\leq \frac{\alpha}{\Delta} \sum_{\ell} A_{i\ell}^{ll} |h_1(x_i, x_\ell) - h_1(\bar{x}_i, \bar{x}_\ell)| + \frac{\alpha}{\Delta} \sum_m A_{im}^{lh} |h_1(x_i, y_m) - h_1(\bar{x}_i, \bar{y}_m)| \\ &\leq \mathcal{O}(\Delta^{-1}\delta) d_\infty(\mathbf{z}, \bar{\mathbf{z}}). \end{aligned}$$

which implies

$$\begin{aligned}
 \|\text{Col}^i[\mathbf{B} - \bar{\mathbf{B}}]\|_p &= \left(\sum_{k \in [L]} |B_{ik} - \bar{B}_{ik}|^p \right)^{\frac{1}{p}} + \left(\sum_{k \in [L+1, N]} |B_{ik} - \bar{B}_{ik}|^p \right)^{\frac{1}{p}} \\
 &\leq \left(\sum_{k \in [L] \setminus \{i\}} |B_{ik} - \bar{B}_{ik}|^p \right)^{\frac{1}{p}} + \left(\sum_{k \in [L+1, N]} |B_{ik} - \bar{B}_{ik}|^p \right)^{\frac{1}{p}} + \mathcal{O}(\Delta^{-1}\delta)d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \\
 &\leq 2 \left(\sum_{k \neq i} A_{ik} \right)^{\frac{1}{p}} \mathcal{O}(\Delta^{-1})d_\infty(\mathbf{z}, \bar{\mathbf{z}}) + \mathcal{O}(\Delta^{-1}\delta)d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \\
 &\leq \mathcal{O}(\Delta^{-1}\delta)d_\infty(\mathbf{z}, \bar{\mathbf{z}}).
 \end{aligned}$$

For $j \in [M]$, looking again at (3.23) the non-vanishing entries of $[\mathbf{B}(\mathbf{z})]_{(j+L)k}$ for $k \neq j + N$ are Lipschitz functions with Lipschitz constants of the order $\mathcal{O}(\Delta^{-1})$, while $[\mathbf{B}(\mathbf{z})]_{(j+L)(j+L)}$ has Lipschitz constant of order $\mathcal{O}(1)$, thus

$$\begin{aligned}
 \|\text{Col}^{j+L}[\mathbf{B} - \bar{\mathbf{B}}]\|_p &= \left(\sum_{k \in [L]} |B_{(j+L)k} - \bar{B}_{(j+L)k}|^p \right)^{\frac{1}{p}} + \left(\sum_{k \in [L+1, N]} |B_{(j+L)k} - \bar{B}_{(j+L)k}|^p \right)^{\frac{1}{p}} \\
 &\leq \left(\sum_{k \in [L]} |B_{(j+L)k} - \bar{B}_{(j+L)k}|^p \right)^{\frac{1}{p}} + \\
 &\quad + \left(\sum_{k \in [L+1, N] \setminus \{j+L\}} |B_{(j+L)k} - \bar{B}_{(j+L)k}|^p \right)^{\frac{1}{p}} + \mathcal{O}(1)d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \\
 &\leq 2 \left(\sum_{k \neq j+L} A_{(j+L)k} \right)^{\frac{1}{p}} \mathcal{O}(\Delta^{-1})d_\infty(\mathbf{z}, \bar{\mathbf{z}}) + \mathcal{O}(1)d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \\
 &\leq \mathcal{O}(1)d_\infty(\mathbf{z}, \bar{\mathbf{z}}).
 \end{aligned}$$

Proposition A.1 from Appendix A.2 now implies that

$$\frac{|\mathcal{D}(\mathbf{z})|}{|\mathcal{D}(\bar{\mathbf{z}})|} \leq \exp \left\{ \sum_{k=1}^N \text{Col}^k[\mathbf{B} - \bar{\mathbf{B}}] \right\} \leq \exp \{ (\mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M))d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \}. \quad (3.26)$$

□

3.3.3 Invariant Cones of Densities

Define the cone of functions

$$C_{a,p} := \left\{ \varphi : \mathbb{T}^N \rightarrow \mathbb{R}^+ : \frac{\varphi(\mathbf{z})}{\varphi(\bar{\mathbf{z}})} \leq \exp[ad_p(\mathbf{z}, \bar{\mathbf{z}})], \forall \mathbf{z}, \bar{\mathbf{z}} \in \mathbb{T}^N \right\}.$$

This 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 restricted to this set and one can show that the unique fixed point 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 3.4. *Under conditions (H1)-(H3), for every $a > a_c$, where a_c is of the form*

$$a_c = \frac{\mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)}{1 - \bar{\sigma}}, \quad (3.27)$$

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

Proof. Since $\mathbf{F}_\varepsilon$ is a local expanding diffeomorphism, its transfer operator, P_ε , has expression

$$(P_\varepsilon \varphi)(z) = \sum_i \varphi(\mathbf{F}_{\varepsilon,i}^{-1}(z)) \left| \mathbf{D}_{\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z})} \mathbf{F}_\varepsilon \right|^{-1}$$

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

$$\begin{aligned} \frac{\varphi(\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z}))}{\varphi(\mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}}))} \frac{\left| \mathbf{D}_{\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z})} \mathbf{F}_\varepsilon \right|}{\left| \mathbf{D}_{\mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}})} \mathbf{F}_\varepsilon \right|} &\leq \\ &\leq \exp \left\{ ad_p(\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z}), \mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}})) \right\} \exp \left\{ [\mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_\infty(\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z}), \mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}})) \right\} \\ &\leq \exp \left\{ [a + \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_p(\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z}), \mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}})) \right\} \\ &\leq \exp \left\{ [\bar{\sigma}^{-1}a + \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_p(\mathbf{z}, \bar{\mathbf{z}}) \right\}. \end{aligned}$$

Here we used that $d_\infty(\mathbf{z}, \bar{\mathbf{z}}) \leq d_p(\mathbf{z}, \bar{\mathbf{z}})$ for every $1 \leq p < \infty$. Hence

$$\begin{aligned} \frac{(P_\varepsilon \varphi)(\mathbf{z})}{(P_\varepsilon \varphi)(\bar{\mathbf{z}})} &= \frac{\sum_i \varphi(\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z})) \left| \mathbf{D}_{\mathbf{F}_{\varepsilon,i}^{-1}(\mathbf{z})} \mathbf{F}_\varepsilon \right|^{-1}}{\sum_i \varphi(\mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}})) \left| \mathbf{D}_{\mathbf{F}_{\varepsilon,i}^{-1}(\bar{\mathbf{z}})} \mathbf{F}_\varepsilon \right|^{-1}} \\ &\leq \exp \left[(\bar{\sigma}^{-1}a + \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)) d_p(\mathbf{z}, \bar{\mathbf{z}}) \right]. \end{aligned}$$

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

Proof of Theorem 3.1. The existence of the absolutely continuous invariant probability measure is standard from the expansivity of $\mathbf{F}_\varepsilon$. The regularity bound on the density immediately follows from Proposition 3.4 and from the observation (that can be found in [Via97]) that $P_\varepsilon^n \varphi$ converges uniformly to the invariant density of $\mathbf{F}_\varepsilon$ for any density in the invariant cone C_a . $\square$

3.3.4 Proof of Theorem A in the Expanding Case

Property (3.19) 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 3.4 holds, the invariant density for $\mathbf{F}_\varepsilon$ belongs to the cone $C_{a,p}$, $\rho \in C_{a,p}$, for $a > a_c$. Since ρ is a continuous density, it has to take value one at some point in its domain. This together with the regularity condition given by the cone implies that

$$\sup_{\mathbf{z} \in \mathbb{T}^N} \rho(\mathbf{z}) \leq \exp \left\{ \mathcal{O}(\Delta^{-1} \delta L^{1/p}) + \mathcal{O}(ML^{1/p}) \right\}. \quad (3.28)$$

Using the upper bound (3.18),

$$\begin{aligned} \nu(B_\varepsilon) &= \int_{B_\varepsilon} \rho(\mathbf{z}) dm_N(\mathbf{z}) \\ &\leq m_N(B_\varepsilon) \sup_{\mathbf{z} \in \mathbb{T}^N} \rho(\mathbf{z}) \\ &\leq \exp \left\{ -\Delta \varepsilon^2 / K^2 + \mathcal{O}(\Delta^{-1} \delta L^{1+1/p}) + \mathcal{O}(ML^{1/p}) \right\}. \end{aligned}$$

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

$$\begin{aligned} \nu \left(\bigcup_{t=0}^T \mathbf{F}_\varepsilon^{-t}(\mathcal{B}_\varepsilon) \right) &\leq (T+1) \nu(\mathcal{B}_\varepsilon) \\ &\leq (T+1) \exp \left\{ -\Delta \varepsilon^2 / K^2 + \mathcal{O}(\Delta^{-1} \delta L^{1+1/p}) + \mathcal{O}(ML^{1/p}) \right\}. \end{aligned}$$

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

$$m_N \left(\bigcup_{t=0}^T \mathbf{F}_\varepsilon^{-t}(\mathcal{B}_\varepsilon) \right) = \int_{\bigcup_{t=0}^T \mathbf{F}_\varepsilon^{-t}(\mathcal{B}_\varepsilon)} \rho^{-1} d\nu$$

$$\begin{aligned}
&\leq \nu \left(\bigcup_{t=0}^T \mathbf{F}_\varepsilon^{-t}(\mathcal{B}_\varepsilon) \right) \exp \left\{ -\Delta \varepsilon^2 / K^2 + \mathcal{O}(\Delta^{-1} \delta L^{1+1/p}) + \mathcal{O}(ML^{1/p}) \right\} \\
&\leq (T+1) \exp \left\{ -\Delta \varepsilon^2 / K^2 + \mathcal{O}(\Delta^{-1} \delta L^{1+1/p}) + \mathcal{O}(ML^{1/p}) \right\} \\
&\leq (T+1) \exp \left\{ -\Delta \varepsilon^2 / K^2 + \mathcal{O}(\eta) \Delta \right\}.
\end{aligned}$$

Where we used (H4) to obtain the last inequality. Hence, the set

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

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

3.4 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 $j \in [M_u]$, while for $j \in [M_u + 1, M]$, 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 $n = 1$. We will show how to drop this assumption in Lemma 3.6.

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 $\mathbf{F}_\varepsilon$ 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 3.4.8 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 $(\mathbf{z}_u, \mathbf{z}_s)$ for a point in $\mathbb{T}^{L+M_u} \times \mathbb{T}^{M_s} = \mathbb{T}^N$ and $\mathbf{z}_u = (\mathbf{x}, \mathbf{y}_u) \in \mathbb{T}^L \times \mathbb{T}^{M_u}$. Let

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

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

We begin by pointing out the existence of the forward invariant set $\mathcal{S}$.

Lemma 3.1. *As before, for $j \in [M_u + 1, M]$, let Λ_j be the attracting sets of g_j and $\Upsilon = \mathbb{T} \setminus W_s(\Lambda_j)$. There exist $\lambda \in (0, 1)$, $\varepsilon_\Lambda > 0$, $r_0 > 0$ so that for each $j \in [M_u + 1, M]$ and each $|r| < r_0$,*

(i) $|Dg_j(y)| < \lambda < 1$ for every $y \in U_j$ and $g_j(x) + r \in U_j$, $\forall x \in U_j$, where U_j is the ε_Λ -neighborhood of Λ_j .

(ii) $|Dg_j| > \lambda^{-1}$ on the ε_Λ -neighborhood of Υ_j , $\forall j \in [M_u, M]$, where Υ_j is the complement of the basin of Λ_j in $\mathbb{T}$.

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 $r < (1 - \lambda)\varepsilon_\Lambda$,

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

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

Let

$$\mathcal{R} := U_{M_u+1} \times \dots \times U_M \quad \text{and} \quad \mathcal{S} := \mathbb{T}^{L+M_u} \times \mathcal{R} \subset \mathbb{T}^N. \quad (3.29)$$

Lemma 3.1 implies that provided the ε from the truncated system is below r_0 , the set $\mathcal{S}$ is forward invariant under $\mathbf{F}_\varepsilon$. It follows that for each attracting periodic orbit $O(\mathbf{z}_s)$ of $g_{M_u+1} \times \dots \times g_M: \mathbb{T}^{M_s} \rightarrow \mathbb{T}^{M_s}$, the endomorphism $\mathbf{F}_\varepsilon$ has a fat solenoidal invariant set. Indeed, take the union U of the connected component of $\mathcal{R}$ containing $O(\mathbf{z}_s)$. Then by the previous lemma, $\mathbf{F}_\varepsilon(\mathbb{T}^{L+M_u} \times U) \subset \mathbb{T}^{L+M_u} \times U$. The set $\bigcap_{n \geq 0} \mathbf{F}_\varepsilon^n(\mathbb{T}^{L+M_u} \times U)$ is the analogue of the usual solenoid but with self-intersections, see Figure 3.2. An analogous situation, but where the map is a skew product is studied in [Tsu01]. The set $\bigcap_{n \geq 0} \mathbf{F}_\varepsilon^n(\mathbb{T}^{L+M_u} \times U)$ will support an invariant measure:

Theorem 3.2. *Under the hypotheses of Theorem A with $\eta > 0$ sufficiently small*

- *for every attracting periodic orbit of $g_{M_u+1} \times \dots \times g_M$, $\mathbf{F}_\varepsilon$ 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 $\forall \mathbf{z}_u, \bar{\mathbf{z}}_u \in \mathbb{T}^{L+M_u} \times \{0\}$,*

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

- *these are the only physical measures for $\mathbf{F}_\varepsilon$.*

This theorem will be proved in Subsection 3.4.5.

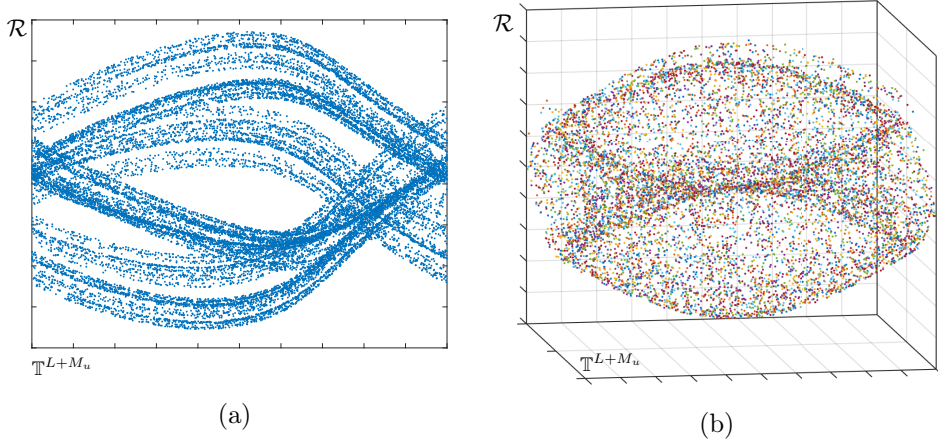


Figure 3.2: Approximate 2D and 3D representations of one component of the attractor of $\mathbf{F}_\varepsilon$.

3.4.1 Strategy of the proof of Theorem A in the Presence of Hyperbolic Attractors

Supposing that the threshold of the fluctuations is below r_0 as defined in Lemma 3.1, we restrict our attention, for the time being, to the map $\mathbf{F}_\varepsilon|_{\mathcal{S}} : \mathcal{S} \rightarrow \mathcal{S}$ that we will still call $\mathbf{F}_\varepsilon$ with an abuse of notation. The expression for $\mathbf{F}_\varepsilon$ is the same as in equations (3.12) and (3.13), but now the local phase space for the hubs with a non-empty attractor, $\{L + M_u + 1, \dots, L + M = N\}$, is restricted to the open set $\mathcal{R}$.

The proof of Theorem A will follow from the following proposition.

Proposition 3.5. *For every $s_1 \in \mathbb{Z}$ and $j \in [M]$*

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

is bounded as

$$m_N \left(\mathcal{B}_{\varepsilon, T}^{(s_1, j)} \right) \leq T \exp \left[-C\Delta\varepsilon^2 + \mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p}) \right]. \quad (3.30)$$

To prove the above result, we first build families of stable and unstable invariant cones for $\mathbf{F}_\varepsilon$ in the tangent bundle of $\mathcal{S}$ (Proposition 3.6) which correspond to contracting and expanding directions for the dynamics, thus proving hyperbolic behaviour of the map. In Section 3.4.3 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 action of $\mathbf{F}_\varepsilon$.

Bounding the Jacobian of the map restricted to the manifolds (Proposition 3.8) one can prove the existence of an invariant cone of densities (Proposition 3.9) 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 pre-images of $\mathbf{F}_\varepsilon^t$, we approximate it with a set which is the union of global stable manifolds (Lemma 3.3). This last property is preserved taking pre-images. The bound in (3.30) will then be a consequence of estimates on the distortion of the holonomy map along stable leaves of $\mathbf{F}_\varepsilon$ (Proposition 3.10).

3.4.2 Invariant Cone Fields for $\mathbf{F}_\varepsilon$

Proposition 3.6. *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 $\varepsilon > 0$*

$$\varepsilon < \min \left\{ \frac{1 - \lambda}{C_\#}, \frac{\lambda^{-1} - 1}{C_\#}, \varepsilon_0 \right\} \quad (3.31)$$

(i) the constant cone fields

$$\mathcal{C}_p^u := \left\{ (\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathbb{R}^{L+M_u+M_s} \setminus \{0\} : \frac{\|\mathbf{v}\|_{p, M_s}}{\|\mathbf{u}\|_{p, L} + \|\mathbf{w}\|_{p, M_u}} < \beta_{u, p} \right\} \quad (3.32)$$

and

$$\mathcal{C}_p^s := \left\{ (\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathbb{R}^{L+M_u+M_s} \setminus \{0\} : \frac{\|\mathbf{v}\|_{p, M_s}}{\|\mathbf{u}\|_{p, L} + \|\mathbf{w}\|_{p, M_u}} > \frac{1}{\beta_{s, p}} \right\} \quad (3.33)$$

with

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

satisfy $\forall \mathbf{z} \in \mathbb{T}^N$ $\mathbf{D}_\mathbf{z} \mathbf{F}_\varepsilon(\mathcal{C}^u) \subset \mathcal{C}^u$ and $\mathbf{D}_\mathbf{z} \mathbf{F}_\varepsilon^{-1}(\mathcal{C}^s) \subset \mathcal{C}^s$.

(ii) there exists $\bar{\sigma}$ and $\bar{\lambda}$ such that, for every $\mathbf{z} \in \mathbb{T}^N$

$$\frac{\|\mathbf{D}_\mathbf{z} \mathbf{F}_\varepsilon(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p}{\|(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p} \geq \bar{\sigma} > 1, \quad \forall (\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathcal{C}_p^u \quad (3.34)$$

$$\frac{\|\mathbf{D}_\mathbf{z} \mathbf{F}_\varepsilon(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p}{\|(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p} \leq \bar{\lambda} < 1, \quad \forall (\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathcal{C}_p^s. \quad (3.35)$$

Corollary 3.1. *Under the assumptions of the previous proposition, $\pi_u \circ \mathbf{F}_\varepsilon^n: \mathbb{T}^{L+M_u} \times \{0\} \rightarrow \mathbb{T}^{L+M_u}$ 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 previous proposition, because $\mathbb{T}^{L+M_u} \times \{0\}$ is tangent to the unstable cone. $\square$

Proof of Proposition 3.4.2. (i) The expression for the differential of the map $\mathbf{F}_\varepsilon$ is the same as in (A.2). Take $(\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathbb{R}^L \times \mathbb{R}^{M_u} \times \mathbb{R}^{M_s}$, and suppose $(\mathbf{u}', \mathbf{w}', \mathbf{v}')^t := \mathbf{D}_z \mathbf{F}_\varepsilon(\mathbf{u}, \mathbf{v}, \mathbf{w})^t$. Then

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

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

$$\begin{aligned}
\|\mathbf{u}'\|_p &\geq (\sigma - \mathcal{O}(\Delta^{-1}\delta)) \|\mathbf{u}\|_p - \mathcal{O}(\Delta^{-1}N^{1/p}) \max_i \left[\sum_{\ell=1}^L A_{i\ell}^{ll} |u_\ell| + \sum_{m=1}^{M_u} A_{im}^{lh} |w_m| \right. \\
&\quad \left. + \sum_{m=M_u+1}^M A_{im}^{lh} |v_m| \right] \\
&\geq (\sigma - \mathcal{O}(\Delta^{-1}\delta)) \|\mathbf{u}\|_p - \mathcal{O}(\Delta^{-1}L^{1/p}\delta^{1/q})(\|\mathbf{u}\|_p + \|\mathbf{w}\|_p + \|\mathbf{v}\|_p)
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{w}'\|_p &\geq (\lambda^{-1} - C_\# \varepsilon - \mathcal{O}(\Delta^{-1}M)) \|\mathbf{w}\|_p - \\
&\quad - \mathcal{O}(\Delta^{-1}M_u^{1/p}) \max_{1 \leq j \leq M_u} \left[\sum_{\ell=1}^L A_{j\ell}^{hl} |u_\ell| + \sum_{m=1}^{M_u} A_{jm}^{hh} |w_m| + \sum_{m=M_u+1}^M A_{jm}^{hh} |v_m| \right] + \\
&\geq (\lambda^{-1} - C_\# \delta - \mathcal{O}(\Delta^{-1}M)) \|\mathbf{w}\|_p - \mathcal{O}(\Delta^{-1/p}M_u^{1/p})(\|\mathbf{u}\|_p + \|\mathbf{w}\|_p + \|\mathbf{v}\|_p)
\end{aligned}$$

and analogously

$$\|\mathbf{v}'\|_p \leq (\lambda + C_\# \varepsilon + \mathcal{O}(\Delta^{-1}M)) \|\mathbf{v}\|_p + \mathcal{O}(\Delta^{-1/p}M_s^{1/p})(\|\mathbf{u}\|_p + \|\mathbf{w}\|_p + \|\mathbf{v}\|_p)$$

Suppose that $(\mathbf{u}, \mathbf{w}, \mathbf{v})$ satisfies the cone condition $\|\mathbf{u}\|_p + \|\mathbf{w}\|_p \geq \tau\|\mathbf{v}\|_p$ for some τ . Then

$$\begin{aligned} \frac{\|\mathbf{u}'\|_p + \|\mathbf{w}'\|_p}{\|\mathbf{v}'\|_p} &\geq \frac{\mathcal{F}_{11}(\|\mathbf{u}\|_p + \|\mathbf{w}\|_p) - \mathcal{F}_{12}\|\mathbf{v}\|_p}{\mathcal{F}_{21}\|\mathbf{v}\|_p + \mathcal{F}_{22}(\|\mathbf{u}\|_p + \|\mathbf{w}\|_p)} \\ &\geq \frac{\mathcal{F}_{11} - \tau^{-1}\mathcal{F}_{12}}{\tau^{-1}\mathcal{F}_{21} + \mathcal{F}_{22}} \end{aligned}$$

with

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

where we used (H1)-(H4). The cone $\mathcal{C}_p^u$ is forward invariant iff $\|\mathbf{u}'\|_p + \|\mathbf{w}'\|_p \geq \tau\|\mathbf{v}'\|_p$ and therefore iff

$$\mathcal{F}_{11} - \tau^{-1}\mathcal{F}_{12} \geq \mathcal{F}_{21} + \tau\mathcal{F}_{22}. \quad (3.36)$$

Hence we find $C_* > 0$, so that if $\tau = C_*\mathcal{F}_{12}$ the inequality (3.36) is satisfied provided (3.31) 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 $\|\mathbf{u}'\|_p + \|\mathbf{w}'\|_p \leq \tau\|\mathbf{v}'\|_p$, thus

$$\begin{aligned} \mathcal{F}_{11} \frac{\|\mathbf{u}\|_p + \|\mathbf{w}\|_p}{\|\mathbf{v}\|_p} - \mathcal{F}_{12} &\leq \tau\mathcal{F}_{21} + \tau\mathcal{F}_{22} \frac{\|\mathbf{u}\|_p + \|\mathbf{w}\|_p}{\|\mathbf{v}\|_p} \\ \frac{\|\mathbf{u}\|_p + \|\mathbf{w}\|_p}{\|\mathbf{v}\|_p} &\leq \frac{\mathcal{F}_{12} + \tau\mathcal{F}_{21}}{\mathcal{F}_{11} - \tau\mathcal{F}_{22}} \end{aligned}$$

and imposing, yet again,

$$\tau^{-1}\mathcal{F}_{12} + \mathcal{F}_{21} \leq \mathcal{F}_{11} - \tau\mathcal{F}_{22}, \quad (3.37)$$

implies that $\|\mathbf{u}\|_p + \|\mathbf{w}\|_p \leq \tau\|\mathbf{v}\|_p$. Taking $\tau = C_*/\mathcal{F}_{22}$ with $C_* > 0$ small, we obtain that $\mathcal{C}_p^s$ is backward invariant (provided as before that (3.31) holds and $\eta > 0$ is small).

(ii) Take $(\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathcal{C}_p^u$ such that $\|(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p = 1$. From the above computa-

tions, and applying the cone condition

$$\begin{aligned}
\|\mathbf{u}'\|_p + \|\mathbf{w}'\|_p + \|\mathbf{v}'\|_p &\geq \|\mathbf{u}'\|_p + \|\mathbf{w}'\|_p \\
&\geq \mathcal{F}_{11}(\|\mathbf{u}\|_p + \|\mathbf{w}\|_p) - \mathcal{F}_{12}\|\mathbf{v}\|_p \\
&\geq \mathcal{F}_{11}(1 - \beta_{u,p}) - \mathcal{F}_{12}\beta_{u,p} \\
&\geq \min\{\sigma, \lambda^{-1} - C_{\#}\varepsilon\} - \mathcal{O}(\Delta^{-1/p}\delta^{1/p}) - \mathcal{O}(\Delta^{-1}L^{1/p}\delta^{1/q}) \\
&\geq \min\{\sigma, \lambda^{-1} - C_{\#}\varepsilon\} - \eta
\end{aligned} \tag{3.38}$$

where to obtain (3.38) 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 in conditions (H1)-(H4), if $\eta > 0$ is sufficiently small, (3.34) will be satisfied. Choosing, now, $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathcal{C}_p^s$ of unit norm we get

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

and again whenever conditions (H1)-(H4) are satisfied with $\eta > 0$ sufficiently small, (3.35) is verified. $\square$

3.4.3 Admissible Manifolds for $\mathbf{F}_\varepsilon$

As in the diffeomorphism case, the existence of the stable and unstable cone fields implies that the endomorphism $\mathbf{F}_\varepsilon$ admits a natural measure.

To determine the measure of the set $\mathcal{B}_\varepsilon$ with respect to 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 (Appendix A.3) which implies that manifolds tangent to the unstable cone which are locally graph 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 3.3 (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 $\mathbf{E} : I^{L+M_u} \rightarrow \mathcal{R}$ with Lipschitz differential so that

- W is the graph $(id, \mathbf{E})(I^{L+M_u})$ of E ,
- $\mathbf{D}_{\mathbf{z}_u} \mathbf{E}(\mathbb{R}^{L+M_u}) \subset \mathcal{C}_p^u$, $\forall \mathbf{z}_u \in I^{L+M_u}$,
- and

$$\|\mathbf{D}\mathbf{E}\|_{\text{Lip}} := \sup_{\mathbf{z}_u \neq \bar{\mathbf{z}}_u} \frac{\|\mathbf{D}_{\mathbf{z}_u} \mathbf{E} - \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E}\|_p}{d_p(\mathbf{z}_u, \bar{\mathbf{z}}_u)} \leq K_0.$$

Proposition 3.7. *Under conditions (H1)-(H4), for $\eta > 0$ sufficiently small, there is K_u uniform on the network parameters such that for all $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{S}$*

$$\|\mathbf{D}_{\mathbf{z}_1} \mathbf{F}_\varepsilon - \mathbf{D}_{\mathbf{z}_2} \mathbf{F}_\varepsilon\|_{u,p} := \sup_{(\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathcal{C}_p^u} \frac{\|(\mathbf{D}_{\mathbf{z}_1} \mathbf{F}_\varepsilon - \mathbf{D}_{\mathbf{z}_2} \mathbf{F}_\varepsilon)(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p}{\|(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p} \leq K_u d_\infty(\mathbf{z}_1, \mathbf{z}_2).$$

Proof. Notice that from the regularity assumptions on the coupling function h , we can write the entries $[\mathbf{D}_{\mathbf{z}_1} \mathbf{F}_\varepsilon - \mathbf{D}_{\mathbf{z}_2} \mathbf{F}_\varepsilon]_{k\ell} =$

$$= \begin{cases} \left[\sum_{\ell=1}^L \mathcal{O}(\Delta^{-1}) A_{k\ell}^{ll} + \sum_{m=1}^M \mathcal{O}(\Delta^{-1}) A_{km}^{lh} \right] d_\infty(\mathbf{z}_1, \mathbf{z}_2) & k = \ell \leq L \\ \mathcal{O}(\Delta^{-1}) A_{k\ell}^{ll} d_\infty(\mathbf{z}_1, \mathbf{z}_2) & k \neq \ell \leq L \\ \mathcal{O}(\Delta^{-1}) A_{k(\ell-L)}^{lh} d_\infty(\mathbf{z}_1, \mathbf{z}_2) & k \leq L, \ell > L \\ \mathcal{O}(\Delta^{-1}) A_{k\ell}^{hl} d_\infty(\mathbf{z}_1, \mathbf{z}_2) & k > L, \ell \leq L \\ \mathcal{O}(\Delta^{-1}) A_{(k-L)(\ell-L)}^{hh} d_\infty(\mathbf{z}_1, \mathbf{z}_2) & k \neq \ell > L \\ [\mathcal{O}(1) + \mathcal{O}(\Delta^{-1}M)] d_\infty(\mathbf{z}_1, \mathbf{z}_2) & k = \ell > L \end{cases} \quad (3.39)$$

Take $(\mathbf{u}, \mathbf{w}, \mathbf{v}) \in \mathcal{C}_p^u$ such that $\|(\mathbf{u}, \mathbf{w}, \mathbf{v})\|_p = 1$ and $(\mathbf{u}', \mathbf{w}', \mathbf{v}')^t = (\mathbf{D}_{\mathbf{z}_1} \mathbf{F}_\varepsilon - \mathbf{D}_{\mathbf{z}_2} \mathbf{F}_\varepsilon)(\mathbf{u}, \mathbf{w}, \mathbf{v})^t$.

$$\begin{aligned} u'_i &= \mathcal{O}(\Delta^{-1}) \left[\sum_{\ell=1}^L A_{k\ell}^{ll} + \sum_{m=1}^M A_{km}^{lh} \right] u_i d_\infty(\mathbf{z}_1, \mathbf{z}_2) + \\ &\quad + \mathcal{O}(\Delta^{-1}) \left[\sum_{\ell=1}^L A_{i\ell}^{ll} u_n + \sum_{m=1}^{M_u} A_{im}^{lh} w_m + \sum_{m=1}^{M_s} A_{i(m+M_u)}^{lh} v_m \right] d_\infty(\mathbf{z}_1, \mathbf{z}_2) \quad 1 \leq i \leq L \\ w'_j &= [\mathcal{O}(1) + \mathcal{O}(\Delta^{-1}M)] w_j d_\infty(\mathbf{z}_1, \mathbf{z}_2) \\ &\quad + \mathcal{O}(\Delta^{-1}) \left[\sum_i A_{ji}^{hl} u_i + \sum_{m=1}^{M_u} A_{jm}^{hh} w_m + \sum_{m=1}^{M_s} A_{j(m+M_u)}^{hh} v_m \right] d_\infty(\mathbf{z}_1, \mathbf{z}_2) \quad 1 \leq j \leq M_u \\ v'_j &= [\mathcal{O}(1) + \mathcal{O}(\Delta^{-1}M)] v_j d_\infty(\mathbf{z}_1, \mathbf{z}_2) + \\ &\quad + \mathcal{O}(\Delta^{-1}) \left[\sum_{\ell=1}^L A_{j\ell}^{hl} u_\ell + \sum_{m=1}^{M_u} A_{jm}^{hh} w_m + \sum_{m=M_u+1}^M A_{jm}^{hh} v_m \right] d_\infty(\mathbf{z}_1, \mathbf{z}_2) \quad 1 \leq j \leq M_s \end{aligned}$$

$$\|\mathbf{u}'\|_p \leq \mathcal{O}(\Delta^{-1}\delta N^{1/p})d_\infty(\mathbf{z}_1, \mathbf{z}_2) = \mathcal{O}(\eta)d_\infty(\mathbf{z}_1, \mathbf{z}_2)$$

$$\|\mathbf{w}'\|_p \leq \mathcal{O}(1)d_\infty(\mathbf{z}_1, \mathbf{z}_2)$$

$$\|\mathbf{v}'\|_p \leq \mathcal{O}(1)d_\infty(\mathbf{z}_1, \mathbf{z}_2)$$

which implies the proposition. $\square$

Lemma 3.2. *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 $n \in \mathbb{N}$, $\mathbf{F}_\varepsilon^n(W)$ 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 $\mathbf{F}_\varepsilon^n(W) \setminus \cup\{W_{n,k}\}_{k \in \mathcal{K}_n}$ consists of finite union of manifolds of lower dimension).*

Proof. As in Corollary 3.1, since $\pi_u|_{W_0}$ is a diffeomorphism, the map $\pi_u \circ \mathbf{F}_\varepsilon^n \circ \pi_u|_{W_0}^{-1} : \mathbb{T}^{L+M_u} \rightarrow \mathbb{T}^{L+M_u}$ is a well defined local diffeomorphism between compact manifolds, and, therefore, 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 $\pi_u \circ \mathbf{F}_\varepsilon^n \circ \pi_u|_{W_0}^{-1}(R_{n,k}) = \mathbb{T}^{L+M_u}$ and, defining $W_{n,k} := \pi_u \circ \mathbf{F}_\varepsilon^n \circ \pi_u|_{W_0}^{-1}(R_{n,k}^o)$, where $R_{n,k}^o$ is the interior of $R_{n,k}$, $\pi_u(W_{n,k}) = I^{L+M_u}$. From Proposition A.2 in Appendix A 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. $\square$

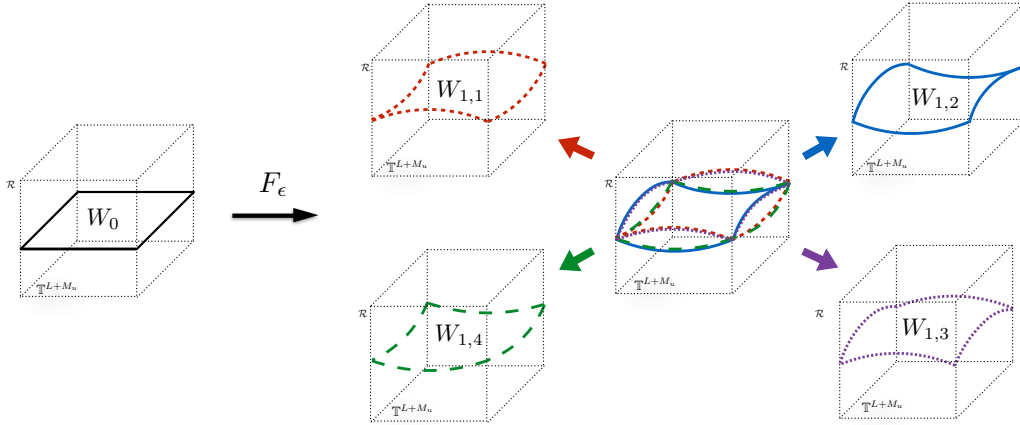


Figure 3.3: The admissible manifold W_0 is mapped under F_ε to the union of sub-manifolds $W_{1,1}$, $W_{1,2}$, $W_{1,3}$, and $W_{1,4}$.

3.4.4 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 $\mathbf{E} : I^{L+M_u} \rightarrow \mathcal{S}$, for every $\mathbf{z}_u \in I^{L+M_u}$ the map

$$\pi_u \circ \mathbf{F}_\varepsilon \circ (id, \mathbf{E})(\mathbf{z}_u)$$

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

$$J(\mathbf{z}_u) = |\Pi_u \cdot \mathbf{D}_{(id, \mathbf{E})(\mathbf{z}_u)} \mathbf{F}_\varepsilon \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})|.$$

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

Proposition 3.8. *Let $W \in W_{p,L}$ be an admissible manifold and suppose $\mathbf{z}_u, \bar{\mathbf{z}}_u \in I^{L+M_u}$, then*

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

Proof.

$$\begin{aligned} \frac{|J(\mathbf{z}_u)|}{|J(\bar{\mathbf{z}}_u)|} &= \frac{|\Pi_u \cdot \mathbf{D}_{(id, \mathbf{E})(\mathbf{z}_u)} \mathbf{F}_\varepsilon \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})|}{|\Pi_u \cdot \mathbf{D}_{(id, \mathbf{E})(\bar{\mathbf{z}}_u)} \mathbf{F}_\varepsilon \cdot (\text{Id}, \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})|} \frac{|\Pi_u \cdot \mathbf{D}_{(id, \mathbf{E})(\bar{\mathbf{z}}_u)} \mathbf{F}_\varepsilon \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})|}{|\Pi_u \cdot \mathbf{D}_{(id, \mathbf{E})(\bar{\mathbf{z}}_u)} \mathbf{F}_\varepsilon \cdot (\text{Id}, \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})|} \\ &=: (A) \cdot (B) \end{aligned}$$

(A) can be bounded with computations similar to the ones carried on in Proposition 4.6:

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

To estimate (B) we also factor out the number $Df = \sigma$ from the first L columns of $\Pi_u \mathbf{D}_{(id, \mathbf{E})(\bar{\mathbf{z}}_u)} \mathbf{F}_\varepsilon$, and $Dg_j(\bar{y}_{u,j})$ from the $(L+j)$ -th column when $j \in [M_u]$ and thus obtain

$$(B) = \frac{\sigma^L}{\sigma^L} \cdot \frac{\prod_{j=1}^{M_u} Dg_j}{\prod_{j=1}^{M_u} Dg_j} \cdot \frac{|\Pi_u \mathcal{D}((id, \mathbf{E})(\bar{\mathbf{z}}_u)) \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})|}{|\Pi_u \mathcal{D}((id, \mathbf{E})(\bar{\mathbf{z}}_u)) \cdot (\text{Id}, \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})|}$$

where $\mathcal{D}(\cdot)$ is the same matrix defined in (3.23) apart from the last M_s columns which are kept equal to the corresponding columns of $\mathbf{D} \cdot \mathbf{F}_\varepsilon$. The first two ratios trivially cancel. For the third factor we proceed in a fashion similar to previous computations using Proposition A.1 in the appendix. Defining $\mathbf{B} := \mathcal{D}((id, \mathbf{E})(\bar{\mathbf{z}}_u)) - \text{Id}$, we are reduced to estimate

$$\frac{|\text{Id} + \Pi_u \cdot \mathbf{B} \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})|}{|\text{Id} + \Pi_u \cdot \mathbf{B} \cdot (\text{Id}, \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})|}$$

where we used that $\Pi_u D \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E}) - \text{Id} = \Pi_u \mathbf{B} \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})$.

Since $\|(\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})\|_p \leq (1 + \beta_{u,p})$ for any $\mathbf{z}_u \in \mathcal{S}$, it follows, choosing $\eta > 0$ sufficiently small in (H1)-(H4) and from equation (3.25) that the operator norm

$$\|\Pi_u \cdot \mathbf{B} \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})\|_p < \lambda < 1 \quad (3.40)$$

It is also rather immediate to upper bound the column norms of $\Pi_u \cdot \mathbf{B} \cdot (0, \mathbf{D}_{\mathbf{z}_u} \mathbf{E} - \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})$ and obtain

$$\begin{aligned} \|\text{Col}^i[\Pi_u \mathbf{B}(0, \mathbf{D}_{\mathbf{z}_u} \mathbf{E} - \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})]\|_p &\leq \mathcal{O}(\Delta^{-1} L^{1/p} \delta^{1/q}) \|\mathbf{D}\mathbf{E}\|_{\text{Lip},p} d_p(\mathbf{z}_u, \bar{\mathbf{z}}_u) \\ &\leq \mathcal{O}(\Delta^{-1} M^{1/p}) d_p(\mathbf{z}_u, \bar{\mathbf{z}}_u) \end{aligned}$$

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

$$\frac{|\Pi_u \cdot \mathcal{D}((id, \mathbf{E})(\bar{\mathbf{z}}_u)) \cdot (\text{Id}, \mathbf{D}_{\mathbf{z}_u} \mathbf{E})|}{|\Pi_u \cdot \mathcal{D}((id, \mathbf{E})(\bar{\mathbf{z}}_u)) \cdot (\text{Id}, \mathbf{D}_{\bar{\mathbf{z}}_u} \mathbf{E})|} \leq \exp \left\{ \mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) d_p(\mathbf{z}_u, \bar{\mathbf{z}}_u) \right\}. \quad (3.41)$$

□

3.4.5 Invariant Cone of Densities on Admissible Manifolds for $\mathbf{F}_\varepsilon$

Take $W \in \mathcal{W}_{p,K_0}$. A density ρ on W is a measurable function $\varphi : W \rightarrow \mathbb{R}^+$ such that the integral of φ over W with respect to m_W , 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 $\varphi_u : \mathbb{T}^{L+M_u} \rightarrow \mathbb{R}^+$ is well defined.

Definition 3.4. For every $W \in \mathcal{W}_{p,K_0}$ and for every $\varphi : W \rightarrow \mathbb{R}^+$ we define φ_u as

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

Consider the set of densities

$$\mathcal{C}_{a,p}(W) := \left\{ \varphi : W \rightarrow \mathbb{R}^+ \text{ s.t. } \frac{\varphi_u(\mathbf{z}_u)}{\varphi_u(\bar{\mathbf{z}}_u)} \leq \exp[ad_p(\mathbf{z}_u, \bar{\mathbf{z}}_u)] \right\}.$$

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

Proposition 3.9. For every $a > a_c$, where

$$a_c = \mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M) \quad (3.42)$$

$W \in \mathcal{W}_{p,K_0}$ and $\varphi \in \mathcal{C}_{a,p}(W)$ the following holds. Suppose that $\{W'_k\}_k$ is the partition of $\mathbf{F}_\varepsilon(W)$ given by Lemma 3.2 and that W_k is a manifold of W such that $\mathbf{F}_\varepsilon(W_k) = W'_k \in \mathcal{W}_{p,K_0}$, 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 $\mathbf{G}_k$ be the inverse of the map $\mathbf{F}_\varepsilon|_{W_k} : W_k \rightarrow W'_k$. From Definition 3.4 follows that

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

where $\mathbf{E}$ is the map whose graph equals W . This implies that

$$(\varphi'_k)_u(z_u) = \frac{\varphi_u(\mathbf{G}_k(\mathbf{z}_u))}{J(\mathbf{G}_k(\mathbf{z}_u))}$$

and therefore

$$\begin{aligned} \frac{(\varphi'_k)_u(\mathbf{z}_u)}{(\varphi'_k)_u(\bar{\mathbf{z}}_u)} &= \frac{\varphi_u(\mathbf{G}_k(\mathbf{z}_u))}{\varphi_u(\mathbf{G}_k(\bar{\mathbf{z}}_u))} \frac{J(\mathbf{G}_k(\bar{\mathbf{z}}_u))}{J(\mathbf{G}_k(\mathbf{z}_u))} \\ &\leq \exp[\bar{\sigma}^{-1} ad_p(\mathbf{z}_u, \bar{\mathbf{z}}_u)] \exp\left\{[\mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M)] d_p(\mathbf{z}_u, \bar{\mathbf{z}}_u)\right\} \\ &\leq \exp\left\{[\bar{\sigma}^{-1} a + \mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M)] d_p(\mathbf{z}_u, \bar{\mathbf{z}}_u)\right\}. \end{aligned}$$

Taking a_c as in (3.42), the proposition is verified. □

At this point we can prove that the system admits invariant physical measures and that their marginals on the first $L + M_u$ coordinates is in the cone $\mathcal{C}_{a,p}$ for $a > a_c$:

Proof of Theorem 3.2. Pick a periodic orbit, $O(\mathbf{z}_s)$ 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 $O(\mathbf{z}_s)$. Pick $\mathbf{y}_s \in U$ and take the admissible manifold $W_0 := \mathbb{T}^{L+M_u} \times \{\mathbf{y}_s\} \in \mathcal{W}_{p,K_0}$. 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 of measures $\{\mu_t\}_{t \in \mathbb{N}_0}$ defined as

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

From Lemma 3.2 we know that $\mathbf{F}_\varepsilon^i(W_0) = \bigcup_{k \in \mathcal{K}_i} W_{i,k}$ modulo a negligible set w.r.t. $\mathbf{F}_{\varepsilon*}^i(\mu_0)$, and that

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

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 3.9 that $\mu_{ik} = \rho_{i,k} \cdot m_{W_{i,k}}$ with $\rho_{i,k} \in C_{a,p}(W_{i,k})$. Since $\mathbf{F}_\varepsilon$ 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 $\bar{\mu}$ be a probability measure which is a limit of a converging subsequence. By convexity of the cone $\mathcal{C}_{a,p}$ the second assertion of the theorem holds for $\bar{\mu}$.

Now let $V_1, \dots, V_n$ be the components of U where n is the period of $O(\mathbf{z}_s)$. Since all stable manifolds are tangent to a constant cone which have 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 $\mathbb{T}^{L+M_u} \times \{\mathbf{y}\}$ with $\mathbf{y} \in V_k$. It follows from the standard arguments that $\bar{\mu}$ 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 set (see [QXZ09] for details). Following the standard Hopf argument ([Wil12, KH95]), one first notice that fixed a point $\mathbf{x} \in V_i$ on the support of $\bar{\mu}$, a history $\bar{x} \in (\mathbb{T}^N)^\mathbb{N}$, and a continuous observable φ , from the definition of $\bar{\mu}$, almost every point on the local unstable manifold associated to the selected history have a well defined forward asymptotic Birkhoff average (computed along $\bar{x}$) 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 $\bar{\mu}$ which implies ergodicity. $\square$

3.4.6 Jacobian of the Holonomy Map along Stable Leaves of $\mathbf{F}_\varepsilon$

In order to prove Proposition 3.5 we need to upper bound the Jacobian of the holonomy map along stable leaves. It is known that for a C^2 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. This can be easily generalised to the non-invertible case. First of all, let us recall the definition of holonomy map. We will consider this map between manifolds tangent to the unstable cone.

Definition 3.5. Given D_1 and D_2 embedded disks of dimension $L + M_u$, tangent

to the unstable cone $\mathcal{C}^u$, we define the holonomy map $\pi : D_1 \rightarrow D_2$

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

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

Remark 3.5. For the truncated dynamical system $\mathbf{F}_\varepsilon$, fixing D_1 , one can always find a sufficiently large D_2 such that the map is well defined everywhere in D_1 .

Proposition 3.10. *Given D_1 and D_2 admissible embedded disks tangent to the unstable cone, the holonomy map $\pi : D_1 \rightarrow D_2$ associated to $\mathbf{F}_\varepsilon$ is absolutely continuous with respect to m_{D_1} and m_{D_2} which are the restrictions of Lebesgue measure to the two embedded disks. Furthermore, if J_s is the Jacobian of π , then*

$$J_s(\mathbf{z}) \leq \exp \left\{ [\mathcal{O}(\Delta^{-1}L\delta) + \mathcal{O}(M)] \frac{1}{1-\lambda} d_\infty(\mathbf{z}, \pi(\mathbf{z})) \right\}, \quad \forall \mathbf{z} \in D_1. \quad (3.43)$$

Proof. The absolute continuity follows from results in [Mn87] (see Appendix A.3), as well as the estimate on the Jacobian. In fact it is proven in [Mn87] that

$$J_s(\mathbf{z}) = \prod_{k=0}^{\infty} \frac{\text{Jac}(\mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon | V_k)}{\text{Jac}(\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon | \bar{V}_k)} \quad \forall \mathbf{z} \in D_1,$$

where $\mathbf{z}_k := \mathbf{F}_\varepsilon^k(\mathbf{z})$, $\bar{\mathbf{z}}_k := \mathbf{F}_\varepsilon^k(\pi(\mathbf{z}))$, $V_k := T_{\mathbf{z}_k} \mathbf{F}_\varepsilon^k(D_1)$ and $\bar{V}_k := T_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon^k(D_2)$. Since D_1 and D_2 are tangent to the unstable cone, one can write $\mathbf{F}_\varepsilon^k(D_1)$ and $\mathbf{F}_\varepsilon^k(D_2)$ locally as graphs of functions $\mathbf{E}_{1,k} : B_\delta^u(\mathbf{z}_k) \rightarrow B_\delta^s(\mathbf{z}_k)$ and $\mathbf{E}_{2,k} : B_\delta^u(\bar{\mathbf{z}}_k) \rightarrow B_\delta^s(\bar{\mathbf{z}}_k)$, with $\mathbf{E}_{i,k}$ given by application of the graph transform on $\mathbf{E}_{i,k-1}$, and $\mathbf{E}_{i,0}$ such that $(\text{Id}, \mathbf{D}_{\mathbf{z}} \mathbf{E}_{i,0})(\mathbb{R}^{L+M_u}) = T_{\mathbf{z}} \mathbf{D}_i$.

For every $k \in \mathbb{N} \cup \{0\}$, $(\text{Id}, D_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k}) \circ \Pi_u|_{V_k} = \text{Id}|_{V_k}$,

$$\text{Jac}(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k}) \text{Jac}(\Pi_u|_{V_k}) = 1$$

and analogously

$$\text{Jac}(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k}) \text{Jac}(\Pi_u|_{\bar{V}_k}) = 1.$$

Since

$$\begin{aligned} |\Pi_u \circ \mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})| &= \text{Jac}(\Pi_u \circ \mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})) \\ &= \text{Jac}(\Pi_u|_{V_k}) \text{Jac}(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k}) \text{Jac}(\mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon | V_k) \end{aligned}$$

and, analogously,

$$\begin{aligned} |\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k})| &= \text{Jac}(\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k})) \\ &= \text{Jac}(\Pi_u|_{\bar{V}_k}) \text{Jac}(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k}) \text{Jac}(\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|_{\bar{V}_k}). \end{aligned}$$

So

$$\begin{aligned} J_s(\mathbf{z}) &= \prod_{k=0}^{\infty} \frac{\text{Jac}(\mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon|_{V_k})}{\text{Jac}(\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|_{V_k})} \frac{\text{Jac}(\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|_{V_k})}{\text{Jac}(\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|_{\bar{V}_k})} \\ &= \prod_{k=0}^{\infty} \frac{|\Pi_u \circ \mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|}{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|} \frac{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|}{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k})|}. \end{aligned}$$

The first ratio can be deduced with minor changes from the estimate (3.41) in the proof of Proposition 3.8. So

$$\begin{aligned} \prod_{k=0}^{\infty} \frac{|\Pi_u \circ \mathbf{D}_{\mathbf{z}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|}{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|} &\leq \exp \left\{ [\mathcal{O}(\Delta^{-1}L\delta) + \mathcal{O}(M)] \sum_{k=0}^{\infty} d_p(\mathbf{z}_k, \bar{\mathbf{z}}_k) \right\} \\ &\leq \exp \left\{ [\mathcal{O}(\Delta^{-1}L\delta) + \mathcal{O}(M)] \frac{1}{1-\bar{\lambda}} d_p(\mathbf{z}_0, \bar{\mathbf{z}}_0) \right\} \end{aligned}$$

where we used the fact that $\mathbf{z}_0$ and $\bar{\mathbf{z}}_0$ lay on the same stable manifold and, by Proposition 3.6: $d_p(\mathbf{z}_k, \bar{\mathbf{z}}_k) \leq \bar{\lambda}^k d_p(\mathbf{z}_0, \bar{\mathbf{z}}_0)$.

To estimate the other ratio we proceed making similar computations leading to the estimate in (3.41). Once more we factor out σ from the first N columns of $\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon$ and, for all $j \in [M]$, $D_{y_j} g(y_j)$ from the $(L+j)$ -th column and thus

$$\begin{aligned} \frac{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|}{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k})|} &= \frac{\sigma^L}{\sigma^L} \cdot \frac{\prod_{j=1}^{M_u} Dg_j |\Pi_u \circ \mathbf{D}(\bar{\mathbf{z}}_k)(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|}{\prod_{j=1}^{M_u} Dg_j |\Pi_u \circ \mathbf{D}(\bar{\mathbf{z}}_k)(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k})|} \\ &= \frac{|\Pi_u \circ \mathcal{D}(\bar{\mathbf{z}}_k)(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k})|}{|\Pi_u \circ \mathcal{D}(\bar{\mathbf{z}}_k)(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k})|} \end{aligned}$$

where $\mathcal{D}(\mathbf{z}_k)$ is defined as in (3.23). Defining for every $k \in \mathbb{N}$

$$\begin{aligned} \mathbf{B}_k &:= \Pi_u \mathcal{D}(\bar{\mathbf{z}}_k)(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k}) - \text{Id} \\ &= \Pi_u (\mathcal{D}(\bar{\mathbf{z}}_k) - \text{Id})(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)} \mathbf{E}_{1,k}) \end{aligned}$$

and analogously

$$\begin{aligned} \bar{\mathbf{B}}_k &:= \Pi_u \mathcal{D}(\bar{\mathbf{z}}_k)(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k}) - \text{Id} \\ &= \Pi_u (\mathcal{D}(\bar{\mathbf{z}}_k) - \text{Id})(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)} \mathbf{E}_{2,k}). \end{aligned}$$

we have proved in (3.40) that $\|\bar{\mathbf{B}}_k\|, \|\mathbf{B}_k\| \leq \lambda < 1$. It remains to estimate the norm of the columns of $\bar{\mathbf{B}}_k - \mathbf{B}_k$. For all $\ell \in [L]$

$$\begin{aligned} \|\text{Col}^\ell[\bar{\mathbf{B}}_k - \mathbf{B}_k]\| &= \|\text{Col}^\ell[\Pi_u(\mathcal{D}(\bar{\mathbf{z}}_k) - \text{Id})(0, \mathbf{D}_{\pi_u(\mathbf{z}_k)}\mathbf{E}_{1,k} - \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)}\mathbf{E}_{2,k})]\| \\ &\leq \|\Pi_u(\mathcal{D}(\bar{\mathbf{z}}_k) - \text{Id})|_{0 \oplus \mathbb{R}^{M_s}}\| d_u(\bar{V}_k, V_k) \\ &\leq \mathcal{O}(\Delta^{-1}\delta) d_u(\bar{V}_k, V_k). \end{aligned}$$

where we used that, as can be easily deduced from the definition of d_u in (A.8) of Appendix A.3, that $\|(0, \mathbf{D}_{\pi_u(\mathbf{z}_k)}\mathbf{E}_{1,k} - \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)}\mathbf{E}_{2,k})\| = d_u(\bar{V}_k, V_k)$. By Proposition A.1 in Appendix A.2, we obtain

$$\frac{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)}\mathbf{E}_{1,k})|}{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)}\mathbf{E}_{2,k})|} \leq \exp \left\{ \mathcal{O}(L\Delta^{-1}\delta) d_u(\bar{V}_k, V_k) \right\}.$$

By Proposition A.4 we know that, if β_u is sufficiently small, then $d_u(\bar{V}_k, V_k) \leq \lambda_* d_u(V_0, W_0)$ for some $\lambda_* < 1$, and this implies that

$$\begin{aligned} \prod_{k=0}^{\infty} \frac{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\mathbf{z}_k)}\mathbf{E}_{1,k})|}{|\Pi_u \circ \mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon(\text{Id}, \mathbf{D}_{\pi_u(\bar{\mathbf{z}}_k)}\mathbf{E}_{2,k})|} &\leq \exp \left\{ \mathcal{O}(L\Delta^{-1}\delta) \sum_{k=0}^{\infty} d_u(\bar{V}_k, V_k) \right\} \\ &\leq \exp \left\{ \mathcal{O}(L\Delta^{-1}\delta) d_u(V_0, W_0) \right\} \\ &\leq \exp \left\{ \mathcal{O}(L\Delta^{-1}\delta\beta_u) \right\}. \end{aligned}$$

□

3.4.7 Proof of Proposition 3.5

The following result shows that the set $\mathcal{B}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}$, which is the set where fluctuations of the dynamics of a given hub exceed a given threshold, is contained in a set, $\tilde{\mathcal{B}}_\varepsilon^{(s,j)}$, that is the union of global stable manifolds. This is important to notice because, even if the product structure of the former set is not preserved taking preimages under $\mathbf{F}_\varepsilon$, the preimage of $\tilde{\mathcal{B}}_\varepsilon^{(s,j)}$ 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 $\tilde{\mathcal{B}}_\varepsilon^{(s,j)}$ will be “close” (topologically and with respect to the right measures) to $\mathcal{B}_\varepsilon^{(s,j)}$.

Lemma 3.3. *Consider $\mathcal{B}_\varepsilon^{(s,j)}$ as in (3.15). Then there exist a constant $C > 0$ such*

that

$$\mathcal{B}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R} \subset \widetilde{\mathcal{B}}_\varepsilon^{(s,j)} := \bigcup_{z \in \mathcal{B}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}} W^s(z) \subset \mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}, \quad (3.44)$$

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

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

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

Since $W^s(\mathbf{z})$ is tangent to the stable cone $\mathcal{C}^s$, by (3.33) for any $\mathbf{z}' \in W^s(\mathbf{z})$, $d_p(\pi_u(\mathbf{z}'), \mathbf{z}) \leq \beta_{s,p}$. This implies that

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

proving the lemma. $\square$

Proof of Proposition 3.5. As in the proof of Theorem 3.2, take an embedded $L + M_u$ torus $W_0 \in \mathcal{W}_{p,K_0}$ such that $\pi_u|_{W_0} : W_0 \rightarrow \mathbb{T}^{L+M_u}$ is a diffeomorphism, a density $\rho \in \mathcal{C}_{a,p}(W_0)$ with $a > a_c$ so that $\rho \mu_W$ is a probability measure and the limit $\bar{\mu}$ of the sequence of measures $\{\mu_t\}_{t \in \mathbb{N}_0}$ defined as

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

is an SRB measure. From Lemma 3.2 we know that $\mathbf{F}_\varepsilon^i(W_0) = \bigcup_{k \in \mathcal{K}_i} W_{i,k}$ modulo a negligible set w.r.t. $\mathbf{F}_{\varepsilon*}^i(\mu_0)$, and that

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

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 3.9 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 $t \in \mathbb{N}_0$,

$$\begin{aligned}
 \mu_t(\tilde{\mathcal{B}}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}) &\leq \mu_t(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}) \\
 &= \sum_{i=0}^t \sum_{k \in \mathcal{K}_i} \frac{\mathbf{F}_{\varepsilon_*}^i \mu_0(W_{i,k})}{t+1} \mu_{i,k}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}) \\
 &= \sum_{i=0}^t \sum_{k \in \mathcal{K}_i} \frac{\mathbf{F}_{\varepsilon_*}^i \mu_0(W_{i,k})}{t+1} \int_{\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}} \rho_{i,k} dm_{L+M_u} \\
 &\leq \sum_{i=0}^t \sum_{k \in \mathcal{K}_i} \frac{\mathbf{F}_{\varepsilon_*}^i \mu_0(W_{i,k})}{t+1} \exp \left[\mathcal{O}(\Delta^{-1} L^{1+2/p} \delta^{1/q}) + \right. \\
 &\quad \left. + \mathcal{O}(M) \right] m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}) \tag{3.45}
 \end{aligned}$$

$$= \exp \left[\mathcal{O}(\Delta^{-1} L^{1+2/p} \delta^{1/q}) + \mathcal{O}(M) \right] m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}) \tag{3.46}$$

To prove the bound (3.45) we used 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 $\exp [\mathcal{O}(\Delta^{-1} L^{1+2/p} \delta^{1/q}) + \mathcal{O}(ML^{1/p})]$. (3.46) follows from the fact that

$$\sum_{i=0}^t \sum_{k \in \mathcal{K}_i} \frac{\mathbf{F}_{\varepsilon_*}^i \mu_0(W_{i,k})}{t+1} = \sum_{i=0}^t \frac{\mathbf{F}_{\varepsilon_*}^i \mu_0(\mathbf{F}_\varepsilon^i(W_0))}{t+1} = \sum_{i=0}^t \frac{1}{t+1} = 1. \tag{3.47}$$

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

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

and since $\bar{\mu}$ is invariant $\forall t \in \mathbb{N}$

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

From (3.47) there exists $i \in \mathbb{N}$ and $k \in \mathcal{K}_i$ such that

$$\mu_{i,k}(\mathbf{F}_\varepsilon^{-t}(\tilde{\mathcal{B}}_\varepsilon^{(s,j)})) \leq \bar{\mu}(\mathbf{F}_\varepsilon^{-t}(\tilde{\mathcal{B}}_\varepsilon^{(s,j)}))$$

and thus

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

Now, pick $y_s \in \mathcal{R}$ and consider the holonomy map along the stable leaves $\pi : W_{i,k} \rightarrow D_{y_s}$ 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 3.10 that the Jacobian of π is bounded by (3.43) and thus

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

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

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

and from the first inclusion in (3.44) we obtain

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

□

3.4.8 Proof of Theorem A

In this section $\mathbf{F}_\varepsilon : \mathbb{T}^N \rightarrow \mathbb{T}^N$ denotes again the truncated map defined on the whole phase space. Define the uncoupled map $\mathbf{f} : \mathbb{T}^N \rightarrow \mathbb{T}^N$

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

The next lemma evaluates the ratios of the Jacobians of $\mathbf{F}_\varepsilon^t$ and $\mathbf{f}^t$ for any fixed $t \in \mathbb{N}$.

Lemma 3.4.

$$\frac{|\mathbf{D}_{\mathbf{z}} \mathbf{f}^t|}{|\mathbf{D}_{\mathbf{z}} \mathbf{F}_\varepsilon^t|} \leq \exp \left[\mathcal{O}(tL\Delta^{-1}\delta) + \mathcal{O}(tM) \right]$$

Proof. For all $i \in [t]$ define $\mathbf{z}_i := \mathbf{f}^i(\mathbf{z})$, $\bar{\mathbf{z}}_i := \mathbf{F}_\varepsilon^i(\mathbf{z})$, and $\mathbf{z}_0 = \bar{\mathbf{z}}_0 := \mathbf{z}$.

$$\begin{aligned} \frac{|\mathbf{D}_{\mathbf{z}} \mathbf{f}^t|}{|\mathbf{D}_{\mathbf{z}} \mathbf{F}_\varepsilon^t|} &= \frac{\prod_{k=0}^t |\mathbf{D}_{\mathbf{z}_k} \mathbf{f}|}{\prod_{k=0}^t |\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|} = \prod_{k=0}^t \frac{\sigma^L \prod_{m=1}^M D_{y_{i,m}} g_m}{|\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|} \\ &= \prod_{k=0}^t \frac{\sigma^L \prod_{m=1}^M D_{\bar{y}_{i,m}} g_m \left(1 + \frac{D_{y_{i,m}} g_m - D_{\bar{y}_{i,m}} g_m}{D_{\bar{y}_{i,m}} g_m} \right)}{|\mathbf{D}_{\bar{\mathbf{z}}_k} \mathbf{F}_\varepsilon|} \\ &\leq \exp [\mathcal{O}(M)] \prod_{k=0}^t \frac{1}{|\mathcal{D}(\bar{\mathbf{z}}_k)|} \end{aligned}$$

where $\mathcal{D}(\mathbf{z}_i)$ is defined as in (3.23). $1/|\mathcal{D}(\bar{\mathbf{z}}_i)|$ can be estimated in the usual way defining $\mathbf{B}(\mathbf{z}_i) := \mathcal{D}(\mathbf{z}_i) - \text{Id}$ and noticing that $1 = |\text{Id} + 0|$. One can obtain, from

the computations leading to (3.26), that

$$\frac{1}{|\mathcal{D}(\bar{\mathbf{z}}_i)|} \leq \exp \left[\sum_{k=1}^N \text{Col}^k[\mathbf{B}(\mathbf{z}_i)] \right] \leq \exp [\mathcal{O}(L\Delta^{-1}\delta) + \mathcal{O}(M)].$$

and thus

$$\frac{|\mathbf{D}_{\mathbf{z}} \mathbf{f}^t|}{|\mathbf{D}_{\mathbf{z}} \mathbf{F}_{\varepsilon}^t|} \leq \exp [\mathcal{O}(M)] \exp [\mathcal{O}(tL\Delta^{-1}\delta) + \mathcal{O}(tM)] \leq \exp [\mathcal{O}(tL\Delta^{-1}\delta) + \mathcal{O}(tM)].$$

□

Lemma 3.5. *Consider the set $A^2(\mathbb{T}, \mathbb{T})$ of C^2 Axiom A endomorphisms on $\mathbb{T}$ endowed with the C^1 topology. Take a continuous curve $\gamma : [\alpha_1, \alpha_2] \rightarrow A^2(\mathbb{T}, \mathbb{T})$. Then, denoting by Λ^α and Υ^α respectively the attractor and repeller of γ_α for all $\alpha \in [\alpha_1, \alpha_2]$,*

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

$$\left| \gamma'_\alpha|_{\Lambda_{\varepsilon_\Lambda}^\alpha} \right| < \lambda \quad \text{and} \quad \left| \gamma'_\alpha|_{\Upsilon_{\varepsilon_\Lambda}^\alpha} \right| > \lambda^{-1}, \quad (3.48)$$

(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 $x \in \mathbb{T} \setminus \Upsilon_{\varepsilon_\Lambda}^\alpha$, the orbit $\{x_i\}_{i=0}^\tau$ defined*

$$x_0 := x \quad \text{and} \quad x_i := \gamma_\alpha(x_{i-1}) + \varepsilon_i, \quad (3.49)$$

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

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.

□

Proof of Theorem A. Step 1 Restricting $\mathbf{F}_\varepsilon$ to $\mathcal{S}$, we can use Proposition 3.5 to get an estimate of the Lebesgue measure of $\mathcal{B}_{\varepsilon, T} \times \mathbb{T}^{M_u} \times \mathcal{R}$. Define

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

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

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

For all $\mathbf{y} \in \mathbb{T}^M$, the map $\mathbf{f}|_{\mathbb{T}^L \times \{\mathbf{y}\}} : \mathbb{T}^L \times \{\mathbf{y}\} \rightarrow \mathbb{T}^L \times \{(g_1(y_1), \dots, g_M(y_M))\}$ is an expanding map with constant Jacobian and thus measure preserving if we endow $\mathbb{T}^L \times \{\mathbf{y}\}$ 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 $t \in [\tau]$ (where τ is uniform as in Lemma 3.4)

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

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

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

Now

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

By Lemma 3.4 we get

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

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

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

Consider the system $\mathbf{G} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ obtained redefining $\mathbf{F}_{\varepsilon}$ on $\mathcal{U}^c$ so that $\mathbf{G}$ is globally expanding. In particular, if $(\mathbf{x}', \mathbf{y}') = \mathbf{G}(\mathbf{x}, \mathbf{y})$, $\pi_u \circ \mathbf{G}(\mathbf{x}, \mathbf{y}) = \pi_u \circ \mathbf{F}_{\varepsilon}(\mathbf{x}, \mathbf{y})$ (the evolution of the “expanding” coordinates is unvaried) and for the hubs $j \in [M_u + 1, M]$ with an attractor

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

where the reduced dynamics is (smoothly) modified to be globally expanding by putting $\hat{f}_j|_{\Upsilon_{\varepsilon_{\Lambda}}^j} := g_j|_{\Upsilon_{\varepsilon_{\Lambda}}^j}$ (unvaried in a neighbourhood of the repeller) and $\hat{f}_j|_{\mathbb{T} \setminus \Upsilon_{\varepsilon_{\Lambda}}^j}$

redefined so that $|\hat{f}_j| \geq \lambda^{-1} > 1$ everywhere on $\mathbb{T}$. By definition, $G|_{\mathcal{U}} = F_\varepsilon|_{\mathcal{U}}$. We can then invoke the results of Section 3.3 to impose conditions on η and ε to deduce global expansion of the map $\mathbf{G}$ (under suitable heterogeneity hypotheses) and the bounds on the invariant density obtained in that section. In particular one has that for al $T \in \mathbb{N}$

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

and this implies

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

And this concludes the proof of the theorem. □

3.4.9 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 3.1 with $n = 1$. We now show that any $n \in \mathbb{N}$ will work.

Lemma 3.6. *It is enough to prove Theorem A for $n = 1$.*

Proof. Assume that g_j , $j = 1, \dots, M$ satisfies the assumptions in Definition 3.1 for some (n, λ, r) . Condition (2) and (3) imply that one can, suppose, that after smooth conjugation each of the maps g_j are such that, $|Dg_j(x)| < \lambda$ for all $x \in B_r(\Lambda_j)$, and $|Dg_j(x)| > \lambda^{-1}$ for all $x \in N_r(\Upsilon_j)$. These conjugations are obtained by changing the metric using the well-known ‘‘Mather trick’’. In other words, there exists a smooth coordinate change $\varphi_j: \mathbb{T} \rightarrow \mathbb{T}$ so that for $\tilde{g}_j := \varphi_j \circ g_j \circ \varphi_j^{-1}$ and $\tilde{g}_j$ satisfies Definition 3.1 for $n = 1$. Moreover, there exists some uniform constant $C_\#$ only depending on (n, λ, r) , so that the C^2 norms of φ_j and φ_j^{-1} are bounded by $C_\#$. Writing $\tilde{y}_j = \varphi_j(y_j)$ and $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n) = (x_1, \dots, x_L, \tilde{y}_1, \dots, \tilde{y}_M)$, in these new coordinates (3.2)-(3.5) become

$$x'_i = f(x_i) + \frac{\alpha}{\Delta} \sum_{\ell=1}^L A_{i\ell}^l h(x_i, x_\ell) + \frac{\alpha}{\Delta} \sum_{m=1}^M A_{im}^{lh} \tilde{h}(x_i, \tilde{y}_m) \mod 1 \quad i \in [L] \quad (3.50)$$

$$\tilde{y}'_j = \tilde{g}_j(\tilde{y}_j) + \tilde{\xi}_j(\tilde{\mathbf{z}}) \mod 1 \quad j \in [M] \quad (3.51)$$

where

$$\tilde{\xi}_j(z) := \int_{\tilde{g}_j(\tilde{y}_j)}^{\tilde{g}_j(\tilde{y}_j) + \xi_j} D_t \varphi_j dt, \quad \text{and} \quad \tilde{h}(\mathbf{x}, \tilde{\mathbf{y}}) := h(\mathbf{x}, \varphi_j^{-1}(\tilde{\mathbf{y}})). \quad (3.52)$$

In fact

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

Then we can define $\tilde{\xi}_{j,\varepsilon}$ as

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

and define the truncated system as

$$x'_i = f(x_i) + \frac{\alpha}{\Delta} \sum_{\ell=1}^L A_{i\ell}^{ll} h(x_i, x_\ell) + \frac{\alpha}{\Delta} \sum_{m=1}^M A_{im}^{lh} \tilde{h}(x_i, \tilde{y}_m) \mod 1 \quad i \in [L] \quad (3.53)$$

$$\tilde{y}'_j = \tilde{g}_j(\tilde{y}_j) + \tilde{\xi}_{j,\varepsilon}(\tilde{\mathbf{z}}) \mod 1 \quad j \in [M]. \quad (3.54)$$

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

$$|y'_j - g_j(y_j)| = \left| \varphi_j^{-1}[\tilde{g}_j(\varphi_j(y_j)) + \tilde{\xi}_{j,\varepsilon}(\tilde{\mathbf{z}})] - g_j(y_j) \right| \leq \mathcal{O}(\tilde{\xi}_{j,\varepsilon}(\tilde{\mathbf{z}})) \leq \mathcal{O}(\xi).$$

□

3.4.10 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 $\mathbf{F} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ we have a stationary Markov chain $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ on some probability space $(\Omega, \mathbb{P})$ with transition kernel

$$\mathbb{P}(\mathcal{F}_{n+1} \in A | \mathcal{F}_n = z) := \int_{A \subset \mathbb{T}^N} \varphi(\mathbf{y} - \mathbf{F}(\mathbf{z})) dm(\mathbf{y})$$

where $\varphi : \mathbb{T}^N \rightarrow \mathbb{R}^+$ 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 $\mathbf{F}$. Take now the stationary Markov chain

$\{\mathcal{F}_{\varepsilon,t}\}_{t \in \mathbb{N}}$ defined by the transition kernel

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

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 $\mathbf{F}_{\varepsilon}$ 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

$$(P_{\varphi}\rho)(\mathbf{x}) = \int \rho(\mathbf{y})\varphi(\mathbf{y} - \mathbf{x})dm(\mathbf{y}).$$

Let $C_{a,p}$ be a cone of densities invariant under P_{ε} as prescribed in Proposition 3.4. It is easy to see that this is also invariant under P_{φ} ³ and thus under $\mathcal{P}_{\varepsilon}$. This means that there exists a stationary measure for the chain with density belonging to $C_{a,p}$ and that the same estimates we have in Section 3.3 for the measure of the set $\mathcal{B}_{\varepsilon}$ hold. This allows to conclude that the hitting times to the set $\mathcal{B}_{\varepsilon}$ satisfy the same type of bound in the proof of Theorem A. Notice the independence of the above on the choice of density φ for the noise. This implies that all the arguments continue to hold independently on the size of the noise which, however, contribute to spoil the low-dimensional approximation for the hubs in that it randomly perturbs it.

3.5 An Excursion into One-Dimensional Dynamics: Proof of Theorem B

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

Proposition 3.11. *Let $g: \mathbb{T} \rightarrow \mathbb{T}$ be C^2 and hyperbolic (in the sense of Definition 3.1), and assume that g has an attracting set Λ (consisting of periodic orbits). Then there exist $\chi \in (0, 1)$, $C > 0$ so that for each $\varepsilon > 0$ and $T = 1/\varepsilon$ the following holds.*

There exists a set $\Omega \subset \mathbb{T}$ of measure $1 - \varepsilon^{1-\chi}$ so 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 $k \geq T_0$,

- $g_{\underline{\omega}}^k$ maps each component J of Ω into components of the immediate basin of

³Take $\rho \in C_{a,p}$. Then

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

the periodic attractor of g ;

- the distance of $g_{\omega}^k(J)$ to a periodic attractor of g is at most ε .

The proof of this proposition follows from the next two lemmas:

Lemma 3.7. *Let $g: \mathbb{T} \rightarrow \mathbb{T}$ be C^2 and hyperbolic (in the sense of Definition 3.1), 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 $B_\varepsilon(\Upsilon)$ 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, see [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, see [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 $\lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} = \chi'$. It follows that $N(\varepsilon) \leq \frac{1}{\varepsilon^{\chi}}$ for $\varepsilon > 0$ small. It follows that the Lebesgue measure of $B_\varepsilon(\Upsilon)$ is at most $N(\varepsilon)\varepsilon \leq \varepsilon^{1-\chi}$ for $\varepsilon > 0$ small. $\square$

For simplicity assume that $n = 1$ in Definition 3.1. As in Subsection 3.4.9 the general proof can be reduced to this case.

Lemma 3.8. *Let g and g_{ω}^k as above. Then there exists $C > 0$ so that for each $\varepsilon > 0$ sufficiently small, and taking $\tilde{B} = B_\varepsilon(\Upsilon)$ and $|\omega_i| < \varepsilon' = C\varepsilon$ we have the following:*

1. $g_{\omega}^k(\mathbb{T} \setminus \tilde{B}) \subset \mathbb{T} \setminus \tilde{B}$ for all $k \geq 1$.
2. $\mathbb{T} \setminus \tilde{B}$ consists of at most $1/\varepsilon$ intervals.
3. Take $T_0 = 2/\varepsilon$. Then for each $k \geq T_0$, g_{ω}^k maps each component J of $\mathbb{T} \setminus \tilde{B}$ into a component of the immediate basin of a periodic attractor of g . Moreover, $g_{\omega}^k(J)$ 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 $\mathbb{T} \setminus B_{\varepsilon/4}(\Upsilon)$. If J_i has length less than ε then J_i is contained in $B_\varepsilon(\Upsilon)$, and so the remaining intervals J_i all have length greater or equal to ε and cover $\mathbb{T} \setminus B_\varepsilon(\Upsilon)$. 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 a component of $\mathbb{T} \setminus \Upsilon$ then there exists

k so that $g^k(J')$ is contained in the immediate basin of a periodic attractor of g and $J', \dots, g^k(J')$ are all contained in different components of $\mathbb{T} \setminus \mathcal{Y}$. This, together with 1) and 2) implies that each component of $\mathbb{T} \setminus \widetilde{B}$ 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 that $\varepsilon > 0$ is sufficiently small so that also $2/\varepsilon > m^4$). $\square$

Proof of Theorem B. a) Fix an integer $\sigma \geq 2$, $\alpha \in \mathbb{R}$, $\kappa \in (0, 1]$. The map $\mathcal{F}: C^k(\mathbb{T} \times \mathbb{T}, \mathbb{R}) \rightarrow C^k(\mathbb{T}, \mathbb{R})$ defined by $\mathcal{F}(h)(x) = \int h(x, y) dy$ is continuous. Since the set of hyperbolic C^k maps $g: \mathbb{T} \rightarrow \mathbb{T}$ is open and dense in the C^k topology, see [KSvS07], it follows that the set of C^k 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, then

$$\exists x \in \mathbb{T} \quad \text{s.t.} \quad \frac{d\mathcal{F}(h)(x)}{dx} < 0 \quad (3.55)$$

Condition (3.55) holds for an open and dense set $\Gamma'' \subset C^k(\mathbb{T} \times \mathbb{T}, \mathbb{R})$. Pick $h \in \Gamma''$, then from (3.55) 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 $\mathcal{I} \times V \rightarrow C^k(\mathbb{T}, \mathbb{T})$ 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 rest of the proof follows from Theorem A and Proposition 3.11. $\square$

The following two propositions contain rigorous statements regarding the example presented in the introduction of the paper (Informal Statement of Theorem B).

Proposition 3.12. *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 $O_1 = -O_2$.*

⁴ m is coming from Definition 3.1 for the map g .

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 one has $O = -O$.) Notice that 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. $\square$

Proposition 3.13. *Given $\kappa_0, \kappa_1, \dots, \kappa_m$, consider the families $T_{\beta,j}(x) = 2x - \beta\kappa_j \sin(2\pi x) \bmod 1$. Then*

1. *there exists an open and dense subset $\mathcal{I}'$ of $\mathbb{R}$ so 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 $(-\infty, -\beta_0) \cup (\beta_0, \infty)$ so 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 $\beta \in \mathbb{R}$ so that $T_\beta(x) = 2x - \beta \sin(2\pi x) \bmod 1$ is hyperbolic. By a result in [RGvS15], the set $\mathcal{H}$ is open and dense. It follows that $(1/\kappa_j)\mathcal{H}$ is also open and dense. Hence $(1/\kappa_1)\mathcal{H} \cap \dots \cap (1/\kappa_m)\mathcal{H}$ is open and dense. It follows in particular that this intersection is open and dense in $\mathbb{R}$.

For each $|\beta| > 2\pi$ the map $T_\beta(x) = 2x - \beta \sin(2\pi x) \bmod 1$ has a critical point, and so if such a map T_β is hyperbolic then, by definition, T_β 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. $\square$

3.6 Homogeneous vs Heterogeneous Networks: Proof of Theorem C

Proof of the Theorem C. First we recall that the manifold $\mathcal{S}$ is invariant $\mathbf{F}(\mathcal{S}) \subset \mathcal{S}$. Indeed, if the system is in $\mathcal{S}$ at a time t_0 , hence $x_1(t_0) = \dots = x_N(t_0)$, 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 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 [PERV14].

3.6. Homogeneous vs Heterogeneous Networks: Proof of Theorem C 79

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 Taylor series, we obtain

$$\begin{aligned} \psi_i(t+1) = f'(s(t))\psi_i(t) + \frac{\alpha}{\Delta} \sum_j A_{ij} [\partial_1 h(s(t), s(t))\psi_i(t) + \partial_2 h(s(t), s(t))\psi_j(t) + \\ + R(\psi_i(t), \psi_j(t))] \end{aligned}$$

where ∂_i stands for the derivative in the i th entry and $R(\psi_i, \psi_j)$ is a nonlinear remainder. By Lagrange Theorem we have $R(\psi_i, \psi_j) < C(|\psi_i|^2 + |\psi_j|^2)$, for some positive constant $C = C(\mathbf{A}, h, f)$. Moreover, because h is diffusive

$$\partial_1 h(s(t), s(t)) = -\partial_2 h(s(t), s(t)).$$

Defining $\omega(s(t)) := \partial_1 h(s(t), s(t))$ and entries of the Laplacian matrix $L_{ij} = A_{ij} - d_i \delta_{ij}$, we can write the first variational equation in compact form by introducing $\Psi = (\psi_1, \dots, \psi_n) \in \mathbb{R}^n$. Indeed,

$$\Psi(t+1) = \left[f'(s(t)) \text{Id}_N - \frac{\alpha}{\Delta} \omega(s(t)) L \right] \Psi(t). \quad (3.56)$$

Because the laplacian is symmetric, it admits a spectral decomposition $\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^*$, where $\mathbf{U}$ is the matrix of eigenvectors and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$ the matrix of eigenvalues. Without loss of generality, the eigenvalues can be organised in increasing order

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

as the operator is positive semi-definite. The eigenvalue $\lambda_1 = 0$ always belongs to the spectrum, since every row of $\mathbf{L}$ sums to zero so $L\mathbf{1} = 0$ with $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$. Notice that $\mathbf{1}$ is tangent to the synchronisation manifold $\mathcal{S}$ and all the remaining eigenvectors correspond to transversal directions to $\mathcal{S}$. The Laplacian L has a spectral gap $\lambda_2 > 0$ because the network is connected, as is shown in Theorem 3.3. So, we introduce new coordinates $\Theta = \mathbf{U} \Psi$ to diagonalize $\mathbf{L}$. Notice that by construction Ψ is not in the subspace generated by $\{\mathbf{1}\}$, and thereby Ψ is associated to the dynamics in the transversal eigenmodes. Writing $\Theta = (\theta_1, \dots, \theta_N)$, we obtain the dynamics for the i -th component

$$\theta_i(t+1) = \left[f'(s(t)) + \frac{\alpha \lambda_i}{\Delta} \omega(s(t)) \right] \theta_i.$$

Thus, we decoupled all transversal modes. Since we are interested in the transverse directions we only care about $\lambda_i > 0$. This is equivalent to the linear evolution of Eq. (3.56) restricted to the subspace orthogonal $\mathbf{1}$.

Step 2. Parametric Equation for Transversal Modes. If the transversal modes θ_i with $i = 2, \dots, N$ 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 a parametric equation

$$\mathbf{z}(t+1) = [f'(s(t)) - \beta\omega(s(t))]\mathbf{z}(t). \quad (3.57)$$

This equation will have a uniformly exponentially attracting trivial solution if

$$\nu := \sup_{t>0} \|f'(s(t)) - \beta\omega(s(t))\| < 1. \quad (3.58)$$

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 $f'(s(t)) = 2$ and

$$\omega(s(t)) = -\frac{d\varphi}{dx}(0) + \partial_1 [h'(x, y) - \varphi(y - x)](s(t), s(t)),$$

the condition in Eq. (3.58) is always satisfied as long as

$$\left| 2 - \beta \frac{d\varphi}{dx}(0) \right| + |\beta|\varepsilon < 1 \quad (3.59)$$

Suppose that $\frac{d\varphi}{dx}(0) > 0$ (the negative case can be dealt with analogously). Choosing β inside the closed interval

$$\mathcal{I} = \left[\frac{1}{2} \left(\frac{d\varphi}{dx}(0) \right)^{-1}, \frac{3}{2} \left(\frac{d\varphi}{dx}(0) \right)^{-1} \right]$$

one has $\left| 2 - \beta \frac{d\varphi}{dx}(0) \right| \leq 1/2$ and then picking $\varepsilon > 0$ so that $|\beta|\varepsilon < 1/4$, Eq. (3.59) 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

$$\frac{1}{2} \left(\frac{d\varphi}{dx}(0) \right)^{-1} < \frac{\alpha}{\Delta} \lambda_2 \leq \dots \leq \frac{\alpha}{\Delta} \lambda_N < \frac{3}{2} \left(\frac{d\varphi}{dx}(0) \right)^{-1}. \quad (3.60)$$

Hence, it is necessary for the eigenvalues to satisfy

$$\frac{\lambda_N}{\lambda_2} < 3, \quad (3.61)$$

to find an interval $I \subset \mathbb{R}$ for the coupling strength, such that Eq. (3.60) is satisfied

3.6. Homogeneous vs Heterogeneous Networks: Proof of Theorem C 81

for every $\alpha \in I$. Furthermore, since $\Delta = d_{\max}$, by point 3. of Theorem 3.3

$$1 \leq \frac{\lambda_N}{\Delta} \leq 2$$

one can choose $I \subset \mathbb{R}$ independently on N .

Step 3. Bounds for Laplacian Eigenvalues. Theorem 3.4 below shows that almost every graph $G \in \mathcal{G}_p$

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

Hence, condition (3.61) is met and the manifold $\mathcal{S}$ is normally attracting. We illustrate such a network in Figure 3.4. Indeed, since the coordinates θ_i of the linear approximation decay to zero as $\theta_i(t) \leq Ce^{-\eta t}$ for all $i = 2, \dots, N$ with $\eta > 0$, then the full nonlinear equations synchronise. Indeed, $\|\Psi(t)\| \leq \tilde{C}e^{-\eta t}$, which means that the first variational equation Eq. 3.56 is uniformly stable. To tackle the nonlinearities in the remainder, we notice that for any $\varepsilon > 0$ there is $\delta_0 > 0$ and $C_\varepsilon > 0$ such that for all $|x_i(t_0) - x_j(t_0)| \leq \delta_0$, the nonlinearity is small and by a Groenwall type estimate we have

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

this will precisely happen when the condition Eq. (3.61) is satisfied. The open set for 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 3.3 we obtain

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

hence as the heterogeneity of the degrees increases, the above ratio becomes larger and condition Eq. (3.61) is never met regardless of the value of α . This implies that there are always unstable modes, and the synchronisation manifold $\mathcal{S}$ is unstable. $\square$

The spectrum of the Laplacian $\mathbf{L}$ of a graph 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 3.3. *Let G be a simple network of size N and $\mathbf{L}$ its associated Laplacian with eigenvalues $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$. Then:*

1. [Moh91] $\lambda_2 \geq \frac{4}{ND}$

2. [Fie73] $\lambda_2 \leq \frac{N}{N-1} d_1$
3. [Fie73] $\frac{N}{N-1} d_{\max} \leq \lambda_N \leq 2d_{\max}$

Theorem 3.4 ([Moh92]). Consider the ensemble of random graphs $\mathcal{G}_p$ ⁵ with $p > \frac{\log N}{N}$, then asymptotically almost surely⁶

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

where

$$f(N) = \sqrt{(3 + \varepsilon)(1 - p)pN \log N}$$

for $\varepsilon > 0$ arbitrary.

Regular Networks. Consider a network with N nodes, each one coupled to its $2K$ nearest neighbours (see an Illustration in Figure 3.4 when $K = 2$). In such regular network every node has the same degree $2K$.

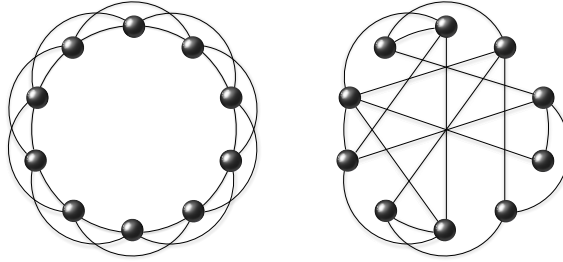


Figure 3.4: On the left panel we present a regular network where every nodes connects to its 2 left and 2 right nearest neighbours. Such networks show poor synchronisation properties in the large N limit if $K \ll N$ as shown in Eq. (3.62). On the right panel, we depict a random (Erdős-Rényi) network 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 synchronisation properties.

Whenever, $K \ll N$ the network will not display synchronisation. 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 $N \rightarrow \infty$. However, as $K \rightarrow N/2$ the network is good for synchronisation. In this case the diameter is extremely small as the graph is close to a full graph.

⁵See Erdős-Rényi in Section 2.1.2

⁶See Definition 2.3

3.6. Homogeneous vs Heterogeneous Networks: Proof of Theorem C 83

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

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

Hence, we can obtain the asymptotics in $K \ll N$ for the synchronisation Eq. (3.61). Using a Taylor expansion in this expression, we obtain

$$\frac{\lambda_N}{\lambda_2} \approx \frac{(3\pi + 2)N^2}{2\pi^3 K^2} \quad (3.62)$$

and when $K \ll N$ synchronisation is never attained.

From a graph theoretic perspective, when $K \ll N$, e.g. K is fixed and $N \rightarrow \infty$ then $\lambda_2 \sim 1/N^2$, implying that the bound in Theorem 3.3 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 nodes 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 $p > \log N/N$).

Coupled Maps with Nonzero Distortion

Contents

4.1 Sparsely Coupled High-Dimensional Systems with Nonzero Distortion	87
4.1.1 Condition for Uniform Expansion of $\mathbf{F}$	88
4.1.2 Estimates on the Entries of $\mathbf{D}_x \mathbf{F}^{-1}$	89
4.1.3 Invariance of Cones of Functions Regular along each Coordinate	91
4.1.4 Decoupling of Nodes from the Network	94
4.1.5 One-Dimensional Marginals along Coordinate Directions	97
4.1.6 One-Dimensional Disintegration along Coordinate Directions	99
4.1.7 Spatial Large Deviations	103
4.2 HCM with Local Dynamics having Nonzero Distortion	103
4.2.1 Global Expansion of $\mathbf{F}_\varepsilon$	104
4.2.2 Distortion of $\mathbf{F}_\varepsilon$	104
4.2.3 Invariant Cones of Functions Regular along Each Coordinate	109
4.2.4 One-Dimensional Disintegration Along the Coordinates of the Low Degree Nodes	112

This chapter is dedicated to a strategy for treating coupled maps in the case where the uncoupled dynamics are C^2 uniformly expanding circle maps with nonzero distortion. The global bounds we obtained for the invariant density of $\mathbf{F}_\varepsilon$ (for example in Eq. (3.28)) do not give any useful information (not even for completely uncoupled maps) due to the exponential increase of the nonlinearities with the dimension. As we have seen in Section 2.4, in the uncoupled case, the set of product measures is invariant under application of the transfer operator, so one can easily obtain statistical stability and spatial large deviations from the one-dimensional case. Whenever the maps are coupled, the product structure is lost and one needs to clarify the role that the coupling has in making density deviate from a pure

product. In this chapter we suggest an approach to deal with this problem and show two examples in which it can be applied to the case that the network map $\mathbf{F} : \mathbb{T}^N \rightarrow \mathbb{T}^N$, or its truncated version, is uniformly expanding. The strategy is the following:

- 1. Distortion of the Jacobian.** Determine how each one of the coordinates, separately, contribute to the distortion of the reference volume.
- 2. Entries of $\mathbf{D}_z \mathbf{F}^{-1}$** Even if most of the entries of $\mathbf{D}_z \mathbf{F}$ (or $\mathbf{D}_z \mathbf{F}_\varepsilon$) equal zero, most entries of the inverse matrix do not. However, given the strength of the coupling, one can estimate the order of magnitude of $(\mathbf{D}_z \mathbf{F}^{-1})_{ij}$ (or $(\mathbf{D}_z \mathbf{F}_\varepsilon^{-1})_{ij}$). This is expected to be different depending on the distance¹ of nodes i and j in the network. Longer distances correspond to smaller orders of magnitude.
- 3. Invariant Cones of Regular Positive Functions** Determine cones of invariant densities whose regularity along each coordinate can be deduced from the entries of $\mathbf{D}_z \mathbf{F}^{-1}$ and the distortion of the Jacobian estimates in the previous steps.
- 4. Lasota-Yorke for One-Dimensional Disintegrations** Define a (very) strong and a weak norm as the suprema of the one-dimensional $W^{1,1}$ and L^1 norms of the disintegration along all leaves of all foliations corresponding to the coordinate directions. Taken a measure with density in the cone found at the previous step, one proves a Lasota-Yorke inequality with respect to these norms.
- 5. Disintegration of the Invariant Measure** Use the spectral gap of the transfer operator for the one-dimensional uncoupled local map and the L-Y inequality obtained at the previous step to show that the disintegration of the invariant measure along one-dimensional leaves of the cartesian foliations is close to the invariant measure for the uncoupled local map.
- 6. Large Deviations** Use the above results and estimates similar to those in Section 2.4.4 to obtain large deviations of empirical averages over the nodes of the network.

In Section 4.1 we consider coupled maps with nonzero distortion and “sparse” interactions. The only parameters that describe the network are the number of nodes N and their maximum (in and out) degree Δ . In Section 4.2 we use the

¹The minimum number of edges one has to walk through to go from one node to another.

same technique to draw similar conclusions on HCM in the case where the reduced dynamics of all the hub nodes is uniformly expanding.

We should stress once again that we consider only two possible scenarios by fixing some structural parameters of the network, but our purpose is to give a flexible methodology that can be applied in different cases to give results which are the more precise, the more information on the structure of the network is known.

4.1 Sparsely Coupled High-Dimensional Systems with Nonzero Distortion

Consider a map $\mathbf{F} : \mathbb{T}^N \rightarrow \mathbb{T}^N$:

$$F^{(i)}(\mathbf{x}) := f(x_i) + \varepsilon \sum_{j=1}^N A_{ij} h(x_i, x_j) \mod 1 \quad (4.1)$$

where $f \in C^2(\mathbb{T}, \mathbb{T})$ is a uniformly expanding map, $h \in C^2(\mathbb{T} \times \mathbb{T}, \mathbb{R})$, and $\varepsilon \in \mathbb{R}^+$. We give a description of the dynamics depending only on the structural parameter

$$\Delta = \max\{d_{O,i}, d_{I,i} : i \in [N]\}. \quad (4.2)$$

This says very little on the network, so the result will be very general. Given a measure μ on $\mathbb{T}^N$ and any $\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}$, we indicate with $\mu_{\hat{\mathbf{x}}_i}$ the conditional expectation along the i -th coordinate where the other are given by the entries of $\hat{\mathbf{x}}_i$. We obtain the following theorem proved in Section 4.1.6.

Theorem 4.1 (Statistical Stability of 1D Disintegrations). *Suppose that ν and μ_0 are the a.c.i.p. measures of $\mathbf{F}$ and f respectively. Then there exists $\eta > 0$, depending on f and h only, such that if $\varepsilon \Delta^2 < \eta$, then*

$$\|\nu_{\hat{\mathbf{x}}_i} - \mu_0\|_{C^0} \leq \mathcal{O}(\varepsilon \Delta^2 \log(\varepsilon \Delta^2))$$

for all $i \in [N]$ and all $\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}$.

Remark 4.1. Notice that we obtain a condition on the coupling strength that does not depend on N , but on Δ only. The result holds for any $\Delta \in [N-1]$, but the most interesting scenario is for $\Delta \ll N$ corresponding to the sparsely coupled case. The case $\Delta = \mathcal{O}(N)$ can be treated in a different way to obtain better estimates.

The above tells us that for the prescribed scaling of ε with Δ , the disintegration of the invariant measure of $\mathbf{F}$ is close to the invariant density for f , depending on the magnitude of ε . This gives sufficiently strong control on the invariant density

to estimate the measure of the set $\mathcal{B}_{\xi,N}$ (Eq. (2.17)). As we have seen in Section 2.4.4, upper bounds on the measure of this set were readily available in the case the measure was a product. However, the geometry of $\mathcal{B}_{\xi,N}$ can be very complicated so it can be in principle difficult to get an upper bound in case the measure is not a product. Theorem 4.1 helps us in this respect and leads us to the following corollary proved in Section 4.1.7.

Corollary 4.1. *Under the hypotheses of Theorem 4.1,*

$$\nu(\mathcal{B}_{\xi,N}) \leq \exp \left[C(\mathcal{O}(\varepsilon\Delta^2 \log(\varepsilon\Delta^2)) - \xi^2)N \right].$$

The above implies that as long as $\varepsilon\Delta^2$ is sufficiently small, the measure of the set $\mathcal{B}_{\xi,N}$ is still exponentially small in N .

4.1.1 Condition for Uniform Expansion of $\mathbf{F}$

Proposition 4.1. *There is $\eta > 0$ depending on f and h only, such that for every $p \in [1, \infty]$, if*

$$\varepsilon(\Delta + N^{1/p}\Delta^{1/q}) < \eta, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

then the map $\mathbf{F}$ is uniformly expanding with respect to the Finsler metric induced by the norm $\|\cdot\|_p$.

Proof. We now show that there is $\sigma > 1$ such that $|D_x f| \geq \sigma \forall x \in \mathbb{T}$. Pick $\mathbf{x} \in \mathbb{T}^N$, $\mathbf{v} = (v_1, \dots, v_N) \in \mathbb{R}^N$, and call $\mathbf{v}' = \mathbf{D}_{\mathbf{x}} \mathbf{F} \mathbf{v}$.

$$v'_i = \left[D_{x_i} f + \varepsilon \sum_{n=1}^N A_{in} h_1(x_i, z_n) \right] v_i + \varepsilon \sum_{n=1}^N A_{in} h_2(x_i, z_n) v_n.$$

Hence

$$\|\mathbf{v}'\|_p \geq (\sigma - \mathcal{O}(\varepsilon\Delta)) \|\mathbf{v}\|_p - \mathcal{O}(\varepsilon N^{1/p}) \max_{i=1, \dots, L} \left[\sum_{n=1}^N A_{in} |v_n| \right].$$

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

$$\|\mathbf{w}\|_1 \leq k^{1/q} \|\mathbf{w}\|_p, \quad \text{with } \frac{1}{p} + \frac{1}{q} = 1 \quad (4.3)$$

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

$$\sum_{n=1}^N A_{in} |v_n| \leq \Delta^{1/q} \left(\sum_{n=1}^N A_{in} |v_n|^p \right)^{1/p} \leq \Delta^{1/q} \|\mathbf{v}\|_p$$

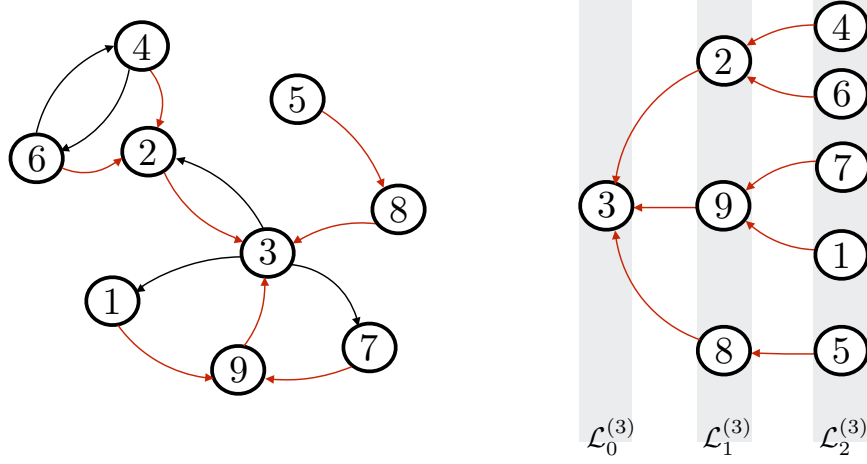


Figure 4.1: A graph and the sets $\mathcal{L}_0^{(3)}$, $\mathcal{L}_1^{(3)}$, and $\mathcal{L}_2^{(3)}$ associated to the node 3.

since at most Δ terms are non-vanishing in the sum $\left(\sum_{n=1}^N A_{in}|v_n|^p\right)$, we can view it as p -norm of a vector in $\mathbb{R}^\Delta$, which implies

$$\|\mathbf{v}'\|_p \geq (\sigma - \mathcal{O}(\varepsilon\Delta)) \|\mathbf{v}\|_p - \mathcal{O}(\varepsilon N^{1/p} \Delta^{1/q}) \|\mathbf{v}\|_p = (\sigma - \mathcal{O}(\varepsilon\Delta) - \mathcal{O}(\varepsilon N^{1/p} \Delta^{1/q})) \|\mathbf{v}\|_p.$$

□

It is a consequence of the proposition above and well known results on uniformly expanding maps (e.g. Theorem 2.1), that $\mathbf{F}$ has an a.c.i.p. measure ν with density $\frac{d\nu}{dm} > 0$ and that $P^n \varphi \rightarrow \frac{d\nu}{dm}$ uniformly for every φ sufficiently regular².

4.1.2 Estimates on the Entries of $\mathbf{D}_\mathbf{x} \mathbf{F}^{-1}$

We now gather information concerning local inverses of $\mathbf{F}$. In particular, we bound the entries of the inverse of the differential $\mathbf{D}_\mathbf{x} \mathbf{F}^{-1}$. It turns out that these are different orders of ε which can be classified with the help of the following definition.

Definition 4.1. Given a directed graph $G = ([N], \mathcal{E})$ and a node $i \in [N]$ one can recursively define disjoint subsets putting $\mathcal{L}_0^{(i)} := \{i\}$ and

$$\mathcal{L}_j^{(i)} := \{\ell \in [N] : \exists k \in \mathcal{L}_{j-1}^{(i)} \text{ with } A_{\ell k} = 1\}.$$

Call $\kappa = \min\{j : \mathcal{L}_j^{(i)} \neq \emptyset\}$, and define $\mathcal{L}_{\kappa+1}^{(i)} := [N] \setminus \bigcup_{j=0}^{\kappa} \mathcal{L}_j^{(i)}$.

We have the following estimates for the entries of $\mathbf{D}_\mathbf{x} \mathbf{F}^{-1}$.

²Lipschitz is enough.

Proposition 4.2. *There is $c_{\#} \in \mathbb{R}^+$, $c > 1$, and $\eta \in \mathbb{R}^+$ depending on f and h only, such that if $\varepsilon\Delta^2 < \eta$, for every $i \in [N]$, taken $\{\mathcal{L}_j^{(i)}\}_{j=1}^{\kappa+1}$ as in Definition 4.1 above,*

$$|(\mathbf{D}_x \mathbf{F}^{-1})_{ik}| \leq \frac{1}{[c_{\#} - \varepsilon^2 \Delta^{3/2}]} \frac{\varepsilon^j \Delta^{j-1/2}}{c^{j-1}}, \quad \text{for } k \in \mathcal{L}_j^{(i)}, \quad j \geq 1 \quad (4.4)$$

and

$$|(\mathbf{D}_x \mathbf{F}^{-1})_{ii}| \leq \frac{1}{c - c_{\#} \varepsilon^2 \Delta^{3/2}}. \quad (4.5)$$

Proof. The matrix $\mathbf{D}_x \mathbf{F}$ satisfies the hypotheses of Lemma B.2 in the Appendix. Pick $i \in [N]$ and consider the sets $\{\mathcal{L}_j^{(i)}\}_{j=1}^{\kappa+1}$ as in Definition 4.1. It follows from Lemma B.2, that there is $K \in \mathbb{R} \setminus \{0\}$ (where $K \neq 0$ since 4.1 $|\mathbf{D}_z \mathbf{F}| \neq 0$) such that

$$|\widehat{\mathbf{DF}}_{ii}| \geq K \left[1/2 - \varepsilon^2 \Delta^{3/2} \right]$$

and for all $k \in \mathcal{L}_j^{(i)}$

$$|\widehat{\mathbf{DF}}_{ki}| \leq K \frac{\varepsilon^j}{c^{j-1}} \Delta^{j-1/2}.$$

Furthermore

$$\begin{aligned} |\mathbf{DF}| &\geq |(\mathbf{DF})_{ii}| |\widehat{\mathbf{DF}}_{ii}| - \left| \sum_{\ell \neq i} (-1)^\ell (\mathbf{DF})_{\ell i} |\widehat{\mathbf{DF}}_{\ell i}| \right| \\ &\geq c |\widehat{\mathbf{DF}}_{ii}| - \sum_{\ell \neq i} A_{\ell i} \varepsilon |\widehat{\mathbf{DF}}_{\ell i}| \\ &\geq cK \left[1/2 - \varepsilon^2 \Delta^{3/2} \right] - \Delta \varepsilon K \varepsilon \Delta^{-1/2} \\ &\geq K \left[c_{\#} - \varepsilon^2 \Delta^{3/2} \right]. \end{aligned} \quad (4.6)$$

This allows to conclude that if $k \in \mathcal{L}_j^{(i)}$ with $j \neq 0$

$$|(\mathbf{DF}^{-1})_{ik}| = \frac{|\widehat{\mathbf{DF}}_{ki}|}{|\mathbf{DF}|} \leq \frac{\varepsilon^j \Delta^{j-1/2}}{c^{j-1} [c_{\#} - \varepsilon^2 \Delta^{3/2}]}.$$

and

$$\begin{aligned} |(\mathbf{DF}^{-1})_{ii}| &= \frac{|\widehat{\mathbf{DF}}_{ii}|}{|\mathbf{DF}|} = \frac{|\widehat{\mathbf{DF}}_{ii}|}{c |\widehat{\mathbf{DF}}_{ii}| - \sum_{\ell \neq i} A_{\ell i} \varepsilon |\widehat{\mathbf{DF}}_{\ell i}|} \\ &\leq \frac{1}{c - \sum_{\ell \neq i} A_{\ell i} \varepsilon \frac{|\widehat{\mathbf{DF}}_{\ell i}|}{|\widehat{\mathbf{DF}}_{ii}|}} \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{c - c_{\#} \sum_{\ell \neq i} A_{\ell i} \varepsilon \varepsilon \sqrt{\Delta}} \\ &\leq \frac{1}{c - c_{\#} \varepsilon^2 \Delta^{3/2}}. \end{aligned}$$

□

4.1.3 Invariance of Cones of Functions Regular along each Coordinate

The next theorem shows that the invariant measures of sparsely coupled maps have disintegrations along every coordinate of regularity comparable to the regularity of the invariant measure of f .

Theorem 4.2. *There are $a \in \mathbb{R}^+$ and $\eta > 0$, depending on f and h only, such that if $\varepsilon \Delta^2 < \eta$, and if φ is the density of the a.c.i.p. measure for the map $\mathbf{F}$, then for all $i \in [N]$,*

$$\frac{\varphi(x_i; \hat{\mathbf{x}}_i)}{\varphi(\bar{x}_i; \hat{\mathbf{x}}_i)} < \exp[ad(\bar{x}_i, x_i)], \quad \forall \hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}.$$

Remark 4.2. Notice that by Proposition D.1

$$\varphi_{\hat{\mathbf{x}}_i}(x) = \frac{\varphi(x; \hat{\mathbf{x}}_i)}{\int_{\mathbb{T}} \varphi(y; \hat{\mathbf{x}}_i) dm(y)}$$

is the density of $\mu_{\hat{\mathbf{x}}_i}$ which is the disintegration at $\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}$ of μ w.r.t the map $\pi_{\{i\}^c} : \mathbb{T}^N \rightarrow \mathbb{T}^{N-1}$ and the measure $\pi_{\{i\}^c} \mu$.

Definition 4.2. Define the convex cone of functions

$$\mathcal{C}_{i,a} := \left\{ \varphi : \mathbb{T}^N \rightarrow \mathbb{R}^+ : \frac{\varphi(x_i; \hat{\mathbf{x}}_i)}{\varphi(\bar{x}_i; \hat{\mathbf{x}}_i)} \leq \exp[ad(x_i, \bar{x}_i)], \forall \hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}, \forall x_i, \bar{x}_i \in \mathbb{T} \right\}$$

and

$$\mathcal{C}_a := \bigcap_{i=1}^N \mathcal{C}_{i,a}$$

The following proposition shows that applying the Perron-Frobenius operator to a function in the cone $\mathcal{C}_a$ for some $a \in \mathbb{R}^+$, one gets a function in $\mathcal{C}_{i,\bar{a}}$, for all $i \in [N]$, where $\bar{a}$ depends on a and can be estimated from the bounds on the entries of $\mathbf{D}_x \mathbf{F}^{-1}$.

Proposition 4.3. *Suppose that P is the Perron-Frobenius operator of $\mathbf{F}$. There exists $\eta > 0$ depending on f and h only such that for all $a > 0$, $\varphi \in \mathcal{C}_a$, and $i \in [N]$,*

$P\varphi \in \mathcal{C}_{i,(\lambda a + C_\#)}$ where $C_\#$ depends on f and h only and

$$\lambda := \left[|(\mathbf{D}\mathbf{F})_{ii}|^{-1} + \mathcal{O}(\varepsilon \Delta^{3/2}) \right].$$

Proof. Take $\varphi \in \mathcal{C}_a$. Call $\mathbf{y} = \mathbf{F}_{(m)}^{-1}(x_1, \dots, x_i, \dots, x_N)$ and $\bar{\mathbf{y}} = \mathbf{F}_{(m)}^{-1}(x_1, \dots, \bar{x}_i, \dots, x_N)$ for some m where $\{\mathbf{F}_{(m)}^{-1}\}_m$ are the inverse branches of the map $\mathbf{F}$. One obtains

$$\frac{|\mathbf{D}_{\mathbf{y}}\mathbf{F}|}{|\mathbf{D}_{\bar{\mathbf{y}}}\mathbf{F}|} \leq \prod_{i=1}^N \frac{|D_{y_i}f| |\mathcal{D}(\mathbf{y})|}{|D_{\bar{y}_i}f| |\mathcal{D}(\bar{\mathbf{y}})|} \quad (4.7)$$

where $\mathcal{D}(\mathbf{y})$ is the matrix $\mathbf{D}_{\mathbf{y}}\mathbf{F}$ to which $|D_{y_i}f|$ has been factored out of the i -th column for all $i \in [N]$. This way we can use Proposition A.1 to upper bound the ratio of the determinants of $\mathcal{D}(\mathbf{y})$ and $\mathcal{D}(\bar{\mathbf{y}})$. First of all, notice that there is $\eta > 0$ depending on f and h only such that if $\varepsilon \Delta < \eta$, then $\mathbf{B}(\mathbf{y}) = \mathcal{D}(\mathbf{y}) - \text{Id}$ has norm less than one with respect to the infinity norm on $\mathbb{R}^N$. In fact pick $\mathbf{v} \in \mathbb{R}^N$ with $\max_k |v_k| = 1$ then

$$\sum_{j=1}^N B_{kj}(\mathbf{y})v_j \leq \mathcal{O}(\varepsilon d_{I,k})|v_k| + \sum_{j \in [N] \setminus \{k\}} A_{kj} \mathcal{O}(\varepsilon)|v_j| \leq \mathcal{O}(\varepsilon d_{I,k}), \quad (4.8)$$

and analogously for $\mathbf{B}(\bar{\mathbf{y}}) = \mathcal{D}(\bar{\mathbf{y}}) - \text{Id}$. Using equations (4.4) and (4.5) one can see that

$$d(y_k, \bar{y}_k) \leq \begin{cases} \mathcal{O}(1)d(x_i, \bar{x}_i) & k = i \\ \mathcal{O}(\varepsilon^j \Delta^{j-1/2})d(x_i, \bar{x}_i) & k \in \mathcal{L}_j^{(i)}, j \geq 1 \end{cases} \quad (4.9)$$

For $k \neq i$ the column $\text{Col}^k[\mathbf{B}(\mathbf{y}) - \mathbf{B}(\bar{\mathbf{y}})]$ has at most Δ entries different from zero. If $k \in \mathcal{L}_1^{(i)}$ then $\|\text{Col}^k[\mathbf{B}(\mathbf{y}) - \mathbf{B}(\bar{\mathbf{y}})]\|_\infty \leq \mathcal{O}(\varepsilon)d(x_i, \bar{x}_i)$ and this is the case for at most Δ different values of k . If $k \in \mathcal{L}_j^{(i)}$ then

$$\begin{aligned} \|\text{Col}^k[\mathbf{B}(\mathbf{y}) - \mathbf{B}(\bar{\mathbf{y}})]\|_\infty &\leq \mathcal{O}(\varepsilon) \max_{\ell \in \mathcal{L}_j^{(i)}} \{\mathcal{O}(\Delta)d(y_k, \bar{y}_k), d(y_\ell, \bar{y}_\ell)\} \\ &\leq \mathcal{O}(\varepsilon^{j+1} \Delta^{j+1/2})d(x_i, \bar{x}_i). \end{aligned}$$

and the cardinality of $\mathcal{L}_j^{(i)}$ is at most Δ^j .

$$\|\text{Col}^i[\mathbf{B}(\mathbf{y}) - \mathbf{B}(\bar{\mathbf{y}})]\|_\infty \leq \mathcal{O}(\varepsilon d_{I,i})d(x_i, \bar{x}_i) \leq \mathcal{O}(\varepsilon \Delta)d(x_i, \bar{x}_i).$$

Plugging these estimates in Proposition A.1 one obtains

$$\begin{aligned}
 \frac{|\mathcal{D}(\mathbf{y})|}{|\mathcal{D}(\bar{\mathbf{y}})|} &\leq \exp \left\{ \left[\mathcal{O}(\varepsilon\Delta) + \sum_{j=1}^m \Delta^j \mathcal{O}(\varepsilon^{j+1}\Delta^j) \right] d(x_i, \bar{x}_i) \right\} \\
 &\leq \exp \left\{ \left[\mathcal{O}(\varepsilon\Delta) + \mathcal{O}(\Delta^{5/2}\varepsilon^2) \sum_{j=0}^{m-1} (\varepsilon\Delta^2)^j \right] d(x_i, \bar{x}_i) \right\} \\
 &\leq \exp \left\{ \left[\mathcal{O}(1) + \mathcal{O}(\Delta^{5/2}\varepsilon^2) \frac{(\varepsilon\Delta^2)^m - 1}{\varepsilon\Delta^2 - 1} \right] d(x_i, \bar{x}_i) \right\} \\
 &\leq \exp\{\mathcal{O}(1)d(x_i, \bar{x}_i)\}.
 \end{aligned}$$

where in the last step we considered $\varepsilon\Delta^2 < \eta < 1$. The other part of the product can be estimated from

$$\frac{|D_{y_j}f|}{|D_{\bar{y}_j}f|} = \left| 1 - \frac{D_{y_j}f - D_{\bar{y}_j}f}{D_{\bar{y}_j}f} \right| \leq (1 + \mathcal{O}(1)d(y_j, \bar{y}_j)),$$

which with analogous computations using again (4.9) implies

$$\begin{aligned}
 \prod_{j=1}^N \frac{|D_{y_j}f|}{|D_{\bar{y}_j}f|} &\leq \exp \left\{ \left[\mathcal{O}(1) + \sum_{j=1}^m \Delta^j \varepsilon^j \Delta^{j-1/2} \right] d(x_i, \bar{x}_i) \right\} \\
 &\leq \exp \left\{ \left[\mathcal{O}(1) + \mathcal{O}(\varepsilon\Delta^{3/2}) \sum_{j=0}^{m-1} (\varepsilon\Delta^2)^j \right] d(x_i, \bar{x}_i) \right\} \\
 &\leq \exp\{\mathcal{O}(1)d(x_i, \bar{x}_i)\}.
 \end{aligned}$$

Taking $\varphi \in \mathcal{C}_a$ one obtains

$$\begin{aligned}
 \frac{\varphi(y)}{\varphi(\bar{y})} &\leq \exp \left\{ \sum_{j=1}^N ad(y_i, \bar{y}_i) \right\} \\
 &\leq \exp \left\{ \left[(\mathbf{D}\mathbf{F}^{-1})_{ii} + \mathcal{O}(\varepsilon\Delta^{3/2}) \frac{(\varepsilon\Delta^2)^{m+1} - 1}{\varepsilon\Delta^2 - 1} \right] ad(x_i, \bar{x}_i) \right\}. \quad (4.10)
 \end{aligned}$$

Putting together all the estimates one obtains

$$\frac{(P\varphi)(x_1, \dots, x_i, \dots, x_N)}{(P\varphi)(x_1, \dots, \bar{x}_i, \dots, x_N)} \leq \exp \left\{ \left[(\mathbf{D}\mathbf{F}^{-1})_{ii} + \mathcal{O}(\varepsilon\Delta^{3/2}) \right] ad(x_i, \bar{x}_i) + \mathcal{O}(1)d(x_i, \bar{x}_i) \right\}.$$

□

The proof of Theorem 4.2 now directly follows from the above proposition.

Proof of Theorem 4.2. We know from Proposition 4.2 that there is $\eta > 0$ such that if $\varepsilon\Delta^2 < \eta$ then $|(\mathbf{D}\mathbf{F}^{-1})_{ii}| < 1$. Then choose η sufficiently small so that if $\varepsilon\Delta^2 < \eta$ implies $\lambda = [(\mathbf{D}\mathbf{F}^{-1})_{ii} + \mathcal{O}(\varepsilon\Delta^{3/2})] \leq \bar{\lambda} < 1$. By Proposition 4.7, for all $a > \frac{C_{\#}}{1-\bar{\lambda}} = \mathcal{O}(1)$, all $\varphi \in \mathcal{C}_a$, and all $i \in [N]$, $P\varphi \in \mathcal{C}_{i,a}$, which implies that for this choice of a , $PC_a \subset \mathcal{C}_a$. For every $\varphi \in \mathcal{C}_a$ with $\int \varphi = 1$, the sequence $P^n\varphi$ in $\mathcal{C}_a$ converges uniformly to the invariant density of $\mathbf{F}$, and this implies that also the invariant density belongs to $\mathcal{C}_a$. $\square$

4.1.4 Decoupling of Nodes from the Network

With a procedure already suggested in [KL06] we define for every $i \in [N]$ the map $\mathbf{F}_i$ component-wise as

$$F_i^{(j)}(x_1, \dots, x_N) = \begin{cases} F^{(j)} - A_{ji}h(x_j, x_i) & j \neq i \\ f(x_i) & j = i \end{cases} \quad (4.11)$$

The above corresponds to the map $\mathbf{F}$ where the i -th node has been decoupled from the network³ (both incoming and outgoing interactions are eliminated) and evolves according to the map f . Define the map $\hat{\mathbf{F}}_i : \mathbb{T}^{N-1} \rightarrow \mathbb{T}^{N-1}$ which gives the action of $\mathbf{F}_i$ on the coordinates different from the i -th one. Call P_f the transfer operator for f , P_i the transfer operator for $\mathbf{F}_i$, and $\hat{P}_i$ the transfer operator for $\hat{\mathbf{F}}_i$. Since the i -th node has been completely decoupled from the network one has

$$P_i = P_f \hat{P}_i = \hat{P}_i P_f. \quad (4.12)$$

In the following lemma we compare the action of $\mathbf{F}$ and $\mathbf{F}_i$ in terms of their Jacobians and action on densities from the cone $\mathcal{C}_a$. The estimates we obtain will be used in Section 4.1.6 to study the disintegration of the invariant measure of $\mathbf{F}$ along coordinate directions.

Lemma 4.1. *There is $\eta > 0$ depending on f and h only such that if $\varepsilon\Delta^2 < \eta$ then*

$$1. \quad \frac{|\mathbf{D}_x \mathbf{F}|}{|\mathbf{D}_x \mathbf{F}_i|} \leq \exp\{\mathcal{O}(\varepsilon\Delta)\}, \quad \forall \mathbf{x} \in \mathbb{T}^N; \quad (4.13)$$

³ $\mathbf{F}_i$ is the network dynamical system with the same local dynamics (same f and h), but on the graph $\mathcal{G}'_i$ whose adjacency matrix is $(A'_{k\ell})$

$$A'_{k\ell} = \begin{cases} A_{k\ell} & k \neq i, \ell \neq i \\ 0 & \text{otherwise} \end{cases}$$

2. if $\varphi \in \mathcal{C}_a$,

$$\frac{P\varphi}{P_i\varphi} \leq \exp\{\mathcal{O}(\varepsilon\Delta^2)\}; \quad (4.14)$$

3. there is $C_\# \in \mathbb{R}^+$ such that if $\varphi \in \mathcal{C}_a$, then

$$\left| \partial_{x_i} \frac{P\varphi}{P_i\varphi} \right| \leq C_\#. \quad (4.15)$$

Proof. To prove 1., we use, as usual Proposition A.1. Fix $\mathbf{x} \in \mathbb{T}^N$. Factor out from the columns of $\mathbf{D}_\mathbf{x}\mathbf{F}$ and $\mathbf{D}_\mathbf{x}\mathbf{F}_i$ the numbers $D_{x_i}f$ to obtain matrices $\mathcal{D}$ and $\mathcal{D}_i$. Observe that

$$\frac{|\mathbf{D}_\mathbf{x}\mathbf{F}|}{|\mathbf{D}_\mathbf{x}\mathbf{F}_i|} = \frac{\prod_{j=1}^N |D_{x_j}f|}{\prod_{j=1}^N |D_{x_j}f|} \frac{|\mathcal{D}|}{|\mathcal{D}_i|}$$

Now define $\mathbf{B} := \mathcal{D} - \text{Id}$ and $\mathbf{B}_i := \mathcal{D}_i - \text{Id}$. One can verify as in Eq. (4.8) that $\|\mathbf{B}\|_\infty, \|\mathbf{B}_i\|_\infty \leq \mathcal{O}(\varepsilon d_{I,i})$ which is strictly less than one for sufficiently small η . Furthermore

$$\|\text{Col}^j[\mathbf{B} - \mathbf{B}_i]\|_\infty = \begin{cases} \mathcal{O}(\varepsilon d_{I,i}) & j = i \\ \mathcal{O}(\varepsilon) & A_{ji} = 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.16)$$

Plugging the estimates in Proposition A.1, we obtain

$$\frac{|\mathcal{D}|}{|\mathcal{D}_i|} \leq \exp \left\{ \frac{\sum_{j=1}^N \|\text{Col}^j[\mathbf{B} - \mathbf{B}_i]\|_\infty}{1 + \mathcal{O}(\varepsilon d_{I,i})} \right\} \leq \exp \{ \mathcal{O}(\varepsilon d_{I,i}) \}.$$

and so point 1. follows.

Now let's prove the second point. First of all recall that

$$P\varphi = \sum_k \frac{\varphi}{|\mathbf{D}\mathbf{F}|} \circ \mathbf{F}_{(k)}^{-1} \quad \text{and} \quad P_i\varphi = \sum_k \frac{\varphi}{|\mathbf{D}\mathbf{F}_i|} \circ \mathbf{F}_{i(k)}^{-1}.$$

Consider functions $\mathbf{F}_{(k)} = \tilde{\mathbf{F}}_{i,0}, \tilde{\mathbf{F}}_{i,1}, \dots, \tilde{\mathbf{F}}_{i,d_{O,i}+1} = \mathbf{F}_{i,(k)}$, where $\tilde{\mathbf{F}}_{i,j}$ and $\tilde{\mathbf{F}}_{i,j+1}$ differ for their action on just one of the coordinates so that $d_{C^1}(\tilde{\mathbf{F}}_{i,j}^{k_j}, \tilde{\mathbf{F}}_{i,j+1}^{k_j}) < \xi$ and $\tilde{\mathbf{F}}_{i,j}^\ell = \tilde{\mathbf{F}}_{i,j+1}^\ell$ for $\ell \neq k_j$.

$$\frac{\varphi \circ \mathbf{F}_{(k)}^{-1}}{\varphi \circ \mathbf{F}_{i,(k)}^{-1}} = \frac{\varphi \circ \tilde{\mathbf{F}}_{i,0}^{-1}}{\varphi \circ \tilde{\mathbf{F}}_{i,1}^{-1}} \cdots \frac{\varphi \circ \tilde{\mathbf{F}}_{i,j}^{-1}}{\varphi \circ \tilde{\mathbf{F}}_{i,j+1}^{-1}}. \quad (4.17)$$

Lemma B.4 along with the fact that $\varphi \in \mathcal{C}_a$, yields

$$\frac{\varphi \circ \tilde{\mathbf{F}}_{i,j}^{-1}}{\varphi \circ \tilde{\mathbf{F}}_{i,j+1}^{-1}} \leq \exp \{a\mathcal{O}(\varepsilon\Delta)\}$$

which along with Eq. (4.17) implies

$$\frac{\varphi \circ \mathbf{F}_{(k)}^{-1}}{\varphi \circ \mathbf{F}_{\mathbf{i}(k)}^{-1}} \leq \exp \{a\mathcal{O}(\varepsilon\Delta^2)\}. \quad (4.18)$$

Now

$$\frac{|\mathbf{DF}| \circ \mathbf{F}_{(k)}^{-1}}{|\mathbf{DF}_{\mathbf{i}}| \circ \mathbf{F}_{\mathbf{i}(k)}^{-1}} = \frac{|\mathbf{DF}| \circ \mathbf{F}_{(k)}^{-1}}{|\mathbf{DF}| \circ \mathbf{F}_{\mathbf{i}(k)}^{-1}} \frac{|\mathbf{DF}| \circ \mathbf{F}_{\mathbf{i}(k)}^{-1}}{|\mathbf{DF}_{\mathbf{i}}| \circ \mathbf{F}_{\mathbf{i}(k)}^{-1}}.$$

The second factor has already been estimated in point 1. For the first factor we can again apply the same trick as in Eq. (4.17) and combining estimates in Lemma B.4 and estimates analogous to Eq. (4.7) one obtains

$$\frac{|\mathbf{DF}| \circ \tilde{\mathbf{F}}_{i,j}}{|\mathbf{DF}| \circ \tilde{\mathbf{F}}_{i,j+1}} \leq \exp\{\mathcal{O}(\varepsilon\Delta)\}$$

which implies

$$\frac{|\mathbf{DF}| \circ \mathbf{F}_{(k)}^{-1}}{|\mathbf{DF}_{\mathbf{i}}| \circ \mathbf{F}_{\mathbf{i}(k)}^{-1}} \leq \exp\{\mathcal{O}(\varepsilon\Delta^2)\}. \quad (4.19)$$

Putting Eq. (4.19) and Eq. (4.18) together one obtains 2.

To prove the last point,

$$\begin{aligned} \frac{(P\varphi)(x_i + t; \hat{\mathbf{x}}_{\mathbf{i}})}{(P_i\varphi)(x_i + t; \hat{\mathbf{x}}_{\mathbf{i}})} - \frac{(P\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})}{(P_i\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})} &= \frac{(P_i\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})}{(P\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})} \left[\frac{(P\varphi)(x_i + t; \hat{\mathbf{x}}_{\mathbf{i}})}{(P\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})} \frac{(P_i\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})}{(P_i\varphi)(x_i + t; \hat{\mathbf{x}}_{\mathbf{i}})} - 1 \right] \\ &\leq \frac{(P_i\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})}{(P\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})} [\exp\{2at\} - 1] \end{aligned}$$

so taking the limit of the increment for $t \rightarrow 0$ one gets

$$\left| \partial_{x_i} \frac{P\varphi}{P_i\varphi} \right| \leq 2a \left| \frac{(P_i\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})}{(P\varphi)(x_i; \hat{\mathbf{x}}_{\mathbf{i}})} \right| \leq 2a \frac{P\varphi}{P_i\varphi} \leq 2a \exp\{[\mathcal{O}(\varepsilon\Delta^2)]\}.$$

□

4.1.5 One-Dimensional Marginals along Coordinate Directions

We now discuss the evolution of one-dimensional marginals under the transfer operator P of $\mathbf{F}$. The results of this section could be easily deduced from similar results in [KL06], but we include them to offer a comparison between the study of marginals and the study of disintegrations we undertake in the next section. As usual $|\cdot|_1$ denote the L^1 norm and $\|\cdot\|$ the Sobolev semi-norm for measures on $\mathbb{T}$

$$\|\mu\| := \sup_{\substack{\psi \in C^\infty(\mathbb{T}, \mathbb{R}) \\ |\psi|_\infty \leq 1}} \mu(D_x \psi).$$

Definition 4.3. Given a density $\varphi : \mathbb{T}^N \rightarrow \mathbb{R}^+$, for every $i \in [N]$ the i -th marginal is given by

$$\varphi_i(x_i) := \int_{\mathbb{T}^{N-1}} \varphi(x_1, \dots, x_N) \prod_{j \in [N] \setminus \{i\}} dm(x_j).$$

The one-dimensional marginals happen to stay close to the invariant measure of f whenever $\varepsilon\Delta$ is sufficiently small. The bound one can obtain is the following.

Proposition 4.4. *Suppose that φ is the invariant density of P and φ_0 is the invariant density of P_f . Then, for all $i \in [N]$*

$$|\varphi_i - \varphi_0|_1 \leq \mathcal{O}(\varepsilon\Delta \log(\varepsilon\Delta)).$$

First of all we obtain a formula for the evolution of the one-dimensional marginals in light of the above decoupling scheme.

Lemma 4.2. *Let $\varphi \in L^1(\mathbb{T}^N, \mathbb{R})$, then*

$$(P\varphi)_i = P_f\varphi_i + \int_{\mathbb{T}^{N-1}} (P - P_i)\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j). \quad (4.20)$$

Proof.

$$\begin{aligned} (P\varphi)_i &= \int_{\mathbb{T}^{N-1}} P\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j) \\ &= \int_{\mathbb{T}^{N-1}} P_i\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j) + \int_{\mathbb{T}^{N-1}} (P - P_i)\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j) \\ &= \int_{\mathbb{T}^{N-1}} \hat{P}_i P_f\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j) + \int_{\mathbb{T}^{N-1}} (P - P_i)\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j) \\ &= P_f\varphi_i + \int_{\mathbb{T}^{N-1}} (P - P_i)\varphi \prod_{j \in [N] \setminus \{i\}} dm(x_j). \end{aligned}$$

□

Proof of Proposition 4.4. From Eq. (4.20) by induction one can easily obtain

$$\varphi_i = (P^n \varphi)_i = P_f^n \varphi_i + \sum_{k=0}^{n-1} P_f^k \int_{\mathbb{T}^{N-1}} (P - P_i) \varphi dm_{N-1}.$$

Now we estimate

$$\begin{aligned} |\varphi_i - \varphi_0|_1 &\leq \|P_f^n(\varphi_i - \varphi_0)\| + \left\| \sum_{k=0}^{n-1} P_f^k \int_{\mathbb{T}^{N-1}} (P - P_i) \varphi dm_{N-1} \right\|_1 \\ &\leq K\lambda^n \|\varphi_i - \varphi_0\| + n \left\| \int_{\mathbb{T}^{N-1}} (P - P_i) \varphi dm_{N-1} \right\|_1. \end{aligned} \quad (4.21)$$

where we used the fact that for a uniformly expanding C^2 map f , P_f has a spectral gap. Let $\psi \in C^\infty(\mathbb{T}, [-1, 1])$

$$\begin{aligned} \int_{\mathbb{T}} dm(x_i) \int_{\mathbb{T}^{N-1}} \psi(x_i) (P - P_i) \varphi dm_{N-1}(\hat{\mathbf{x}}_i) &= \int_{\mathbb{T}^N} (\psi \circ \mathbf{F} - \psi \circ \mathbf{F}_i) \varphi dm_N \\ &= \int_{\mathbb{T}^N} \int_0^1 \partial_t (\psi \circ \mathbf{F}_t) \varphi dt dm_N \\ &= \int_{\mathbb{T}^N} \int_0^1 \psi' \circ \mathbf{F}_t \partial_t F_t^{(i)} \varphi dt dm_N \end{aligned}$$

where we defined $\mathbf{F}_t = t\mathbf{F} + (1-t)\mathbf{F}_i$. Notice that $\mathbf{F}_t$ is again a local diffeomorphism, and the entries of its differential $D\mathbf{F}_t$ satisfy the same bounds of those of $D\mathbf{F}$ so that Proposition 4.2 can be applied. Using Eq. (B.10) we obtain

$$\begin{aligned} \psi' \circ \mathbf{F}_t \partial_t F_t^{(i)} &= \sum_j \partial_j (\psi \circ \mathbf{F} (D\mathbf{F}_t^{-1})_{ij} \partial_t F_t^{(i)}) - \\ &\quad - \psi \circ \mathbf{F}_t [\partial_j (D\mathbf{F}_t^{-1})_{ij} \partial_t F_t^{(i)} + (D\mathbf{F}_t^{-1})_{ij} \partial_j \partial_t F_t^{(i)}] \end{aligned}$$

that allows us to conclude that

$$\sup_{\psi} \int_{\mathbb{T}^N} \psi(x_i) (P - P_i) \varphi dm_N \leq \mathcal{O}(\varepsilon \Delta) \|\varphi\| + \mathcal{O}(\varepsilon \Delta) |\varphi|_1 \leq \mathcal{O}(\varepsilon \Delta).$$

Now consider the function $g(x) = b_1 \lambda^x + b_2 x$, for parameters $b_1 > 0$, $b_2 > 0$ and $\lambda \in (0, 1)$. As can be easily verified, for $b_2 \ll b_1$, g has a minimum at $x_{min} \in [1, \infty)$ given by

$$\lambda^{x_{min}} = -\frac{\mathcal{O}(\varepsilon \Delta)}{K \log \lambda} \quad \Rightarrow \quad x_{min} = \frac{1}{\log \lambda} \log \frac{\mathcal{O}(\varepsilon \Delta)}{-K \log \lambda}.$$

Plugging $n = \lfloor x_{\min} + 1 \rfloor$ in Eq. (4.21), one obtains

$$|\varphi_i - \varphi_0|_1 = \mathcal{O}(\varepsilon \Delta \log(\varepsilon \Delta)).$$

□

4.1.6 One-Dimensional Disintegration along Coordinate Directions

First of all let us recall some basic notation on marginals and disintegrations of measures (standard results are listed in Appendix D). Suppose that μ is a signed measure on $\mathbb{T}^N$ and $\Lambda = \{i_1, \dots, i_k\} \subset [N]$ is a subset of indices with $i_1 < \dots < i_k$. Then we denote by $\pi_\Lambda : \mathbb{T}^N \rightarrow \mathbb{T}^{N-|\Lambda|}$ the projection on the coordinates indexed by Λ , $\pi_\Lambda(x_1, \dots, x_N) = (x_{i_1}, \dots, x_{i_k})$, and by $\mu_\Lambda = \pi_{\Lambda*}(\mu)$ the *marginal* of μ on those coordinates. If $\varphi \in C^0(\mathbb{T}^{N-|\Lambda|}, \mathbb{R})$, then $\mu_\Lambda(\varphi) = \mu(\varphi)$. In particular, we denote by μ_i the marginal on the i -th coordinate and by μ_{i^c} the marginal on all coordinates, but the i -th one.

We will denote by $\{\mu_{\hat{\mathbf{x}}_i}\}_{\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}}$ the disintegration⁴ of μ with respect to $\pi_{\{i\}^c} : \mathbb{T}^N \rightarrow \mathbb{T}^{N-1}$ and the measure $\pi_{\{i\}^c*}(\mu)$. By Proposition D.1, if μ is positive, then $\mu_{\hat{\mathbf{x}}_i}$ is a probability measure on $\mathbb{T}$ for all $\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}$. The following norms control the regularity of $\mu_{\hat{\mathbf{x}}_i}$ uniformly on $i \in [N]$ and $\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}$.

Definition 4.4. Consider μ a measure on $\mathbb{T}^N$. The norms $|\cdot|_{1D}$ ⁵ and $\|\cdot\|_{1D}$ are defined as

$$|\mu|_{1D} := \sup_{i \in [N]} \sup_{\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}} \sup_{\substack{\psi \in C^\infty(\mathbb{T}, \mathbb{R}) \\ |\psi|_\infty < 1}} \mu_{\hat{\mathbf{x}}_i}(\psi) \quad (4.22)$$

$$\|\mu\|_{1D} := \sup_{i \in [N]} \sup_{\hat{\mathbf{x}}_i \in \mathbb{T}^{N-1}} \sup_{\substack{\psi \in C^\infty(\mathbb{T}, \mathbb{R}) \\ |\psi|_\infty < 1}} \mu_{\hat{\mathbf{x}}_i}(D_x \psi). \quad (4.23)$$

Proposition 4.5. *There exists $A \in \mathbb{R}^+$ and $\lambda \in (0, 1)$ depending on f and h only, such that if $\mu \ll m_N$ is a positive measure, and $d\mu/dm_N = \varphi \in \mathcal{C}_a$, then*

$$\|\mathbf{F}_* \mu\|_{1D} \leq \lambda \|\mu\|_{1D} + A |\mu|_{1D}.$$

Remark 4.3. Notice that the condition $d\mu/dm_N = \varphi \in \mathcal{C}_a$ restricts the validity of the Lasota-Yorke to a set of measures that do not form a linear space. This means

⁴See Definition D.1 in the appendix.

⁵ $1D$ stands for one-dimensional disintegration.

that all the usual machinery built around this inequality is not readily available. However, the L-Y gives a control uniform in $i \in [N]$ and $\hat{\mathbf{x}}_i$ on the $W^{1,1}$ and L^1 norms of the disintegrations.

Proof of Proposition 4.5. Consider the disintegration $\{(\mathbf{F}_*\mu)_{\hat{\mathbf{x}}_i}\}_{\hat{\mathbf{x}}_i}$ of $\mathbf{F}_*\mu$ with respect to $\pi_{\{i\}^c}$ and $\pi_{\{i\}^c,*}(\mathbf{F}_*\mu)$. Fix $\psi \in C^\infty(\mathbb{T}, [-1, 1])$. We can use Theorem D.2 to get

$$(\mathbf{F}_*\mu)_{\hat{\mathbf{x}}_i}(\psi') = \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbf{F}_*\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_*\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}$$

Consider the map $\mathbf{F}_i$ defined in Eq. (4.11), then

$$\frac{\mathbf{F}_*\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_*\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})} = \frac{\mathbf{F}_{i*}\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_{i*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})} + \frac{(\mathbf{F}_* - \mathbf{F}_{i*})\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_*\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}. \quad (4.24)$$

For the first factor write

$$\frac{\mathbf{F}_{i*}\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_{i*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})} = \frac{\mathbf{F}_{i*}\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_{i*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})} \frac{\mathbf{F}_{i*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_*\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})},$$

with the bound, obtained applying Lemma 4.1,

$$\frac{\mathbf{F}_{i*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_*\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})} \leq \sup_{\mathbf{x} \in \mathbb{T}^N} \frac{|\mathbf{D}_\mathbf{x} \mathbf{F}_i|}{|\mathbf{D}_\mathbf{x} \mathbf{F}|} \leq \exp\{\mathcal{O}(\varepsilon \Delta)\}, \quad (4.25)$$

and the limit

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\mathbf{F}_{i*}\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})}{\mathbf{F}_{i*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_i)} \circ \pi_{\{i\}^c})} = (\mathbf{F}_{i*}\mu)_{\hat{\mathbf{x}}_i}(\psi').$$

obtained again by the characterisation of disintegration as in Theorem D.2. Since the evolution of the i -th coordinate under $\mathbf{F}_i$ is uncoupled from all the others, one has

$$(\mathbf{F}_{i*}\mu)_{\hat{\mathbf{x}}_i} = \sum_k \alpha_k f_* \mu_{\hat{\mathbf{F}}_{i_k}^{-1}(\hat{\mathbf{x}}_i)} \quad (4.26)$$

where $\{\hat{\mathbf{F}}_{i_k}^{-1}\}_k$ are the inverse branches of $\hat{\mathbf{F}}_i$, and $\{\alpha_k\} \subset [0, 1]$ with $\sum_k \alpha_k = 1$. In fact, since the i -th component is decoupled, the disintegrated measure $\mathbf{F}_{i*}\mu$ at $\hat{\mathbf{x}}_i$ is going to be a superposition of the push forward of disintegrations of μ at the preimages of $\hat{\mathbf{x}}_i$ under $\hat{\mathbf{F}}_i$. That the coefficients have to sum to one is an immediate consequence of the fact that all the disintegrations give probability measures. From

Eq. (4.26) one obtains

$$\begin{aligned} (\mathbf{F}_{\mathbf{i}*}\mu)_{\hat{\mathbf{x}}_{\mathbf{i}}}(\psi') &= \sum_k \alpha_k f_* \left(\mu_{\hat{\mathbf{F}}_{\mathbf{i}k}^{-1}(\hat{\mathbf{x}}_{\mathbf{i}})} \right) (\psi') \leq \sum_k \alpha_k \|f_* \left(\mu_{\hat{\mathbf{F}}_{\mathbf{i}k}^{-1}(\hat{\mathbf{x}}_{\mathbf{i}})} \right)\| \\ &\leq \sum_k \alpha_k \left[\lambda \|\mu_{\hat{\mathbf{F}}_{\mathbf{i}k}^{-1}(\hat{\mathbf{x}}_{\mathbf{i}})}\| + A |\mu_{\hat{\mathbf{F}}_{\mathbf{i}k}^{-1}(\hat{\mathbf{x}}_{\mathbf{i}})}| \right] \end{aligned} \quad (4.27)$$

$$\leq \lambda \|\mu\|_{1D} + A |\mu|_{1D} \quad (4.28)$$

where to get (4.28) we used the Lasota-Yorke inequality for f . Combining Eq. (4.25) and Eq. (4.28) we obtain

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\mathbf{F}_{\mathbf{i}*}\mu(\psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c})}{\mathbf{F}_{\mathbf{i}*}\mu(1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c})} \leq \exp\{\mathcal{O}(\varepsilon\Delta)\} [\lambda \|\mu\|_{1D} + A |\mu|_{1D}].$$

Now we estimate the second term in Eq. (4.24).

$$\begin{aligned} \frac{\int_{\mathbb{T}^N} (P\varphi - P_i\varphi) \psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P\varphi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu} &= \frac{\int_{\mathbb{T}^N} P\varphi (1 - \frac{P_i\varphi}{P\varphi}) \psi' 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P\varphi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu} \\ &= \frac{\int_{\mathbb{T}^N} P\varphi \partial_{x_i} [(1 - \frac{P_i\varphi}{P\varphi}) \psi] 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P\varphi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu} + \\ &\quad - \frac{\int_{\mathbb{T}^N} P\varphi \partial_{x_i} (\frac{P_i\varphi}{P\varphi}) \psi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P\varphi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}. \end{aligned}$$

Using Eq. (4.44) and Eq. (4.45) one gets

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\int_{\mathbb{T}^N} P\varphi \partial_{x_i} [(1 - \frac{P_i\varphi}{P\varphi}) \psi] 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P\varphi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu} \leq [\mathcal{O}(\varepsilon\Delta^2)] \|\mu\|_{1D} \quad (4.29)$$

and

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\int_{\mathbb{T}^N} P\varphi \partial_{x_i} (\frac{P_i\varphi}{P\varphi}) \psi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P\varphi 1_{B_\varepsilon(\hat{\mathbf{x}}_{\mathbf{i}})} \circ \pi_{\{i\}^c} d\mu} \leq A |\mu|_{1D}.$$

All these estimates together prove the Lasota-Yorke inequality. $\square$

The next result shows that for sparsely coupled uniformly expanding maps the disintegrations are close in the C^0 norm. Recall that μ_0 is the a.c.i.p. measure for f and φ_0 is its density.

Proof of Theorem 4.1. Pick $i \in [N]$, $\hat{\mathbf{x}}_{\mathbf{i}} \in \mathbb{T}^{N-1}$ then for every $n \in \mathbb{N}$

$$\|\nu_{\hat{\mathbf{x}}_{\mathbf{i}}} - \mu_0\|_{C^0} = \|(\mathbf{F}_{\mathbf{i}*}^n \nu)_{\hat{\mathbf{x}}_{\mathbf{i}}} - f_*^n \mu_0\|_{C^0}$$

$$\leq \|(\overline{\mathbf{F}_{i*}^n \nu})_{\hat{\mathbf{x}}_i} - f_*^n \mu_0\|_{C^0} + \|(\overline{\mathbf{F}_{i*}^n \nu})_{\hat{\mathbf{x}}_i} - (\mathbf{F}_{i*}^n \nu)_{\hat{\mathbf{x}}_i}\|_{C^0}$$

where we put a bar on top of $(\mathbf{F}_{i*}^n \nu)$ to indicate that the disintegration is with respect to $\mathbf{F}_{i*}^n \nu$ (rather than the usual $\mathbf{F}_{i*}^n$). Recalling that the Sobolev functions $W^{1,1}$ on a one-dimensional space are compactly embedded into C^0 , there exists a $C_{\#} \in \mathbb{R}^+$ such that

$$\|\nu_{\hat{\mathbf{x}}_i} - \mu_0\|_{C^0} \leq C_{\#} \|(\overline{\mathbf{F}_{i*}^n \nu})_{\hat{\mathbf{x}}_i} - f_*^n \mu_0\|_{W^{1,1}} + \|(\mathbf{F}_{i*}^n \nu - \mathbf{F}_{i*}^n \nu)_{\hat{\mathbf{x}}_i}\|_{C^0}.$$

We get that

$$\begin{aligned} \|(\overline{\mathbf{F}_{i*}^n \nu})_{\hat{\mathbf{x}}_i} - f_*^n \mu_0\|_{W^{1,1}} &= \|(\overline{\mathbf{F}_{i*}^n \nu})_{\hat{\mathbf{x}}_i} - (\mathbf{F}_{i*}^n \nu)_{\hat{\mathbf{x}}_i} + (\mathbf{F}_{i*}^n \nu)_{\hat{\mathbf{x}}_i} - f_*^n \mu_0\|_{W^{1,1}} \\ &\leq n\mathcal{O}(\varepsilon\Delta) \|(\mathbf{F}_{i*}^n \nu)_{\hat{\mathbf{x}}_i}\|_{W^{1,1}} + \left\| \sum_k \alpha_k [f_*^n \nu_{\hat{\mathbf{F}}_{i_k}^{-n}(\hat{\mathbf{x}}_i)} - f_*^n \mu_0] \right\|_{W^{1,1}} \end{aligned}$$

with $\sum_k \alpha_k = 1$. Since the transfer operator f_* , has a spectral gap on $W^{1,1}$ one has that

$$\begin{aligned} \left\| \sum_k \alpha_k [f_*^n \nu_{\hat{\mathbf{F}}_{i_k}^{-n}(\hat{\mathbf{x}}_i)} - f_*^n \mu_0] \right\|_{W^{1,1}} &\leq \sum_k \alpha_k \|f_*^n [\nu_{\hat{\mathbf{F}}_{i_k}^{-n}(\hat{\mathbf{x}}_i)} - \mu_0]\|_{W^{1,1}} \\ &\leq M_{\#} \lambda^n \|\nu_{\hat{\mathbf{F}}_{i_k}^{-n}(\hat{\mathbf{x}}_i)} - \mu_0\|_{W^{1,1}} \\ &\leq M_{\#} \lambda^n \end{aligned}$$

where for the last inequality we used that $\|\nu_{\hat{\mathbf{F}}_{i_k}^{-n}(\hat{\mathbf{x}}_i)} - \mu_0\|_{W^{1,1}} < \|\nu\|_{1D} + \|\mu_0\|_{W^{1,1}}$ and thus is uniformly bounded. Now we need the following estimate

$$\begin{aligned} \|(\overline{\mathbf{F}_{i*}^n \nu})_{\hat{\mathbf{x}}_i} - (\mathbf{F}_{i*}^n \nu)_{\hat{\mathbf{x}}_i}\|_{C^0} &= \sup_{x \in \mathbb{T}} \left| \frac{(P_i^n \varphi)(x; \hat{\mathbf{x}}_i)}{\int_{\mathbb{T}} P^n \varphi(y; \hat{\mathbf{x}}_i) dy} - \frac{(P^n \varphi)(x; \hat{\mathbf{x}}_i)}{\int_{\mathbb{T}} P^n \varphi(y; \hat{\mathbf{x}}_i) dy} \right| \\ &\leq \sup_{x \in \mathbb{T}} \left| \frac{(P_i^n \varphi)(x; \hat{\mathbf{x}}_i)}{\int_{\mathbb{T}} P^n \varphi(y; \hat{\mathbf{x}}_i) dy} \right| \left| \frac{(P_i^n \varphi)(x; \hat{\mathbf{x}}_i)}{(P^n \varphi)(x; \hat{\mathbf{x}}_i)} - 1 \right| \\ &\leq n[\mathcal{O}(\varepsilon\Delta^2)]. \end{aligned}$$

where we used Lemma 4.1, and the fact that one has an n -fold composition of the transfer operators. Putting all estimates together, for every $n \in \mathbb{N}$

$$\|\nu_{\hat{\mathbf{x}}_i} - \mu_0\|_{C^0} \leq M_{\#} \lambda^n + n[\mathcal{O}(\varepsilon\Delta^2)]$$

and the conclusion follows. $\square$

4.1.7 Spatial Large Deviations

Now we use Theorem 4.1 on the a.c.i.p. measure ν for $\mathbf{F}$, to estimate $\nu(\mathcal{B}_{\xi,N})$ comparing $\varphi(x_1, \dots, x_N) := d\nu/dm$ with the product density $\prod_{j \in [N]} \varphi_0(x_j)$.

Proof of Corollary 4.1.

$$\int_{\mathcal{B}_{\xi,N}} \varphi(\mathbf{x}) dm_N(\mathbf{x}) = \int_{\mathcal{B}_{\xi,N}} \prod_{j \in [N]} \varphi_0(x_j) \left[\frac{\varphi(\mathbf{x})}{\prod_{j \in [N]} \varphi_0(x_j)} \right] dm_N(\mathbf{x}).$$

We can write the telescopic product

$$\frac{\varphi}{\prod_{j \in [N]} \varphi_0(x_j)} = \prod_{k=0}^{N-1} \frac{\int \varphi(x_1, \dots, x_N) dm_k(x_1, \dots, x_k)}{\varphi_0(x_{k+1}) \int \varphi(x_1, \dots, x_N) dm_{k+1}(x_1, \dots, x_{k+1})}.$$

Given $k \in [N]$, we know from the estimates on the one-dimensional disintegrations in Theorem 4.1 that, for every $(x_{k+2}, \dots, x_N) \in \mathbb{T}^{N-k-2}$,

$$\frac{\int \varphi(x_1, \dots, x_N) dm_k(x_1, \dots, x_k)}{\int \varphi(x_1, \dots, x_N) dm_{k+1}(x_1, \dots, x_{k+1})} = \varphi_0(x_{k+1}) + *,$$

where $*$ is a function bounded by $\mathcal{O}(\varepsilon \Delta^2 \log(\varepsilon \Delta^2))$. Thus

$$\frac{\int \varphi(x_1, \dots, x_N) dm_k(x_1, \dots, x_k)}{\varphi_0(x_{k+1}) \int \varphi(x_1, \dots, x_N) dm_{k+1}(x_1, \dots, x_{k+1})} = 1 + *.$$

where with an abuse of notation we used the same symbol $*$ to indicate $*/\varphi_0$. The above implies that for every $\mathbf{x} \in \mathbb{T}^N$

$$\left| \frac{\varphi(\mathbf{x})}{\prod_{j \in [N]} \varphi_0(x_j)} \right| \leq (1 + \mathcal{O}(\varepsilon \Delta^2 \log(\varepsilon \Delta^2)))^N = \exp\{\mathcal{O}(\varepsilon \Delta^2 \log(\varepsilon \Delta^2))N\},$$

and so

$$\int_{\mathcal{B}_{\xi,N}} \varphi(\mathbf{x}) dm_N(\mathbf{x}) \leq \exp [C(\mathcal{O}(\varepsilon \Delta^2 \log(\varepsilon \Delta^2)) - \xi^2)N].$$

□

4.2 HCM with Local Dynamics having Nonzero Distortion

We now move to the case of HCM with uniformly expanding local maps having nonzero distortion, and with reduced dynamics of the hub nodes being uniformly

expanding. The notation that we use is the same of Chapter 3 with the difference that now f is any C^2 uniformly expanding map. The strategy is going to be the same that was outlined at the beginning of the chapter. We are going to exploit the extra knowledge that we have on HCM to infer the result. We prove the analogous of Theorem 4.1.

Theorem 4.3. *There exists $\eta > 0$ depending on f , α , and h only such that if $\Delta^{-4}L^3M^2\delta^{3/2} < \eta$, and (H1) and (H3) are satisfied, then, for all $\varepsilon < \varepsilon_0$ (as in Proposition 3.2), $\mathbf{F}_\varepsilon$ has an a.c.i.p. measure ν and*

$$\|\nu_{\hat{\mathbf{z}}_i} - \mu_0\|_{C^0} \leq \mathcal{O}(\Xi \log(\Xi))$$

for all $i \in [L]$ and all $\hat{\mathbf{z}}_i \in \mathbb{T}^{N-1}$ where μ_0 is the a.c.i.p. measure for f and

$$\Xi := \mathcal{O}(\Delta^{-1}\delta) + \mathcal{O}(\Xi_l \Delta^{-1}L\delta) + \mathcal{O}(\Xi_H \Delta^{-1}M\delta)$$

and

$$\Xi_l := \Delta^{-3/2}\delta M + \Delta^{-3}L\delta^{5/2}, \quad (4.30)$$

$$\Xi_H := \Delta^{-3/2}M^2 + \Delta^{-2}L\delta^{3/2}. \quad (4.31)$$

Once again one can use the estimates on one-dimensional disintegrations to obtain results on spatial large deviations. In the case of HCM these estimates can be applied to infer information on the mean-field, as was done in Section 3.3, and prove Theorem A in the case the local map f has nonzero distortion.

4.2.1 Global Expansion of $\mathbf{F}_\varepsilon$

It is easy to verify that the proof of Proposition 3.2 does not rely on the distortionless property of the local maps f . This implies that as long as there exists $\sigma > 1$ such that $|Df| \geq \sigma$ the same conclusions on expansivity of $\mathbf{F}_\varepsilon$ hold. From now on we suppose that $\varepsilon < \varepsilon_0$ as prescribed by 3.2 and that the heterogeneity conditions ensure $\mathbf{F}_\varepsilon$ is expanding.

4.2.2 Distortion of $\mathbf{F}_\varepsilon$

In the following proposition we compute the distortion of the Jacobian of the map $\mathbf{F}_\varepsilon$.

Proposition 4.6. *There exists $\eta > 0$ depending only on f , α and h , such that if*

$\Delta^{-4}L^3M^2\delta^{3/2} < \eta$ and if $\varepsilon < \varepsilon_0$ then for every $\mathbf{z}, \bar{\mathbf{z}} \in \mathbb{T}^N$

$$\frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}|} \leq \exp \left\{ \mathcal{O}(1) \sum_{k=1}^N d(z_k, \bar{z}_k) \right\}. \quad (4.32)$$

Remark 4.4. Notice that $d_{\infty}(\mathbf{z}, \bar{\mathbf{z}}) := \max_{n \in [N]} d(z_n, \bar{z}_n)$, so from the above we have

$$\frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}|} \leq \exp \{ \mathcal{O}(N) d_{\infty}(\mathbf{z}, \bar{\mathbf{z}}) \}$$

and from $d_1(\mathbf{z}, \bar{\mathbf{z}}) := \sum_{n=1}^N d(z_n, \bar{z}_n)$ we get

$$\frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}|} \leq \exp \{ \mathcal{O}(1) d_1(\mathbf{z}, \bar{\mathbf{z}}) \}.$$

This informs us that, considered the very high order of distortions, the usual cones of functions on $\mathbb{T}^N$ are only going to give very weak regularity property of the invariant density, in particular, not sharp enough to be useful for the large deviations estimates we are interested in.

We are going to use the following lemma which is obtained applying Lemma B.1 on a matrix $\mathbf{M}$ of the form of the differential of $\mathbf{F}_{\varepsilon}$.

Lemma 4.3. *Suppose that $\mathbf{M} \in \mathcal{M}(N, N)$ is a matrix whose entries satisfy the following conditions:*

1. *there exists $c \in \mathbb{R}^+$ such that $|M_{ii}| \geq c$ for all $i \in [N]$,*
2. *there exists $\mathbf{A} \in \mathcal{M}(N, N)$ the adjacency matrix of an heterogeneous graph with global parameters L, M, Δ, δ as in Section 3.1 such that $|M_{ij}| \leq \mathcal{O}(\Delta^{-1})A_{ij}$ for all $i \neq j$, and $A_{jj} = 0 \ \forall j \in [N]$.*

then there exists $\eta > 0$ uniform on the structural parameters such that if $\Delta^{-4}L^3M^2\delta^{3/2} < \eta$ then if $\mathbf{v} \in \text{Span}(\mathbf{M}_1, \dots, \mathbf{M}_{j-1}, \mathbf{M}_{j+1}, \dots, \mathbf{M}_N)^{\perp}$ and $\|\mathbf{v}\|_2 = 1$.

$$|v_k| \leq \begin{cases} \mathcal{O}(\Delta^{-2}L\delta^{3/2}) + \mathcal{O}(\Delta^{-3/2}M\delta) & k \in [L] \setminus \{j\} \\ \mathcal{O}(\Delta^{-1/2}M) + \mathcal{O}(\Delta^{-2}L\delta^{3/2}) & k \in [L+1, N] \setminus \{j\} \end{cases} \quad (4.33)$$

and $|v_j| \geq c_{\#}$, where $c_{\#} > 0$ depend on f, α and h only.

Proof. From Lemma B.1 we obtain estimates for the entries v_k of $\mathbf{v}$ with $k \neq j$.

$$|v_k| \leq \mathcal{O}(\Delta^{-1})A_{kj} + \mathcal{O}(\Delta^{-2}) \sum_{\ell_1 \neq j} \sum_{\ell_2 \neq j} A_{k\ell_1} A_{\ell_1\ell_2} \sqrt{d_{\mathcal{O}, \ell_2}} \quad (4.34)$$

$$\leq \mathcal{O}(\Delta^{-1})A_{kj} + \mathcal{O}(\Delta^{-2}) \left[\sum_{\ell_2=1}^L \sum_{\ell_1 \neq j} A_{k\ell_1} A_{\ell_1\ell_2} \sqrt{\delta} + \sum_{\ell_2=L+1}^N \sum_{\ell_1 \neq j} A_{k\ell_1} A_{\ell_1\ell_2} \sqrt{\Delta} \right]. \quad (4.35)$$

If $k \in [L] \setminus \{j\}$

$$\begin{aligned} |v_k| &\leq \mathcal{O}(\Delta^{-1})A_{kj} + \mathcal{O}(\Delta^{-2})[L\delta\sqrt{\delta} + M\delta\sqrt{\Delta}] \\ &\leq \mathcal{O}(\Delta^{-2}L\delta^{3/2}) + \mathcal{O}(\Delta^{-3/2}M\delta) \end{aligned} \quad (4.36)$$

in fact for every $\ell_2 \in [N]$, $\sum_{\ell_1=1}^N A_{k\ell_1} A_{\ell_1\ell_2}$ counts the paths of length 2 going from ℓ_2 to k . Since the node $k \in [L]$ has at most degree δ , for every ℓ_2 the maximum number of such paths is δ . If $k \in [L+1, N] \setminus \{j\}$ one can analogously find

$$\begin{aligned} |v_k| &\leq \mathcal{O}(\Delta^{-1})A_{kj} + \mathcal{O}(\Delta^{-2}) \left[M\Delta^{3/2} + L\delta^{3/2} \right] \\ &\leq \mathcal{O}(\Delta^{-1/2}M) + \mathcal{O}(\Delta^{-2}L\delta^{3/2}). \end{aligned} \quad (4.37)$$

The lower bound for $|v_j|$ is

$$\begin{aligned} |v_j| &= \left(1 - \sum_{k \neq j} |v_k|^2 \right)^{\frac{1}{2}} \\ &\geq \left(1 - \sum_{k=1}^N A_{jk} \mathcal{O}(\Delta^{-2}) - \sum_{k=1}^L [\mathcal{O}(\Delta^{-3/2}M\delta) + \mathcal{O}(\Delta^{-2}L\delta^{3/2})]^2 \right. \\ &\quad \left. - \sum_{k=L+1}^M [\mathcal{O}(\Delta^{-1/2}M) + \mathcal{O}(\Delta^{-2}L\delta^{3/2})]^2 \right)^{\frac{1}{2}} \end{aligned}$$

so if $j \in [L]$

$$\begin{aligned} |v_j| &\geq (1 - \mathcal{O}(\Delta^{-2}\delta) - \mathcal{O}(L\Delta^{-3}M^2\delta^2) - \mathcal{O}(L^3\Delta^{-4}\delta^3) - \mathcal{O}(\Delta^{-1}M^3) - \mathcal{O}(\Delta^{-4}L^2M^2\delta^3))^{\frac{1}{2}} \\ &\geq (1 - \mathcal{O}(L^3\Delta^{-4}\delta^3M^2))^{\frac{1}{2}}. \end{aligned}$$

So for any fixed $c_{\#} > 0$ there is $\eta > 0$ uniform in the structural parameters such that $L^3\Delta^{-4}\delta^3M^2 < \eta$ implies $|v_j| > c_{\#}$. $\square$

Proof of Proposition 4.6. Define matrices $\{\mathbf{D}_k\}_{k=0}^N$ such that $\mathbf{D}^{(0)} := \mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}$, $\mathbf{D}^{(N)} := \mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}$ and $\mathbf{D}^{(k)}$ and $\mathbf{D}^{(k+1)}$ differ only for the $(k+1)$ -th row. One has the

telescopic product

$$\frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}|} = \prod_{k=1}^N \frac{|\mathbf{D}^{(k-1)}|}{|\mathbf{D}^{(k)}|}.$$

For every $k \in [N]$

$$\frac{|\mathbf{D}^{(k-1)}|}{|\mathbf{D}^{(k)}|} = \frac{\mathbf{D}_k^{(k-1)} \mathbf{v}^{(k)}}{\mathbf{D}_k^{(k)} \mathbf{v}^{(k)}}$$

that is the ratio of the scalar products between the k -th rows of the two matrices and $\mathbf{v}^{(k)}$, which is the vector orthogonal to all the rows of $\mathbf{D}^{(k)}$ apart from the k -th one. This means that

$$\frac{|\mathbf{D}^{(k-1)}|}{|\mathbf{D}^{(k)}|} = \frac{|\sum_i D_{ki}^{(k-1)} v_i^{(k)}|}{|\sum_i D_{ki}^{(k)} v_i^{(k)}|} \leq 1 + \frac{\sum_i |D_{ki}^{(k-1)} - D_{ki}^{(k)}| |v_i^{(k)}|}{|\sum_i D_{ki}^{(k)} v_i^{(k)}|}$$

and so

$$\begin{aligned} \frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}|} &\leq \exp \left\{ \mathcal{O} \left(\sum_{k,i} |D_{ki}^{(k-1)} - D_{ki}^{(k)}| |v_i^{(k)}| \right) \right\} \\ &\leq \exp \left\{ \mathcal{O} \left(\sum_{k \neq i} |D_{ki}^{(k-1)} - D_{ki}^{(k)}| |v_i^{(k)}| + \sum_k |D_{kk}^{(k-1)} - D_{kk}^{(k)}| |v_k^{(k)}| \right) \right\} \\ &\leq \exp \left\{ \mathcal{O} \left(\sum_{k \neq i} |D_{ki}^{(k-1)} - D_{ki}^{(k)}| [\mathcal{O}(\Delta^{-1/2}M) + \mathcal{O}(\Delta^{-2}L\delta^{3/2})] + \sum_n |D_{kk}^{(k-1)} - D_{kk}^{(k)}| \right) \right\} \end{aligned}$$

Notice that $D_{ki}^{(k-1)} - D_{ki}^{(k)} = (\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon})_{ki} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon})_{ki}$ for every $k, i \in [N]$. In Lemma B.3 we prove that, $|(\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon})_{ki} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon})_{ki}| \leq$

$$\begin{cases} A_{ki} \mathcal{O}(\Delta^{-1}) [d(z_k, \bar{z}_k) + d(z_i, \bar{z}_i)] & k \in [L], i \neq k \\ \mathcal{O}(1) d(z_k, \bar{z}_k) + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d(z_n, \bar{z}_n) & k \in [L], i = k \\ A_{ki} \mathcal{O}(\Delta^{-1}) [d(z_i, \bar{z}_i) + d(z_k, \bar{z}_k) + \mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L A_{k\ell} d(z_{\ell}, \bar{z}_{\ell})] & k \in [L+1, N], i \in [L] \\ A_{ki} \mathcal{O}(\Delta^{-1}) [d(z_k, \bar{z}_k) + d(z_i, \bar{z}_i)] & k, i \in [L+1, N], i \neq k \\ \mathcal{O}(1) d(z_k, \bar{z}_k) + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d(z_n, \bar{z}_n) & k \in [L+1, N], i = k \end{cases} \quad (4.38)$$

Let d_n be the shorthand notation for $d_n := d(z_n, \bar{z}_n)$ for all $n \in [N]$. Then

$$\sum_{k \neq i} |D_{ki}^{(k-1)} - D_{ki}^{(k)}| = \sum_{k=1}^N \sum_{\substack{i=1 \\ i \neq k}}^N |(\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon})_{ki} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon})_{ki}|$$

$$\begin{aligned}
&\leq \sum_{k=1}^L \sum_{\substack{i=1 \\ i \neq k}}^N A_{ki} \mathcal{O}(\Delta^{-1}) [d_k + d_i] + \\
&\quad + \sum_{k=L+1}^N \sum_{i=1}^L A_{ki} \mathcal{O}(\Delta^{-1}) [d_i + d_k + \mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L A_{k\ell} d_\ell] + \\
&\quad + \sum_{k=L+1}^N \sum_{\substack{i=L+1 \\ i \neq k}}^N A_{ki} \mathcal{O}(\Delta^{-1}) [d_k + d_i] \\
&\leq \sum_{k=1}^N \sum_{i=1}^N A_{ki} \mathcal{O}(\Delta^{-1}) [d_k + d_i] + \mathcal{O}(\Delta^{-2}) \sum_{\ell=1}^L d_\ell \sum_{k=L+1}^N A_{k\ell} \sum_{i=1}^L A_{ki} \\
&\leq \sum_{k=1}^N \sum_{i=1}^N A_{ki} \mathcal{O}(\Delta^{-1}) [d_k + d_i] + \mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L d_\ell \sum_{k=L+1}^N A_{k\ell} \\
&\leq \sum_{k=1}^N \sum_{i=1}^N A_{ki} \mathcal{O}(\Delta^{-1}) [d_k + d_i] + \mathcal{O}(\Delta^{-1} \delta) \sum_{\ell=1}^L d_\ell.
\end{aligned}$$

The double sum on the last line can be simplified further observing that

$$\begin{aligned}
\sum_{k=1}^N \sum_{i=1}^N A_{ki} [d_k + d_i] &= \sum_{k=1}^N d_k \sum_{i=1}^N A_{ki} + \sum_{i=1}^N d_i \sum_{k=1}^N A_{ki} \\
&= \sum_{k=1}^N d_k d_{I,k} + \sum_{i=1}^N d_i d_{O,i} \\
&= \sum_{k=1}^N d_k [d_{I,k} + d_{O,k}] \\
&= \mathcal{O}(\delta) \sum_{\ell=1}^L d_\ell + \mathcal{O}(\Delta) \sum_{m=1}^M d_m.
\end{aligned}$$

The term that remain to be estimated is

$$\begin{aligned}
\sum_{k=1}^N |D_{kk}^{(k-1)} - D_{kk}^{(k)}| &= \sum_{k=1}^N |(\mathbf{D}_z \mathbf{F}_\varepsilon)_{kk} - (\mathbf{D}_{\bar{z}} \mathbf{F}_\varepsilon)_{kk}| \leq \\
&\leq \sum_{k=1}^L \mathcal{O}(1) d_k + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d_n + \sum_{k=L+1}^N \mathcal{O}(1) d_k + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d_n \\
&\leq \mathcal{O}(1) \sum_{k=1}^N d_k + \mathcal{O}(\Delta^{-1}) \sum_{k,n=1}^N A_{kn} d_n \\
&\leq \mathcal{O}(1) \sum_{k=1}^N d_k + \mathcal{O}(\delta \Delta^{-1}) \sum_{\ell=1}^L d_\ell + \mathcal{O}(1) \sum_{m=1}^M d_m.
\end{aligned}$$

Putting everything together we obtain

$$\frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_{\varepsilon}|} \leq \exp\{\Xi_J\}$$

with Ξ_J equal to

$$[\mathcal{O}(\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L\delta^{3/2}) + \mathcal{O}(\Delta^{-1})] \left[\mathcal{O}(\delta) \sum_{\ell=1}^L d_{\ell} + \mathcal{O}(\Delta) \sum_{m=1}^M d_m \right] + \mathcal{O}(1) \sum_{k=1}^N d_k$$

so there is η' such that if $\Delta^{-1/2}M + \Delta^{-2}L\delta^{3/2} < \eta'$, then the bound in Eq. (4.32) follows. In particular, choosing η sufficiently small and $L^3\Delta^{-4}\delta^3M^2 < \eta$ implies the previous bound. $\square$

4.2.3 Invariant Cones of Functions Regular along Each Coordinate

As we have done for the sparsely coupled systems we can estimate the entries of $\mathbf{D}_{\mathbf{x}}\mathbf{F}_{\varepsilon}^{-1}$. The proof follows the same steps of Proposition 4.2, but we are going to keep into account the heterogeneous structure of the network. Notice that since the network is heterogeneous its diameter can be very small, so the construction in Definition 4.1 does not provide much information on the dependencies since, in principle, all the nodes can be included in the sets $\mathcal{L}_1^{(i)}$ and $\mathcal{L}_2^{(i)}$. In this section we obtain the following result

Theorem 4.4. *There is $\eta > 0$ depending on f , h , and α only such that if $L^3\Delta^{-4}\delta^3M^2 < \eta$ then there is $a \in \mathbb{R}^+$, depending on f and h only such that if φ is the density of the a.c.i.p. measure for the map $\mathbf{F}_{\varepsilon}$, then for all $i \in [N]$,*

$$\frac{\varphi(z_i; \hat{\mathbf{z}}_i)}{\varphi(\bar{z}_i; \hat{\mathbf{z}}_i)} < \exp[ad(\bar{z}_i, z_i)], \quad \forall \hat{\mathbf{z}}_i \in \mathbb{T}^{N-1}.$$

Lemma 4.4. *There is $\eta > 0$, depending on f , h and α only, such that if $\Delta^{-4}L^3M^2\delta^{3/2} < \eta$, then for all $i, j \in [N]$, $i \neq j$, and $\mathbf{z} \in \mathbb{T}^N$*

$$(\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}^{-1})_{ij} \leq \begin{cases} \mathcal{O}(\Xi_l) & i \in [L] \\ \mathcal{O}(\Xi_H) & i \in [L+1, N] \end{cases}$$

and

$$(\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}^{-1})_{ii} \leq \frac{1}{c - \mathcal{O}(\Xi_H)}.$$

Proof. Call $\mathbf{M} := \mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}$. Having assumed that $\varepsilon < \varepsilon_0$, the entries of $\mathbf{M}$ satisfy $|M_{ii}| \geq c$, and any other nonzero entry has modulus less than $\mathcal{O}(\Delta^{-1})$. Now we

compare the determinants of $\widehat{\mathbf{M}}_{\mathbf{ii}}$ and $\widehat{\mathbf{M}}_{\mathbf{ji}}$ for a fixed $i \in [N]$ (one can relabel the nodes and obtain a comparison for any index i). Notice that these two matrices are the same except for one row. Call $\widetilde{\mathbf{M}} \in \mathcal{M}(N, N-1)$ the matrix obtained from $\mathbf{M}$ deleting the i -th column, and $\overline{\mathbf{M}} \in \mathcal{M}(N-1, N-2)$ obtained from $\mathbf{M}$ eliminating the i -th column and row, and the j -th row. Define the linear application $\mathcal{L} : \mathbb{R}^{N-1} \rightarrow \mathbb{R}$

$$\mathcal{L}(\mathbf{w}) = \det \begin{pmatrix} \overline{\mathbf{M}}_1 \\ \vdots \\ \overline{\mathbf{M}}_{N-2} \\ \mathbf{w}^t \end{pmatrix}$$

in particular $|\widehat{\mathbf{M}}_{\mathbf{NN}}| = |\mathcal{L}(\mathbf{M}_j)|$ and $|\widehat{\mathbf{M}}_{\mathbf{jN}}| = |\mathcal{L}(\mathbf{M}_N)|$. From the geometric characterisation of the determinant as volume of the polyhedron with sides the row of the matrix, we know that $\|\mathcal{L}\| = |\mathcal{L}(\mathbf{v})|$ where $\mathbf{v} \in \text{Span}(\overline{\mathbf{M}}_1, \dots, \overline{\mathbf{M}}_{N-2})^\perp$ and $\|\mathbf{v}\|_2 = 1$. Furthermore, there exists $K \in \mathbb{R} \setminus \{0\}$ such that $\mathcal{L}(\mathbf{w}) = K \mathbf{w}^t \mathbf{v}$, which means that the value of $\mathcal{L}(\mathbf{w})$ is entirely determined by the scalar product of $\mathbf{w}$ and $\mathbf{v}$. ($K \neq 0$ from the invertibility of $\mathbf{M}$). In particular,

$$\frac{|\widehat{\mathbf{M}}_{\mathbf{jN}}|}{|\widehat{\mathbf{M}}_{\mathbf{NN}}|} = \left| \frac{\widetilde{\mathbf{M}}_N \mathbf{v}}{\widetilde{\mathbf{M}}_j \mathbf{v}} \right|.$$

Since we are under the hypotheses of Lemma 4.3

$$|\widehat{\mathbf{M}}_{\mathbf{ii}}| = K \mathbf{M}_j \mathbf{v} \geq K \begin{cases} c_\# + \mathcal{O}(\delta \Delta^{-3/2} M) + \mathcal{O}(\Delta^{-3} L \delta^{5/2}) & i \in [L] \\ c_\# + \mathcal{O}(\Delta^{-3/2} M^2) + \mathcal{O}(\Delta^{-2} L \delta^{3/2}) & i \in [L+1, N] \end{cases}$$

and

$$|\widehat{\mathbf{M}}_{\mathbf{ji}}| = K \mathbf{M}_i \mathbf{v} \leq K \begin{cases} \mathcal{O}(\delta \Delta^{-3/2} M) + \mathcal{O}(\Delta^{-3} L \delta^{5/2}) & i \in [L] \\ \mathcal{O}(\Delta^{-3/2} M^2) + \mathcal{O}(\Delta^{-2} L \delta^{3/2}) & i \in [L+1, N] \end{cases}$$

Furthermore, expanding the determinant of M along the i -th column

$$\begin{aligned} |\mathbf{M}| &\geq |M_{ii}| |\widehat{\mathbf{M}}_{\mathbf{ii}}| - \left| \sum_{\ell \neq i} (-1)^\ell M_{\ell i} \det \widehat{\mathbf{M}}_{\ell i} \right| \\ &\geq c |\widehat{\mathbf{M}}_{\mathbf{ii}}| - \sum_{\ell=1}^L A_{\ell i} \mathcal{O}(\Delta^{-1}) |\widehat{\mathbf{M}}_{\ell i}| - \sum_{\ell=L+1}^N A_{\ell i} \mathcal{O}(\Delta^{-1}) |\widehat{\mathbf{M}}_{\ell i}| \end{aligned} \quad (4.39)$$

$$\geq K c_\# - K \Delta \mathcal{O}(\Delta^{-1}) [\mathcal{O}(\delta \Delta^{-3/2} M) + \mathcal{O}(\Delta^{-3} L \delta^{5/2})] - \quad (4.40)$$

$$\begin{aligned}
& -KM\mathcal{O}(\Delta^{-1})[O(\Delta^{-3/2}M^2) + \mathcal{O}(\Delta^{-2}L\delta^{3/2})] \\
& \geq K[c_{\#} - \mathcal{O}(\Delta^{-3}L\delta^{5/2}) + \mathcal{O}(\Delta^{-3/2}M^2) + \mathcal{O}(\Delta^{-3/2}\delta M)].
\end{aligned} \tag{4.41}$$

Using the analytical formula for the inverse of a square matrix we obtain

$$|(\mathbf{M}^{-1})_{ij}| = \frac{|\widehat{\mathbf{M}}_{ji}|}{|\mathbf{M}|} \leq \begin{cases} \mathcal{O}(\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L\delta^{5/2}) & i \in [L] \\ \mathcal{O}(\Delta^{-3/2}M^2) + \mathcal{O}(\Delta^{-2}L\delta^{3/2}) & i \in [L+1, N] \end{cases}$$

$$\begin{aligned}
|(\mathbf{M}^{-1})_{ii}| &= \frac{|\widehat{\mathbf{M}}_{ii}|}{|\mathbf{M}|} \leq \frac{1}{c - \sum_{\ell=1}^L A_{\ell i} \mathcal{O}(\Delta^{-1}) \frac{|\widehat{\mathbf{M}}_{\ell i}|}{|\widehat{\mathbf{M}}_{ii}|} - \sum_{\ell=L+1}^N A_{\ell i} \mathcal{O}(\Delta^{-1}) \frac{|\widehat{\mathbf{M}}_{\ell i}|}{|\widehat{\mathbf{M}}_{ii}|}} \\
&\leq \frac{1}{c - \mathcal{O}(\Delta^{-3/2}M^2) - \mathcal{O}(\Delta^{-2}L\delta^{3/2})}.
\end{aligned}$$

Recalling the expressions (4.30) and (4.31) for Ξ_l and Ξ_H , the claim follows. $\square$

Proposition 4.7. *For all $\varphi \in \mathcal{C}_a$ and $i \in [N]$, $P\varphi \in \mathcal{C}_{i,(\lambda a+C)}$ where*

$$\begin{aligned}
\lambda &:= \left[(\mathbf{D}\mathbf{F}_{\varepsilon}^{-1})_{ii} + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2}) \right] \\
C &:= \left[\mathcal{O}(1) + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2}) \right].
\end{aligned}$$

Proof. Take $\varphi \in \mathcal{C}_a$. Take $\mathbf{z} \in \mathbb{T}^N$ and using the same notation as in Proposition 4.7, call $\mathbf{z}' = \mathbf{F}_{\varepsilon(m)}^{-1}(z_1, \dots, z_i, \dots, z_N)$ and $\bar{\mathbf{z}}' = \mathbf{F}_{\varepsilon(m)}^{-1}(z_1, \dots, \bar{z}_i, \dots, z_N)$ where $\{\mathbf{F}_{\varepsilon(m)}^{-1}\}_m$ are the inverse branches of the map $\mathbf{F}_{\varepsilon}$. From Lemma B.2 we know that, $d(z'_i, \bar{z}'_i) \leq \mathcal{O}(1)d(z_i, \bar{z}_i)$, $d(z'_k, \bar{z}'_k) \leq \mathcal{O}(\Xi_l)d(z_i, \bar{z}_i)$ if $k \in [L] \setminus \{i\}$, and $d(z'_k, \bar{z}'_k) \leq \mathcal{O}(\Xi_H)d(z_i, \bar{z}_i)$ if $k \in [L+1, N] \setminus \{i\}$. Plugging this in to equation (4.32) we obtain

$$\begin{aligned}
\frac{|\mathbf{D}_{\mathbf{z}}\mathbf{F}_{\varepsilon}|}{|\mathbf{D}_{\bar{\mathbf{z}}} \mathbf{F}_{\varepsilon}|} &\leq \exp \{ [\mathcal{O}(1) + L\mathcal{O}(\Xi_l) + M\mathcal{O}(\Xi_H)]d(z_i, \bar{z}_i) \} \\
&\leq \exp \left\{ [\mathcal{O}(1) + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2}) + \mathcal{O}(\Delta^{-3/2}M^3) + \right. \\
&\quad \left. + \mathcal{O}(\Delta^{-2}LM\delta^{3/2})]d(z_i, \bar{z}_i) \right\} \\
&\leq \exp \left\{ [\mathcal{O}(1) + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2})]d(z_i, \bar{z}_i) \right\}.
\end{aligned}$$

Taking $\varphi \in \mathcal{C}_a$, after analogous manipulations

$$\frac{\varphi(\mathbf{z}')}{\varphi(\bar{\mathbf{z}}')} \leq \exp \left\{ \sum_{n=1}^N ad(z'_n, \bar{z}'_n) \right\}$$

$$\leq \exp\left\{\left[(\mathbf{DF}_\varepsilon)_{ii}^{-1} + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2})\right] ad(z_i, \bar{z}_i)\right\}. \quad (4.42)$$

Putting together (4.7) and (4.42)

$$\begin{aligned} \frac{(P\varphi)(z_1, \dots, z_i, \dots, z_N)}{(P\varphi)(z_1, \dots, \bar{z}_i, \dots, z_N)} &\leq \\ &\leq \exp\left\{\left[(\mathbf{DF}_\varepsilon^{-1})_{ii} + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2})\right] ad(z_i, \bar{z}_i) + \right. \\ &\quad \left. + \left[\mathcal{O}(1) + \mathcal{O}(L\delta\Delta^{-3/2}M) + \mathcal{O}(\Delta^{-3}L^2\delta^{5/2})\right] d(z_i, \bar{z}_i)\right\}. \end{aligned}$$

□

Now that we have the above estimates the proof of Theorem 4.4 is the same as the proof of Theorem 4.2.

4.2.4 One-Dimensional Disintegration Along the Coordinates of the Low Degree Nodes

We follow once again the decoupling strategy of Section 4.1.4, in particular, we apply it to the decoupling of low degree nodes. For $\ell \in [L]$ we define $\mathbf{F}_{\varepsilon, \ell}$ the map $\mathbf{F}_\varepsilon$ where the ℓ -th node has been decoupled (see Eq. (4.11)), $P_{\varepsilon, \ell} = f_* \hat{P}_{\varepsilon, \ell} = \hat{P}_{\varepsilon, \ell} f_*$. Most of what we argued in the case of a sparsely coupled network holds also in this case, but we will have to compute again some bounds to account for the heterogeneous structure. We are just going to highlight the changes one needs to consider.

Lemma 4.5. *There is $\eta > 0$ depending on f , h and α only, such that, if $L^3\Delta^{-4}\delta^3M^2 < \eta$, then*

$$1. \quad \frac{|\mathbf{D}_z \mathbf{F}_\varepsilon|}{|\mathbf{D}_z \mathbf{F}_{\varepsilon, \ell}|} \leq \exp\{\mathcal{O}(\Delta^{-1}\delta)\}, \quad \forall z \in \mathbb{T}^N; \quad (4.43)$$

2. if $\varphi \in \mathcal{C}_a$,

$$\frac{P_\varepsilon \varphi}{P_{\varepsilon, \ell} \varphi} \leq \exp\left\{\mathcal{O}(\Delta^{-1}\delta) + \mathcal{O}(\Xi_l \Delta^{-1}L\delta) + \mathcal{O}(\Xi_H \Delta^{-1}M\delta)\right\}. \quad (4.44)$$

3. there is $C_\# \in \mathbb{R}^+$ such that if $\varphi \in \mathcal{C}_a$,

$$\left|\partial_{z_i} \frac{P\varphi}{P_i\varphi}\right| \leq C_\#. \quad (4.45)$$

Proof. The proof of items (1) and (3) can be reproduced verbatim from the proof of Lemma 4.1.

To prove (2), once again define functions, for any fixed $\ell \in [L]$, $\mathbf{F}_\varepsilon =: \tilde{\mathbf{F}}_0, \tilde{\mathbf{F}}_1, \dots, \tilde{\mathbf{F}}_{d_{O,\ell}+1} := \mathbf{F}_{\varepsilon,k}$ where $\tilde{\mathbf{F}}_j$ and $\tilde{\mathbf{F}}_{i,j+1}$ differ only for the coordinate k_j and $d_{C^1}(\tilde{\mathbf{F}}_j^{k_j}, \tilde{\mathbf{F}}_{j+1}^{k_j}) < \mathcal{O}(\Delta^{-1})$. From Lemma 4.6 below one deduces that

$$d_{C^0}((\mathbf{F}_j^{-1})^{(n)}, (\mathbf{F}_{j+1}^{-1})^{(n)}) \leq \mathcal{O}(\Delta^{-1}) \begin{cases} \mathcal{O}(1) & n = k \\ \mathcal{O}(\Xi_l) & n \in [L] \setminus \{k\} \\ \mathcal{O}(\Xi_H) & n \in [L+1, N] \setminus \{k\} \end{cases}$$

We can use the above bounds to conclude that

$$\begin{aligned} \frac{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \mathbf{F}_\varepsilon^{-1}}{|\mathbf{D}\mathbf{F}_{\varepsilon,\ell}| \circ \mathbf{F}_{\varepsilon,\ell}^{-1}} &= \prod_{j=1}^{d_{O,\ell}} \frac{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \tilde{\mathbf{F}}_{j-1}^{-1}}{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \mathbf{F}_j^{-1}} \\ &\leq \prod_{j=1}^{d_{O,\ell}} \exp \{ \mathcal{O}(\Delta^{-1}) + \mathcal{O}(L\Delta^{-1}\Xi_l) + \mathcal{O}(M\Delta^{-1}\Xi_H) \} \\ &\leq \exp \{ \mathcal{O}(\Delta^{-1}\delta) + \mathcal{O}(\Xi_l L\Delta^{-1}\delta) + \mathcal{O}(\Xi_H \Delta^{-1}M\delta) \} \end{aligned} \quad (4.46)$$

where we used Eq. (4.32) to obtain (4.46), and the fact that $d_{O,\ell} \leq \delta$ for the final estimate. The equality

$$\frac{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \mathbf{F}_\varepsilon^{-1}}{|\mathbf{D}\mathbf{F}_{\varepsilon,\ell}| \circ \mathbf{F}_{\varepsilon,\ell}^{-1}} = \frac{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \mathbf{F}_\varepsilon^{-1}}{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \mathbf{F}_{\varepsilon,\ell}^{-1}} \frac{|\mathbf{D}\mathbf{F}_\varepsilon| \circ \mathbf{F}_{\varepsilon,\ell}^{-1}}{|\mathbf{D}\mathbf{F}_{\varepsilon,\ell}| \circ \mathbf{F}_{\varepsilon,\ell}^{-1}}$$

point (1), and

$$\frac{\varphi \circ \mathbf{F}_\varepsilon^{-1}}{\varphi \circ \mathbf{F}_{\varepsilon,\ell}^{-1}} \leq \exp \{ a[\mathcal{O}(\Delta^{-1}\delta) + \mathcal{O}(\Xi_l L\delta) + \mathcal{O}(\Xi_H M\delta)] \}$$

end the proof of point (2). $\square$

Lemma 4.6. *Let $\mathbf{G}, \tilde{\mathbf{G}} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ be any two maps such that the matrices $\mathbf{D}_z \mathbf{G}$ and $\mathbf{D}_z \tilde{\mathbf{G}}$ satisfy the assumptions of Lemma B.2 for all $\mathbf{z} \in \mathbb{T}^N$. Suppose that there is $k \in [N]$ such that $G^{(n)} = \tilde{G}^{(n)}$ for all $n \neq k$ and $d_{C^1}(G^{(k)}, \tilde{G}^{(k)}) < \xi$. Then*

$$d_{C^0}((\mathbf{G}^{-1})^{(n)}(\mathbf{x}), (\tilde{\mathbf{G}}^{-1})^{(n)}(\mathbf{x})) \leq \mathcal{O}(\xi) \begin{cases} \mathcal{O}(1) & n = k \\ \mathcal{O}(\Xi_l) & n \in [L] \setminus \{k\} \\ \mathcal{O}(\Xi_H) & n \in [L+1, N] \setminus \{k\} \end{cases}.$$

for all $\mathbf{x} \in \mathbb{T}^N$, and where the inverses are meant to be local inverses.

Proof. Once again it is clear that the preimages of $\mathbf{x}$, that we call $\mathbf{z}$ and $\bar{\mathbf{z}}$, belong to the set $\Gamma = \cap_{\ell \neq k} (G^{(\ell)})^{-1}(x_\ell)$ which is a one-dimensional submanifold of $\mathbb{T}^N$. A unit

tangent vector $\mathbf{v}_z$ to this curve at any point $\mathbf{y} \in \Gamma$ belongs to $\text{Span} \left(\{ \nabla_{\mathbf{y}} G^{(\ell)} \}_{\ell \neq k} \right)^\perp$. These gradients are the rows of the matrix $\mathbf{D}_{\mathbf{y}} \mathbf{G}$ apart from the k -th one. Since this matrix satisfies the same assumptions valid for $\mathbf{D}\mathbf{F}_\varepsilon$ in Lemma B.2, we can upper bound the entries of $\mathbf{v}_z$ as in (4.36) and (4.37). Furthermore

$$d_p(\mathbf{G}(\bar{\mathbf{z}}), \mathbf{G}(\mathbf{z})) = d_p(\mathbf{G}(\bar{\mathbf{z}}), \mathbf{x}) = d_p(\mathbf{G}(\bar{\mathbf{z}}), \tilde{\mathbf{G}}(\bar{\mathbf{z}})) < \xi$$

since we assume that $\mathbf{G}$ is uniformly expanding and

$$d_p(\bar{\mathbf{z}}, \mathbf{z}) < \mathcal{O}(\xi).$$

Since $\|\mathbf{v}\|_p > c_\# > 0$ (for some $c_\#$ depending on f , h and α only) one has that

$$d(\bar{z}_n, z_n) \leq \mathcal{O}(\xi) |v_{z,\ell}| \leq \mathcal{O}(\xi) \begin{cases} \mathcal{O}(1) & n = k \\ \mathcal{O}(\Xi_l) & n \in [L] \setminus \{k\} \\ \mathcal{O}(\Xi_H) & n \in [L+1, N] \setminus \{k\} \end{cases}$$

□

Once again we define norms that account for the regularity of disintegrations along one-dimensional coordinates. However, this time we restrict only to those coordinates corresponding to low degree nodes. With an abuse of notation, we denote the norms using the same symbols as in Definition 4.4, but notice that the first supremum in both definitions runs only through indices in $[L]$.

Definition 4.5. Consider μ a measure on $\mathbb{T}^N$. The norms $|\cdot|_{1D}$ and $\|\cdot\|_{1D}$ are defined as

$$|\mu|_{1D} := \sup_{\ell \in [L]} \sup_{\mathbf{z}_\ell \in \mathbb{T}^{N-1}} \sup_{\substack{\psi \in C^\infty(\mathbb{T}, \mathbb{R}) \\ |\psi|_\infty < 1}} \mu_{\mathbf{z}_\ell}(\psi) \quad (4.47)$$

$$\|\mu\|_{1D} := \sup_{\ell \in [L]} \sup_{\mathbf{z}_\ell \in \mathbb{T}^{N-1}} \sup_{\substack{\psi \in C^\infty(\mathbb{T}, \mathbb{R}) \\ |\psi|_\infty < 1}} \mu_{\mathbf{z}_\ell}(D_x \psi). \quad (4.48)$$

As before, we prove Lakota-Yorke inequality for densities in $\mathcal{C}_a$.

Proposition 4.8. *There exists $A \in \mathbb{R}^+$ and $\lambda \in (0, 1)$ depending on $(f \text{ and } h \text{ only})$, such that if $\mu \ll m_N$ is a measure with $d\mu/dm_N = \varphi \in \mathcal{C}_a$ then*

$$\|\mathbf{F}_* \mu\|_{1D} \leq \lambda \|\mu\|_{1D} + A |\mu|_{1D}.$$

Proof. The proof proceeds in the same way as for that of Proposition 4.5 for sparsely coupled maps *mutatis mutandis*. One needs to replace $\mathbf{F}$ with $\mathbf{F}_\varepsilon$ and $\mathbf{F}_i$ with $\mathbf{F}_{\varepsilon, \ell}$

(only for $\ell \in [L]$). The bound in (4.25) in this setting reads

$$\frac{\mathbf{F}_{\varepsilon, \ell_*} \mu(1_{B_t(\mathbf{z}_\ell)} \circ \pi_{\{i\}^c})}{\mathbf{F}_{\varepsilon_*} \mu(1_{B_t(\mathbf{z}_\ell)} \circ \pi_{\{i\}^c})} \leq \sup_{\mathbf{z}} \frac{|\mathbf{D}_{\mathbf{z}} \mathbf{F}_{\varepsilon, \ell}|}{|\mathbf{D}_{\mathbf{z}} \mathbf{F}_{\varepsilon}|} \leq \exp\{\mathcal{O}(\Delta^{-1}\delta)\}$$

for small $t > 0$, where we used (4.43). Once again, we can use the fact that the ℓ -th node has been uncoupled to deduce

$$(\mathbf{F}_{\varepsilon, \ell_*} \mu)_{\mathbf{z}_\ell}(\psi') \leq \lambda' \|\mu\|_{1D} + A' |\mu|_{1D}$$

for some $\lambda' \in (0, 1)$ depending on f , and any $\psi \in C^\infty(\mathbb{T}, [-1, 1])$ so

$$\limsup_{t \rightarrow 0^+} \frac{\mathbf{F}_{\varepsilon, \ell_*} \mu(\psi' 1_{B_t(\mathbf{z}_\ell)} \circ \pi_{\{i\}^c})}{\mathbf{F}_{\varepsilon_*} \mu(1_{B_t(\mathbf{z}_\ell)} \circ \pi_{\{i\}^c})} \leq \exp\{\mathcal{O}(\Delta^{-1}\delta)\} [\lambda' \|\mu\|_{1D} + A' |\mu|_{1D}].$$

Eq. (4.29) becomes

$$\begin{aligned} \limsup_{t \rightarrow 0^+} \frac{\int_{\mathbb{T}^N} P_\varepsilon \varphi \partial_{z_i} [(1 - \frac{P_{\varepsilon, \ell} \varphi}{P_\varepsilon \varphi}) \psi] 1_{B_t(\mathbf{z}_i)} \circ \pi_{\{i\}^c} d\mu}{\int_{\mathbb{T}^N} P \varphi 1_{B_t(\mathbf{z}_i)} \circ \pi_{\{i\}^c} d\mu} &\leq [\mathcal{O}(\Delta^{-1}\delta) + \mathcal{O}(\Xi_t \Delta^{-1} L \delta) + \\ &\quad + \mathcal{O}(\Xi_H \Delta^{-1} M \delta)] \|\mu\|_{1D} \end{aligned}$$

so there is an $\eta > 0$ such that if

$$\Delta^{-1}\delta + \Xi_t \Delta^{-1} L \delta + \Xi_H \Delta^{-1} M \delta < \eta$$

then the Lasota-Yorke is satisfied for some $\lambda \in (\lambda', 1)$. $\square$

Theorem 4.3 can now be proved in the same way as Theorem 4.1, and one can thus obtain similar results regarding the measure of sets supporting large spatial deviations.

Conclusions and Further Developments

5.1 Conclusions

Heterogeneously Coupled Maps (HCM) are ubiquitous in applications. Because of the heterogeneous structure and lack of symmetries in the graph, most 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. Using an ergodic description, the dynamics of hubs can be predicted from knowledge of the local dynamics and coupling function. This makes it possible to obtain quantitative theoretical predictions of the behaviour of complex systems. Thereby, we establish that existence of a range of dynamical phenomena quite different from the ones encountered in homogenous networks. This highlights the need of new paradigms when dealing with high-dimensional dynamical systems with a heterogeneous coupling.

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

Synchronisation in random networks - HCM versus Homogeneous. Theorem C shows that synchronisation occurs in random homogeneous networks, but

is rare in HCM. Recent work (for example [GSBC98]) showed that structure influences dynamics. What Theorem B shows is that (probabilistic) homogeneity makes synchronisation possible. In contrast, in presence of heterogeneity the dynamics changes according to connectivity layers.

Importance of Long Transients in High Dimensional Systems. Section 3.2.1 shows how certain behaviour can be sustained by a system only for finite time T that turns out to be exponentially large, in terms of the size of the network, and 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 $N \rightarrow \infty$ ¹. 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.

5.2 Open problems and new research directions

With regard to HCM some problems remain open.

1. 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.*

2. In the proof of Theorems A and C we assume that the reduced dynamics g_j , in Eq. (1.4), of each hub node $j \in \{1, \dots, M\}$ is uniformly hyperbolic.

¹Take f an expanding map and define

$$\mathbf{f} := f \times \dots \times f.$$

Suppose ν is the invariant measure for f absolutely continuous with respect to some reference measure m different from ν . Then the push forward $\mathbf{f}_*^t(m^{\otimes N}) = (f_*^t m)^{\otimes N}$ will converge exponentially fast in some suitable product norm to $\nu^{\otimes N}$ because this is true for each factor separately. However, choosing a large N , the rate can be made arbitrarily slow and in the limit of infinite N , $\mathbf{f}_*^t(m^{\otimes N})$ and $\nu^{\otimes N}$ are singular for all $t \in \mathbb{N}$.

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

In fact, hyperbolicity is a generic condition in dimension one, see [RGvS15, KSvS07] but not in higher dimensions. An answer to this question would be desirable even in the one-dimensional case, but especially when treating heterogeneously coupled multi-dimensional maps.

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 proved in [Tsu01] 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 depend 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 To Chapter 3

A.1 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 suppose that there exists $a_i < b_i$ such that $X_i(\omega) \in [a_i, b_i]$ for all $\omega \in \Omega$, then

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

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

Proof of Proposition 3.18. Hoeffding's inequality can be directly applied to the random variables defined on $(\mathbb{T}^L, \mathcal{B}, m_L)$ by $X_i := \theta_{s_1} \circ \pi_i(\mathbf{x})$ where $\pi_i : \mathbb{T}^L \rightarrow \mathbb{T}$ is the projection on the i -th coordinate ($i \in [L]$). These are in fact independent by construction and bounded since $\{\theta_{s_1}\}_{s_1 \in \mathbb{Z}}$ are trigonometric functions with range in $[-1, 1]$.

Consider the set

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

$\mathcal{B}_\varepsilon^{(s_1, j)}$ defined as in (3.15). Notice that for $s_1 = 0$, $\mathcal{B}_\varepsilon^{(0, j)} = \mathbb{T}^L$. Since $\kappa_j \Delta = d_{I, j}$, we can rewrite $\mathcal{B}_\varepsilon^{(s_1, j)}$ as

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

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

$$m_L(\mathcal{B}_\varepsilon^{(s_1, j)}) \leq 2 \exp \left[- \frac{d_j^2 \varepsilon^2 |s_1|^2}{d_j 2 \kappa_j^2} \right] = 2 \exp \left[- \frac{\Delta \varepsilon^2 |s_1|^2}{2 \kappa_j} \right] \quad (\text{A.1})$$

and this gives

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

since $\kappa_j < 1$, which concludes the proof of the corollary. $\square$

Here follows an expression for $\mathbf{D}_z \mathbf{F}_\varepsilon$. Using (3.12) and (3.13) and writing as before $\mathbf{z} = (\mathbf{x}, \mathbf{y})$, noting that for $k > L$, $z_k = y_{k-L}$ we get

$$[\mathbf{D}_{(\mathbf{x}, \mathbf{y})} \mathbf{F}_\varepsilon]_{k\ell} = \begin{cases} D_{x_k} f + \frac{\alpha}{\Delta} \sum_{n=1}^N A_{kn} h_1(x_k, z_n) & k = \ell \in [L], \\ \frac{\alpha}{\Delta} A_{k\ell} h_2(x_k, z_\ell) & k \neq \ell, k \in [L], \\ \partial_{x_\ell} \xi_{k-L, \varepsilon} & k > L, \ell \in [L] \\ \frac{\alpha}{\Delta} A_{k\ell} h_2(y_{k-L}, y_{\ell-L}) & k \neq \ell \in [L+1, N], \\ D_{y_{k-L}} g_{k-L} + \partial_{y_{k-L}} \xi_{k-L, \varepsilon} & k = \ell \in [L+1, N]. \end{cases} \quad (\text{A.2})$$

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, not to additionally cloud the notation some of the functional dependences.

The following lemma summarises the properties $\xi_{j, \varepsilon}$ satisfies and that will yield good hyperbolic properties for $\mathbf{F}_\varepsilon$.

Lemma A.1. *For every $j \in [M]$ and $\varepsilon > 0$, the functions $\xi_{j, \varepsilon} : \mathbb{T}^N \rightarrow \mathbb{R}$ defined in Eq. (3.14) satisfy*

(i)

$$|\xi_{j, \varepsilon}| \leq C_\#(\varepsilon + \Delta^{-1}M)$$

where $C_\#$ is a constant depending only on σ , h , and α .

(ii)

$$|\partial_{z_n} \xi_{j, \varepsilon}| \leq \begin{cases} \mathcal{O}(\Delta^{-1}) A_{j\ell} & n \in [L] \\ \mathcal{O}(\Delta^{-1}) A_{jn} & n \in [L+1, N] \setminus \{j+L\} \\ C_\# \varepsilon + \mathcal{O}(\Delta^{-1}M) & n = j+L. \end{cases}$$

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

$$|\partial_{z_n} \xi_{j,\varepsilon}(\mathbf{z}) - \partial_{z_n} \xi_{j,\varepsilon}(\bar{\mathbf{z}})| \leq \begin{cases} \mathcal{O}(\Delta^{-1}) A_{jn} d_\infty(z, \bar{z}) & n \in [L] \\ \mathcal{O}(\Delta^{-1}) A_{jn} d_\infty(z, \bar{z}) & n \in [L+1, N] \setminus \{j+L\} \\ \mathcal{O}(1) + \mathcal{O}(\Delta^{-1} M) d_\infty(z, \bar{z}) & n = j+L. \end{cases}$$

Proof. Proof of (i) follows from the following estimates

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

where we used that the sum is absolutely convergent. To prove (ii) notice that for $n \in [L]$

$$\partial_{z_n} \xi_{j,\varepsilon}(\mathbf{z}) = \partial_{x_n} \xi_{j,\varepsilon}(\mathbf{z}) = \alpha \sum_{s \in \mathbb{Z}^2} c_s D_{(\cdot) \zeta_{\varepsilon|s_1|}} \frac{A_{jn}}{\Delta} (D_{x_n} \theta_{s_1}) \theta_{s_2} \quad (\text{A.3})$$

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 $n = j+L$

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

Again the decay of the Fourier coefficients yields the desired bound. For $n \in [L+1, N]$ and different from $j+L$ it is trivial. Point (iii). Suppose $i \in [L]$.

$$\partial_{x_i} \xi_{j,\varepsilon}(\mathbf{z}) = \alpha \sum_{s \in \mathbb{Z}^2} c_s D_{\zeta_{\varepsilon|s_1|}}(\cdot) \frac{\alpha}{\Delta} A_{ji}^h (D_{x_i} \theta_{s_1}) \theta_{s_2}(y_j)$$

then

$$\begin{aligned} |\partial_{x_i} \xi_{j,\varepsilon}(\mathbf{z}) - \partial_{x_i} \xi_{j,\varepsilon}(\bar{\mathbf{z}})| &\leq \mathcal{O}(\Delta^{-1}) \sum_{s \in \mathbb{Z}^2} c_s |s_1| A_{ji}^h \mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L A_{j\ell}^h |\theta_{s_1}(x_\ell) - \theta_{s_1}(\bar{x}_\ell)| + \\ &\quad + \mathcal{O}(\Delta^{-1}) \sum_{s \in \mathbb{Z}^2} c_s |s_1| A_{ji}^h |\theta'_{s_1}(x_i) - \theta'_{s_1}(\bar{x}_i)| + \\ &\quad + \mathcal{O}(\Delta^{-1}) \sum_{s \in \mathbb{Z}^2} c_s |s_1| A_{ji}^h |\theta_{s_2}(y_j) - \theta_{s_2}(\bar{y}_j)| \end{aligned}$$

$$\leq A_{ji}^h \mathcal{O}(\Delta^{-1}) \left[\mathcal{O}(\Delta^{-1}) \sum_{\ell} A_{j\ell}^h d(x_{\ell}, \bar{x}_{\ell}) + d(x_i, \bar{x}_i) + d(y_j, \bar{y}_j) \right].$$

If $m \in [M]$, $m \neq j$

$$\partial_{y_m} \xi_{j,\varepsilon} = \frac{\alpha}{\Delta} h_2(y_j, y_m)$$

so

$$|\partial_{y_m} \xi_{j,\varepsilon}(\mathbf{z}) - \partial_{y_m} \xi_{j,\varepsilon}(\bar{\mathbf{z}})| \leq \mathcal{O}(\Delta^{-1}) [d(y_j, \bar{y}_j) + d(y_m, \bar{y}_m)].$$

and finally

$$\partial_{y_j} \xi_{j,\varepsilon} = \alpha \sum_{\mathbf{s} \in \mathbb{Z}^2} c_{\mathbf{s}} \zeta_{\varepsilon|\mathbf{s}_1|}(\cdot) (D\theta_{s_2}) + \sum_{n=1}^M \frac{\alpha}{\Delta} A_{jn}^h h_1(y_j, y_n)$$

which implies

$$\begin{aligned} |\partial_{y_j} \xi_{j,\varepsilon}(\mathbf{z}) - \partial_{y_j} \xi_{j,\varepsilon}(\bar{\mathbf{z}})| &\leq \mathcal{O}(1) \sum_{\mathbf{s}} c_{\mathbf{s}} |\zeta_{\varepsilon|\mathbf{s}_1|}(\cdot) - \zeta_{\varepsilon|\mathbf{s}_1|}(\bar{\cdot})| |s_2| + \\ &\quad + \mathcal{O}(1) \sum_{\mathbf{s}} c_{\mathbf{s}} \zeta_{\varepsilon|\mathbf{s}_1|}(\cdot) |\theta_{s_2}(y_j) - \theta_{s_2}(\bar{y}_j)| + \\ &\quad + \mathcal{O}(\Delta^{-1}) \sum_{m=1}^M A_{jm}^h |h_1(y_j, y_m) - h_1(\bar{y}_j, \bar{y}_m)| \\ &\leq \mathcal{O}(1) \sum_{\mathbf{s}} c_{\mathbf{s}} |s_2| \mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L A_{j\ell}^h [\theta_{s_1}(x_{\ell}) - \theta_{s_1}(\bar{x}_{\ell})] + \\ &\quad + \mathcal{O}(\varepsilon) d(y_j, \bar{y}_j) + \mathcal{O}(\Delta^{-1}) \sum_{m=1}^M A_{jm}^h [d(y_j, \bar{y}_j) + d(y_m, \bar{y}_m)] \\ &\leq \mathcal{O}(\Delta^{-1}) \left[\sum_{\ell} A_{j\ell}^h d(x_{\ell}, \bar{x}_{\ell}) + \sum_m A_{jm}^h [d(y_j, \bar{y}_j) + d(y_m, \bar{y}_m)] \right] \\ &\quad + \mathcal{O}(\varepsilon) d(y_j, \bar{y}_j) \end{aligned}$$

In all the above computations we made use of the fast decay of the Fourier coefficients ensured by the requirement $h \in C^{10}(\mathbb{T}^2, \mathcal{R})$. $\square$

A.2 Estimate on Ratios of Determinants

In the following proposition $\text{Col}^k[\mathbf{M}]$, with $\mathbf{M} \in \mathcal{M}(N, N)$ a square matrix of dimension n , is the k -th column of the matrix $\mathbf{M}$.

Proposition A.1. *Suppose that $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}^+$ is the p norm ($1 \leq p \leq \infty$) on the euclidean space $\mathbb{R}^n$. Take two square matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ of dimension n . Suppose*

there is constant $\lambda \in (0, 1)$ and such that

$$\|\mathbf{B}_i\|_p := \sup_{\substack{\mathbf{v} \in \mathbb{R}^n \\ \|\mathbf{v}\|_p \leq 1}} \frac{\|\mathbf{B}_i \mathbf{v}\|_p}{\|\mathbf{v}\|_p} \leq \lambda \quad \forall i \in \{1, 2\}, \quad (\text{A.4})$$

Then

$$\frac{|\text{Id} + \mathbf{B}_1|}{|\text{Id} + \mathbf{B}_2|} \leq \exp \left\{ \frac{\sum_{k=1}^n \|\text{Col}^k[\mathbf{B}_1 - \mathbf{B}_2]\|_p}{1 + \lambda} \right\}.$$

Proof. Given a matrix $\mathbf{M} \in \mathcal{M}_{N,N}$ it is a standard formula that

$$|\mathbf{M}| = \exp[\text{Tr} \log(\mathbf{M})].$$

$$\frac{|\text{Id} + \mathbf{B}_1|}{|\text{Id} + \mathbf{B}_2|} = \exp \left\{ \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1}}{\ell} \text{Tr}[\mathbf{B}_1^\ell - \mathbf{B}_2^\ell] \right\}.$$

Substituting the expression

$$\mathbf{B}_1^\ell - \mathbf{B}_2^\ell = \sum_{j=0}^{\ell-1} \mathbf{B}_1^j (\mathbf{B}_1 - \mathbf{B}_2) \mathbf{B}_2^{\ell-j-1}$$

we obtain

$$\begin{aligned} \text{Tr}(\mathbf{B}_1^\ell - \mathbf{B}_2^\ell) &= \sum_{i=0}^{\ell-1} \text{Tr}(\mathbf{B}_1^i (\mathbf{B}_1 - \mathbf{B}_2) \mathbf{B}_2^{\ell-i-1}) \\ &= \sum_{j=0}^{\ell-1} \text{Tr}(\mathbf{B}_2^{\ell-j-1} \mathbf{B}_1^j (\mathbf{B}_1 - \mathbf{B}_2)) \\ &\leq \sum_{j=0}^{\ell-1} \sum_{k=1}^n \|\text{Col}^k[\mathbf{B}_2^{\ell-j-1} \mathbf{B}_1^j (\mathbf{B}_1 - \mathbf{B}_2)]\| \end{aligned} \quad (\text{A.5})$$

where we used 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 (A.4) we obtain

$$\text{Tr}(\mathbf{B}_1^\ell - \mathbf{B}_2^\ell) \leq \ell \lambda^{\ell-1} \sum_{k=1}^n \|\text{Col}^k[\mathbf{B}_1 - \mathbf{B}_2]\|_p. \quad (\text{A.6})$$

To conclude

$$\frac{|\text{Id} + \mathbf{B}_1|}{|\text{Id} + \mathbf{B}_2|} \leq \exp \left\{ \sum_{\ell=1}^{\infty} (-1)^{\ell+1} \lambda^{\ell-1} \sum_{k=1}^n \|\text{Col}^k[\mathbf{B}_1 - \mathbf{B}_2]\|_p \right\}$$

$$= \exp \left\{ \frac{\sum_{k=1}^n \|\text{Col}^k[\mathbf{B}_1 - \mathbf{B}_2]\|_p}{1 + \lambda} \right\}.$$

□

A.3 Graph Transform: Some Explicit Estimates

We go through once again 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 compute explicitly bounds on Lipschitz constants for the invariant set of admissible manifolds, and contraction rate of the graph transform.

Consider the torus $\mathbb{T}^n$ with the trivial tangent bundle $\mathbb{T}^n \times \mathbb{R}^n$. Suppose that $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ is a constant norm on the tangent spaces, and that, with an 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 $\Pi_s : \mathbb{R}^n \rightarrow \mathbb{R}^{n_s}$ $\Pi_u : \mathbb{R}^n \rightarrow \mathbb{R}^{n_u}$ 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 call $\pi_s : \mathbb{T}^n \rightarrow \mathbb{T}^{n_s}$ and π_u the projection on the respective coordinates. Take $F : \mathbb{T}^n \rightarrow \mathbb{T}^n$ a C^2 local diffeomorphism. We will also define $F_u := \pi_u \circ F$ and $F_s := \pi_s \circ F$. Suppose that it satisfies the following assumptions. There are constants $\beta_u, \beta_s > 1$ 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:

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

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

•

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

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

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

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

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

that takes σ and maps it to $\Gamma\sigma$ with the only requirement that the graph of $\Gamma\sigma$, $(id, \Gamma\sigma)(B_r^u(F_u(x, y)))$, is contained in $F \circ (id, \sigma)(B_r^u(x))$. An expression for Γ is given by

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

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 A.2. *Consider $\sigma \in C_{u,K}^1(B_r^u(x), B_r^s(y)) \subset C_u^1(B_r^u(x), B_r^s(y))$ characterized as*

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

Then the graph transform 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 $z = \pi_u \circ F \circ (id, \sigma)(x)$ and suppose that $x_1, x_2 \in B_r^u(x)$ are such that $\pi_u \circ F \circ (id, \sigma)(x_i) = z_i$. 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

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

Now

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

and

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

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

$$\text{Lip}(D(\Gamma\sigma)) \leq \frac{\lambda_1}{\mu_1(1 - \beta_u^{-1})} \text{Lip}(D\sigma) + \frac{\mu_2}{\mu_1} \lambda_2^{-1} K_u \lambda_1 \frac{(1 + \beta_u^{-1})}{(1 - \beta_u^{-1})} + K_u \frac{(1 + \beta_u^{-1})^2}{\mu_1(1 - \beta_u^{-1})}$$

and this gives the condition of invariance of the proposition. $\square$

Proposition A.3. *For all $\sigma_1, \sigma_2 \in C_u^1(B_r^u(x), B_r^s(y))$*

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

Then if

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

$\Gamma : C_u^1(B_r^u(x), B_r^s(y)) \rightarrow C_u^1(B_r^u(F_u(x, y)), B_r^s(F_s(x, y)))$ 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 $F_u(x_2, \sigma_2(x_2)) = z$.

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

The following estimates hold

$$\begin{aligned}
\|x_1 - x_2\| &= \|x_1 - (F_u \circ (id, \sigma_2))^{-1} \circ (F_u \circ (id, \sigma_1))(x_1)\| \\
&= \|x_1 - (F_u \circ (id, \sigma_2))^{-1} [F_u \circ (id, \sigma_2)(x_1) + F_u \circ (id, \sigma_1)(x_1) - F_u \circ (id, \sigma_2)(x_1)]\| \\
&\leq \|x_1 - x_1\| + \|D_{\bar{x}}(F_u \circ (id, \sigma_2))^{-1}\| \|F_u \circ (id, \sigma_1)(x_1) - F_u \circ (id, \sigma_2)(x_1)\| \\
&\leq \|D_{\bar{x}} F_u\|^{-1} \lambda_1 \|\sigma_1(x_1) - \sigma_2(x_1)\| \\
&\leq \mu_1^{-1} \lambda_1 \|\sigma_1(x_1) - \sigma_2(x_1)\|
\end{aligned} \tag{A.7}$$

and hence

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

□

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

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

$$d_u(V_1, V_2) := \sup_{\substack{u \in \mathbb{R}^{n_u} \\ \|u\|=1}} \|L_1(u) - L_2(u)\|, \quad (\text{A.8})$$

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

Proposition A.4. *If*

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

then $D_z F$ is a contraction with respect to d_u for all $z \in \mathbb{T}^n$.

Proof. Pick $L_1, L_2 : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_s}$ with $\|L_i\| < \beta_u$. They define linear subspaces $V_i = (\text{Id}, L_i)(\mathbb{R}^{n_u})$ which, as a consequence of the condition on the norm of L_i , are tangent to the unstable cone. V_1 and V_2 are transformed by $D_z F$ into subspaces V'_1 and V'_2 . These subspaces are the graph of linear transformations $L'_1, L'_2 : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_s}$ ($\|L'_i\| \leq \beta_u$). Analogously to the graph transform one can find explicit expression for L'_i in terms of L_i :

$$L'_i = \Pi_s \circ D_z F \circ (\text{Id}, L_i) \circ [\Pi_u \circ D_z F \circ (\text{Id}, L_i)]^{-1}.$$

To prove the proposition we then proceed analogously to the proof of Proposition A.3. Pick $u \in \mathbb{R}^{n_u}$ and suppose that $u_1, u_2 \in \mathbb{R}^{n_u}$ are such that

$$(\text{Id}, L'_1)(u) = D_z F \circ (\text{Id}, L_1)(u_1) \quad (\text{Id}, L'_2)(u) = D_z F \circ (\text{Id}, L_2)(u_2).$$

With the above definitions

$$\begin{aligned} \|L'_1(u) - L'_2(u)\| &= \|\Pi_s \circ D_z F(\text{Id}, L_1)(u_1) - \Pi_s \circ D_z F(\text{Id}, L_1)(u_2)\| \\ &\leq \|\Pi_s \circ D_z F(\text{Id}, L_1)(u_1) - \Pi_s \circ D_z F(\text{Id}, L_2)(u_1)\| \\ &\quad + \|\Pi_s \circ D_z F(\text{Id}, L_2)(u_1 - u_2)\| \end{aligned}$$

$$\leq \lambda_1 \|L_1 - L_2\| \|u_1\| + \beta_u \mu_2 (1 + \beta_u) \|u_1 - u_2\|.$$

$$\begin{aligned} \|u_1 - u_2\| &= \|u_1 - [\Pi_u \circ D_z F \circ (\text{Id}, L_2)]^{-1} \Pi_u \circ D_z F \circ (\text{Id}, L_1)(u_1)\| \\ &= \|[\Pi_u \circ D_z F \circ (\text{Id}, L_2)]^{-1} \Pi_u \circ D_z F \circ (0, L_1 - L_2)(u_1)\| \\ &\leq \|\Pi_u \circ D_z F \circ (\text{Id}, L_2)^{-1}\| \beta_u \lambda_1 \|L_1 - L_2\| \|u_1\| \\ &\leq \frac{\beta_u \lambda_1}{\mu_1 (1 - \beta_u)} \|L_1 - L_2\| \|u_1\| \end{aligned}$$

The two estimates together imply that

$$\|L'_1 - L'_2\| \leq \mu_1^{-1} \left[\lambda_1 + \frac{\beta_u \lambda_1}{\mu_1 (1 - \beta_u)} \right] \|L_1 - L_2\|.$$

□

A.4 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], adding a layer of hubs 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. such that $w_1 \leq w_2 \leq \dots \leq w_N$. We construct a random graph where the expectation of the degrees is close to the one as listed in $\mathbf{w}(N)$ (see Proposition A.6). Let $\rho = 1/(w_1 + \dots + w_N)$. 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

$$w_N w_L \rho \leq 1. \tag{A.9}$$

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

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

We assigned constant probability π of having a connection among the hubs to simplify computations later, but different probability could have been assigned without

changing the final outcome. Notice that the admissibility condition (A.9) 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 A.5. *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*

$$\lim_{N \rightarrow \infty} w_1^{-1} L^{1/p} \beta^{1/q} = 0 \quad (\text{A.10})$$

$$\lim_{N \rightarrow \infty} w_1^{-1/p} M^{1/p} = 0 \quad (\text{A.11})$$

$$\lim_{N \rightarrow \infty} w_1^{-2} \beta L^{1+2/p} = 0 \quad (\text{A.12})$$

$$\lim_{N \rightarrow \infty} w_1^{-1} M L^{1/p} = 0 \quad (\text{A.13})$$

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, for $N \rightarrow \infty$.

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

Proposition A.6. *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 $C \in \mathbb{R}^+$*

$$\mathbb{P}(|d_k - \mathbb{E}[d_k]| > C) \leq \exp\left\{-\frac{NC^2}{2}\right\},$$

where

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

Proof. Suppose $1 \leq k \leq L$:

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

From Hoeffding inequality we know that

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

Suppose $k > L$.

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

Again by Hoeffding

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

□

Proof of Proposition A.5. For every $N \in \mathbb{N}$ consider the graphs in $\mathcal{G}(N)$

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

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 $g(N) \rightarrow \infty$ at any speed then

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

Taken any graph $G \in Q_N$, 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 (A.10)-(A.13), in the limit for $N \rightarrow \infty$

$$\begin{aligned} \frac{M^{1/p}}{\Delta^{1/p}} &\leq \frac{M^{1/p}}{w_1^{1/p}} \frac{1}{\left[1 - \frac{C(N)}{w_1}\right]^{1/p}} \rightarrow 0 \\ \frac{N^{1/p} \delta^{1/q}}{\Delta} &\leq \frac{N^{1/p} [w_L + C(N)]^{1/q}}{w_1 \left[1 - \frac{C(N)}{w_1}\right]} \leq \frac{\left[\frac{L^{q/p} w_L}{w_1^q} + \frac{L^{q/p} C(N)}{w_1^q}\right]^{1/q}}{\left[1 - \frac{C(N)}{w_1}\right]} \rightarrow 0 \\ \frac{ML^{1/p}}{\Delta} &\leq \frac{ML^{1/p}}{w_1} \frac{1}{\left[1 - \frac{C(N)}{w_1(n)}\right]} \rightarrow 0 \end{aligned}$$

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

which proves the proposition. $\square$

Appendix To Chapter 4

B.1 Estimates on Entries of $(\mathbf{D}_\mathbf{x}\mathbf{F}^{-1})_{ij}$

One can obtain the entries of $\mathbf{D}_\mathbf{x}\mathbf{F}^{-1}$ with the analytic expressions that involve determinant and corresponding cofactors of $\mathbf{D}_\mathbf{x}\mathbf{F}$. From this formula, one can try to retrieve the functional dependence and the desired regularity properties. Recall that given an invertible matrix $\mathbf{M}$, $(\mathbf{M}^{-1})_{ik}$ is given by

$$(\mathbf{M}^{-1})_{ik} := \frac{|\widehat{\mathbf{M}}_{\mathbf{k}\mathbf{i}}|}{|\mathbf{M}|} \quad (\text{B.1})$$

where $\widehat{\mathbf{M}}_{\mathbf{k}\mathbf{i}}$ is the cofactor obtained eliminating the k -th row and the i -th column from the matrix $\mathbf{M}$. We now give a few lemmas to estimate the ratio of the determinants $|\widehat{\mathbf{M}}_{\mathbf{i}\mathbf{i}}|$ and $|\widehat{\mathbf{M}}_{\mathbf{i}\mathbf{k}}|$, $k \neq i$, when the matrix $\mathbf{M}$ has the same form of the matrix $\mathbf{D}_\mathbf{x}\mathbf{F}$ in terms of orders of its entries.

The following lemma estimates the coordinates of a unit vector which is orthogonal to all the rows of a matrix, apart from one. The matrix has the form of the differential of a sparsely coupled network map, namely its diagonal entries are of order one, while the other (few) nonzero entries are of order ε .

Lemma B.1. *Suppose that $\mathbf{M} \in \mathcal{M}(N, N)$ is a matrix whose entries satisfy the following conditions:*

1. $\forall i \in [N]$ there exists $c_i \in \mathbb{R}^+$ such that $|M_{ii}| \geq c_i > c$,
2. there exists $\varepsilon \geq 0$ and $\mathbf{A} \in \mathcal{M}(N, N)$ with entries in $\{0, 1\}$ such that $|M_{ij}| \leq \varepsilon A_{ij}$ for all $i \neq j$, and $A_{jj} = 0$ ($\forall j \in [N]$).

Then, for all $i \in [N]$, if $\mathbf{v} \in \mathbb{R}^N$ is such that

$$\begin{pmatrix} \mathbf{M}_1 \\ \dots \\ \mathbf{M}_{i-1} \\ \mathbf{M}_{i+1} \\ \dots \\ \mathbf{M}_N \end{pmatrix} \mathbf{v} = 0 \quad \text{and} \quad \|\mathbf{v}\|_2 = 1,$$

then for all $k \neq i$ and $n \in \mathbb{N}$

$$|v_k| \leq \frac{\varepsilon}{c_k} A_{ki} + \sum_{j=2}^{n-1} \left[\varepsilon^j \sum_{\ell_1 \neq i} \dots \sum_{\ell_{j-1} \neq i} \frac{A_{k\ell_{j-1}} \dots A_{\ell_2 \ell_1} A_{\ell_1 i}}{c_{\ell_{j-1}} \dots c_{\ell_1} c_k} \right] + \\ + \varepsilon^n \sum_{\ell_1 \neq i} \dots \sum_{\ell_{n-1} \neq i} \frac{A_{k\ell_1} A_{\ell_1 \ell_2} \dots A_{\ell_{n-2} \ell_{n-1}} \sqrt{d_{O, \ell_{n-1}}}}{c_{\ell_{n-1}} \dots c_{\ell_1} c_k} \quad (\text{B.2})$$

where

$$d_{O,j} := \sum_{m=1}^N A_{mj}.$$

Proof. For every $k \neq i$ one has

$$M_{kk} v_k + \sum_{\ell_1 \neq k} M_{k\ell_1} v_{\ell_1} = 0. \quad (\text{B.3})$$

This implies that

$$|v_k| \leq \frac{1}{|M_{kk}|} \sum_{\ell_1=1}^N A_{k\ell_1} \varepsilon |v_{\ell_1}| \\ \leq \frac{\varepsilon}{c_k} \sqrt{d_{O,k}}$$

where we used that $\|\mathbf{v}\|_2 = 1$ and the general fact about norms of $\mathbb{R}^n$:

$$\|\mathbf{w}\|_1 \leq n^{1/q} \|\mathbf{w}\|_p \quad \text{with} \quad \frac{1}{p} + \frac{1}{q} = 1, \quad \forall \mathbf{w} \in \mathbb{R}^n.$$

Now we can iterate inequality (B.4) to obtain, for every $k \neq i$

$$|v_k| \leq \frac{\varepsilon}{c_k} A_{ki} + \frac{\varepsilon}{c_k} \sum_{\ell_1 \neq i} A_{k\ell_1} |v_{\ell_1}| \leq \frac{\varepsilon}{c_k} A_{ki} + \varepsilon^2 \sum_{\ell_1 \neq i} \frac{A_{k\ell_1} \sqrt{d_{O, \ell_1}}}{c_k c_{\ell_1}}.$$

Combining all this inequalities, one obtains inductively that, for all $n \in \mathbb{N}$ and $k \neq i$,

$$|v_k| \leq \frac{\varepsilon}{c_k} A_{ki} + \sum_{j=2}^{n-1} \left[\varepsilon^j \sum_{\ell_1 \neq i} \dots \sum_{\ell_{j-1} \neq i} \frac{A_{k\ell_{j-1}} \dots A_{\ell_2 \ell_1} A_{\ell_1 i}}{c_k c_{\ell_{j-1}} \dots c_{\ell_1}} \right] + \\ + \varepsilon^n \sum_{\ell_1 \neq i} \dots \sum_{\ell_{n-1} \neq i} \frac{A_{k\ell_1} A_{\ell_1 \ell_2} \dots A_{\ell_{n-2} \ell_{n-1}} \sqrt{d_{O, \ell_{n-1}}}}{c_k c_{\ell_{n-1}} \dots c_{\ell_1}}.$$

□

Remark B.1. There is a graph theoretical interpretation of the terms in the estimate

for $|v_k|$ above. In fact, the sums

$$\sum_{\ell_1 \neq i} \dots \sum_{\ell_{j-1} \neq i} A_{k\ell_{j-1}} \dots A_{\ell_2 \ell_1} A_{\ell_1 i}$$

count the paths in G of given length j connecting i to k that do not pass through i .

Lemma B.2. *Suppose that $\mathbf{M}$ is in $\mathcal{M}(N, N)$ and its entries satisfy the following conditions:*

1. *there exists $c > 1$ such that $|M_{ii}| > c$ for all $i \in [N]$*
2. *there exists $\varepsilon \geq 0$ and $\mathbf{A} \in \mathcal{M}(N, N)$, the adjacency matrix of a sparse graph with Δ as in (4.2), such that $|M_{ij}| \leq \varepsilon A_{ij}$ for all $i \neq j$.*

Call G the graph defined by the adjacency matrix A and, fixed $i \in [N]$, define the set of nodes $\{\mathcal{L}_j^{(i)}\}_{j=0}^{\kappa}$ as in Definition 4.1. There is $\eta > 0$ depending on h and f only such that if $\varepsilon \Delta^2 < \eta$ then there exists $K > 0$ such that for all $0 \leq m \leq \kappa + 1$ and $k \in \mathcal{L}_m^{(i)}$

$$|\widehat{\mathbf{M}}_{ii}| \geq K \left[1/2 - \varepsilon^2 \Delta^{3/2} \right] \quad \text{and} \quad |\widehat{\mathbf{M}}_{ji}| \leq K \frac{\varepsilon^m \Delta^{m-1/2}}{c^{m-1}}.$$

Proof. Without loss of generality, relabel the nodes so that, $i \rightarrow N$, the nodes in $\mathcal{L}_1^{(N)}$ are labeled with numbers from $N - 1$ to $N - 1 - |\mathcal{L}_1^{(N)}|$, and so on. Call $n_i := \min \mathcal{L}_i^{(N)}$ (after the relabelling). With an abuse of notation we will still call $\mathbf{M}$ the matrix obtained after the relabelling. The entries of $\mathbf{M}$ satisfy: $|M_{ii}| \geq c$, as before; if $k \in \mathcal{L}_{m_1}^{(N)}$, then $M_{kj} = 0$ for all $j \in \mathcal{L}_{m_2}^{(N)}$ with $m_2 > m_1 + 1$; any other nonzero entry is less than ε . Now we compare the determinants of $\widehat{\mathbf{M}}_{\mathbf{NN}}$ and $\widehat{\mathbf{M}}_{\mathbf{jN}}$ for some j . Notice that these two matrices are the same except for one row. Call $\widetilde{\mathbf{M}} \in \mathcal{M}(N, N - 1)$ the matrix obtained from $\mathbf{M}$ deleting the N -th column, and $\overline{\mathbf{M}} \in \mathcal{M}(N - 1, N - 2)$ obtained from $\mathbf{M}$ eliminating the N -th column and row, and the j -th row. Define the linear application $L : \mathbb{R}^{N-1} \rightarrow \mathbb{R}$

$$L(\mathbf{w}) = \det \begin{pmatrix} \overline{\mathbf{M}}_1 \\ \vdots \\ \overline{\mathbf{M}}_{N-2} \\ \mathbf{w}^t \end{pmatrix}$$

in particular $|\widehat{\mathbf{M}}_{\mathbf{NN}}| = |L(\widetilde{\mathbf{M}}_j)|$ and $|\widehat{\mathbf{M}}_{\mathbf{jN}}| = |L(\widetilde{\mathbf{M}}_N)|$. From the geometric characterisation of the determinant as volume of the polyhedron with sides the row of the matrix, we know that $\|L\| = |L(\mathbf{v})|$ where $\mathbf{v} \in \text{Span}(\overline{\mathbf{M}}_1^t, \dots, \overline{\mathbf{M}}_{N-2}^t)^\perp$ and

$\|\mathbf{v}\|_2 = 1$. Furthermore, there exists $K \in \mathbb{R}$ such that $L(\mathbf{w}) = K\mathbf{w}^t\mathbf{v}$, which means that the value of $L(\mathbf{w})$ is entirely determined by the scalar product of $\mathbf{w}$ and $\mathbf{v}$. In particular,

$$\frac{|\widehat{\mathbf{M}}_{\mathbf{jN}}|}{|\widehat{\mathbf{M}}_{\mathbf{NN}}|} = \left| \frac{\widetilde{\mathbf{M}}_{N\mathbf{v}}}{\widetilde{\mathbf{M}}_{\mathbf{jv}}} \right|.$$

From Lemma B.1 we obtain estimates for the entries of $\mathbf{v}$. Suppose that $j \in \mathcal{L}_n^{(k)}$ $n \geq 1$ where the layer now is with respect to the node k . Then from Eq. (B.2)

$$\begin{aligned} |v_k| &\leq \frac{\varepsilon}{c_k} A_{kj} + \sum_{j=2}^{n-1} \left[\varepsilon^j \sum_{\ell_1 \neq j} \dots \sum_{\ell_{j-1} \neq j} \frac{A_{k\ell_{j-1}} \dots A_{\ell_2 \ell_1} A_{\ell_1 j}}{c_k c_{\ell_{j-1}} \dots c_{\ell_1}} \right] + \\ &\quad + \varepsilon^n \sum_{\ell_1 \neq j} \dots \sum_{\ell_{n-1} \neq j} \frac{A_{k\ell_1} A_{\ell_1 \ell_2} \dots A_{\ell_{n-2} \ell_{n-1}} \sqrt{d_{O, \ell_{n-1}}}}{c_k c_{\ell_{n-1}} \dots c_{\ell_1}} \\ &= \varepsilon^n \sum_{\ell_1 \neq j} \dots \sum_{\ell_{n-1} \neq j} \frac{A_{k\ell_1} A_{\ell_1 \ell_2} \dots A_{\ell_{n-2} \ell_{n-1}} \sqrt{d_{O, \ell_{n-1}}}}{c_k c_{\ell_{n-1}} \dots c_{\ell_1}} \end{aligned} \quad (\text{B.5})$$

$$\leq \left(\frac{\varepsilon}{c}\right)^n \Delta^{n-1/2} \quad (\text{B.6})$$

where (B.5) follows from the fact that $j \in \mathcal{L}_n^{(k)}$ and so all the other terms in the other sums equal zero¹ and (B.6) follows from the condition that Δ is the maximum degree. We can also lower bound $|v_j|$, in fact

$$|v_j| = \sqrt{1 - \sum_{k \neq j} |v_k|^2} \geq \sqrt{1 - \Delta \frac{\varepsilon^2 \Delta}{c^2} - \Delta^2 \frac{\varepsilon^4 \Delta^3}{c^4} - \Delta^3 \frac{\varepsilon^6 \Delta^5}{c^6} - \dots}$$

In the above, we used the fact that imposing only the conditions $d_{O,\ell}, d_{I,\ell} \leq \Delta$ for all $\ell \in [N]$, one can upper bound the number of nodes in $\mathcal{L}_m$ with Δ^m .²

$$\begin{aligned} |v_j| &\geq 1 - \mathcal{O}(\varepsilon^2 \Delta^2) \sum_{\ell=0}^m \Delta^\ell \frac{\varepsilon^{2\ell} \Delta^{2\ell}}{c^{2\ell}} \\ &= 1 - \mathcal{O}(\varepsilon^2 \Delta^2) \frac{\varepsilon^{2m+2} \Delta^{3m+3}}{c^{2m+2}} \\ &= 1 - \mathcal{O}\left(\frac{\varepsilon^{2m+4} \Delta^{3m+5}}{c^{2m+2}}\right) \end{aligned}$$

¹Recall Remark B.1

²Once again, more information on the network structure can yield better estimates.

where m is the minimum integer such that $\sum_{\ell=1}^m \Delta^\ell \geq N$, so

$$\frac{\Delta^{m+1} - 1}{\Delta - 1} \geq N \Rightarrow m \leq c_{\#} \frac{\log N}{\log \Delta}$$

and

$$|v_j| \geq 1 - \mathcal{O}\left(\frac{\varepsilon^{2m+4} \Delta^{2m+4}}{c^{2m+2}}\right) \Delta^{m+1}$$

which means that there is $\eta' > 0$ such that if

$$\left(\frac{\varepsilon \Delta}{c}\right)^{2m+4} N < \eta' \quad (\text{B.7})$$

then ³ $|v_j| > 1/2$. Eq. (B.7) implies

$$\begin{aligned} (2m+4) \log \left(\frac{\varepsilon \Delta}{c}\right) &< -\log(\eta' N) \\ \left(2 \frac{\log N}{\log \Delta} + 4\right) \log \left(\frac{\varepsilon \Delta}{c}\right) &< -\log N - \log \eta' \\ \log \left(\frac{\varepsilon \Delta}{c}\right) &< \log \Delta \frac{-\log N - \log \eta'}{2 \log N + 4 \log \Delta} \end{aligned}$$

so there is $\eta > 0$ which does not depend on ε , Δ or N , such that, if

$$\frac{\varepsilon \Delta^2}{c} < \eta \quad (\text{B.8})$$

then Eq. (B.7) holds.

We can now estimate the absolute values of the determinants supposing that Eq.(B.8) holds. For $|\widehat{\mathbf{M}}_{\mathbf{NN}}|$ we obtain

$$\begin{aligned} K|\widetilde{\mathbf{M}}_j \mathbf{v}| &= K \left| \sum_{\ell=1}^{N-1} \widetilde{M}_{j\ell} v_\ell \right| \geq K |v_j| c - K \left| \sum_{\ell \neq j} \widetilde{M}_{j\ell} v_\ell \right| \\ &\geq K \left[1/2 - \sum_{\ell=1}^{N-1} \varepsilon A_{j\ell} |v_\ell| \right] \\ &\geq K \left[1/2 - \varepsilon^2 \Delta^{3/2} \right] \end{aligned}$$

³The number 1/2 is arbitrarily chosen as $\mathcal{O}(1)$ term. Any other choice give the same condition on ε and Δ .

and analogously for $|\widehat{\mathbf{M}}_{\mathbf{jN}}|$

$$\begin{aligned} K|\widetilde{\mathbf{M}}_N \mathbf{v}| &= K \left| \sum_{\ell=1}^{N-1} \widetilde{M}_{N\ell} v_\ell \right| = K \left| \sum_{\ell=n_1}^N \widetilde{M}_{N\ell} v_\ell \right| \leq K\varepsilon \sum_{\ell=n_1}^N A_{N\ell} |v_\ell| \\ &\leq K\varepsilon \sum_{\ell=n_1}^{N-1} A_{\ell j} \left(\frac{\varepsilon}{c} \right)^{n-1} \Delta^{n-3/2} \\ &\leq K \frac{\varepsilon^n \Delta^{n-1/2}}{c^{n-1}}. \end{aligned}$$

□

B.2 Lipschitz Constants of the Entries of $\mathbf{D}_z \mathbf{F}_\varepsilon - \mathbf{D}_{\bar{z}} \mathbf{F}_\varepsilon$

In the following lemma we prove the bounds in Eq. (B.9).

Lemma B.3. *For every $\mathbf{z}, \bar{\mathbf{z}} \in \mathbb{T}^N$, $|(\mathbf{D}_z \mathbf{F}_\varepsilon)_{ki} - (\mathbf{D}_{\bar{z}} \mathbf{F}_\varepsilon)_{ki}| \leq$*

$$\begin{cases} A_{ki} \mathcal{O}(\Delta^{-1}) [d(z_k, \bar{z}_k) + d(z_i, \bar{z}_i)] & k \in [L], i \neq k \\ \mathcal{O}(1) d(z_k, \bar{z}_k) + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d(z_n, \bar{z}_n) & k \in [L], i = k \\ A_{ki} \mathcal{O}(\Delta^{-1}) [d(z_i, \bar{z}_i) + d(z_k, \bar{z}_k) + \mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L A_{k\ell} d(z_\ell, \bar{z}_\ell)] & k \in [L+1, N], i \in [L] \\ A_{ki} \mathcal{O}(\Delta^{-1}) [d(z_k, \bar{z}_k) + d(z_i, \bar{z}_i)] & k, i \in [L+1, N], i \neq k \\ \mathcal{O}(1) d(z_k, \bar{z}_k) + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d(z_n, \bar{z}_n) & k \in [L+1, N], i = k \end{cases} \quad (\text{B.9})$$

Proof. Suppose that $1 \leq k \leq L$ then if $i \neq k$

$$|(\mathbf{D}_z \mathbf{F}_\varepsilon)_{ki} - (\mathbf{D}_{\bar{z}} \mathbf{F}_\varepsilon)_{ki}| \leq A_{ki} \mathcal{O}(\Delta^{-1}) [d(x_k, \bar{x}_k) + d(z_i, \bar{z}_i)]$$

and if $i = k$

$$\begin{aligned} |(\mathbf{D}_z \mathbf{F}_\varepsilon)_{kk} - (\mathbf{D}_{\bar{z}} \mathbf{F}_\varepsilon)_{kk}| &\leq \mathcal{O}(1) d(x_k, \bar{x}_k) + \frac{\alpha}{\Delta} \left| \sum_{\ell} A_{k\ell}^u (h_1(x_k, x_\ell) - h_1(\bar{x}_k, \bar{x}_\ell)) + \right. \\ &\quad \left. + \sum_m A_{km}^{lh} (h_1(x_k, y_m) - h_1(\bar{x}_k, \bar{y}_m)) \right| \\ &\leq \mathcal{O}(1) d(x_k, \bar{x}_k) + \frac{\alpha}{\Delta} \sum_{\ell} A_{k\ell}^u |h_1(x_k, x_\ell) - h_1(\bar{x}_k, \bar{x}_\ell)| + \\ &\quad + \frac{\alpha}{\Delta} \sum_m A_{km}^{lh} |h_1(x_k, y_m) - h_1(\bar{x}_k, \bar{y}_m)| \\ &\leq \mathcal{O}(1) d(x_k, \bar{x}_k) + \mathcal{O}(\Delta^{-1}) \sum_{\ell} A_{k\ell}^u [d(x_k, \bar{x}_k) + d(x_\ell, \bar{x}_\ell)] + \\ &\quad + \mathcal{O}(\Delta^{-1}) \sum_m A_{km}^{lh} [d(x_k, \bar{x}_k) + d(y_m, \bar{y}_m)] \end{aligned}$$

which means

$$|(\mathbf{D}_{\mathbf{z}}\mathbf{F}_\varepsilon)_{kk} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon)_{kk}| \leq \mathcal{O}(1)d(x_k, \bar{x}_k) + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{kn} d(z_n, \bar{z}_n)$$

Now suppose that $k \in [L+1, M]$ and call $j = k - L$. Using formulae in Lemma A.1, if $i \in [L]$

$$\begin{aligned} |(\mathbf{D}_{\mathbf{z}}\mathbf{F}_\varepsilon)_{ki} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon)_{ki}| &\leq \\ &\leq \|f\|_{C^2} |\partial_{x_i} \xi_{j,\varepsilon}(z)| d(x_i, \bar{x}_i) + \mathcal{O}(1) |\partial_{x_i} \xi_{j,\varepsilon}(z) - \partial_{x_i} \xi_{j,\varepsilon}(\bar{z})| \\ &\leq A_{ji}^h \mathcal{O}(\Delta^{-1}) \left[\mathcal{O}(\Delta^{-1}) \sum_{\ell=1}^L A_{j\ell}^h d(x_\ell, \bar{x}_\ell) + d(x_i, \bar{x}_i) + d(y_j, \bar{y}_j) \right] \end{aligned}$$

if $i \in [L+1, N]$, but $i \neq k$,

$$|(\mathbf{D}_{\mathbf{z}}\mathbf{F}_\varepsilon)_{ki} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon)_{ki}| \leq A_{ki} \mathcal{O}(\Delta^{-1}) [d(y_j, \bar{y}_j) + d(z_i, \bar{z}_i)]$$

and if $i = k$

$$\begin{aligned} |(\mathbf{D}_{\mathbf{z}}\mathbf{F}_\varepsilon)_{kk} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon)_{kk}| &\leq \mathcal{O}(1)d(y_j, \bar{y}_j) + \mathcal{O}(1)d(y_j, \bar{y}_j) |\partial_{y_j} \xi_{j,\varepsilon}| + \\ &\quad + \mathcal{O}(1) |\partial_{y_j} \xi_{j,\varepsilon}(z) - \partial_{y_j} \xi_{j,\varepsilon}(\bar{z})| \\ &\leq \mathcal{O}(1)d(y_j, \bar{y}_j) + [\mathcal{O}(\varepsilon) + \mathcal{O}(M\Delta^{-1})]d(y_j, \bar{y}_j) + \\ &\quad + \mathcal{O}(\Delta^{-1}) \left[\sum_{\ell} A_{j\ell}^h d(x_\ell, \bar{x}_\ell) + \sum_m A_{jm}^h [d(y_j, \bar{y}_j) + d(y_m, \bar{y}_m)] \right] + \\ &\quad + \mathcal{O}(\varepsilon)d(y_j, \bar{y}_j) \\ &\leq \mathcal{O}(1)d(y_j, \bar{y}_j) + [\mathcal{O}(\varepsilon) + \mathcal{O}(M\Delta^{-1})]d(y_j, \bar{y}_j) + \\ &\quad + \mathcal{O}(\Delta^{-1}) \left[\sum_{\ell} A_{j\ell}^h d(x_\ell, \bar{x}_\ell) + \sum_m A_{jm}^h d(y_m, \bar{y}_m) \right]. \end{aligned}$$

and so

$$|(\mathbf{D}_{\mathbf{z}}\mathbf{F}_\varepsilon)_{kk} - (\mathbf{D}_{\bar{\mathbf{z}}}\mathbf{F}_\varepsilon)_{kk}| \leq \mathcal{O}(1)d(y_j, \bar{y}_j) + \mathcal{O}(\Delta^{-1}) \sum_{n=1}^N A_{jn}^h d(z_n, \bar{z}_n).$$

□

B.3 Distances of Preimages of Nearby Functions

Suppose that $\mathbf{G}, \overline{\mathbf{G}} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ are local diffeomorphisms and network maps on a graph $([N], \mathcal{E})$ such that $G^{(\ell)}(\mathbf{x}) = \overline{G}^{(\ell)}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{T}^N$ and $j \neq i$, and $d_{C^1}(G^{(j)}(\mathbf{x}), \overline{G}^{(j)}(\mathbf{x})) < \varepsilon$. So basically $\mathbf{G}$ and $\overline{\mathbf{G}}$ act in the same way on $N - 1$ of the coordinates, and on their first coordinate they are ε -close in the C^1 topology. Now let $\mathbf{x} \in \mathbb{T}^N$ and suppose that $U \subset \mathbb{T}^N$ is an open set on which both $\mathbf{G}$ and $\overline{\mathbf{G}}$ are invertible, and suppose that $\mathbf{x} \in \mathbf{G}(U) \cap \overline{\mathbf{G}}(U)$ has preimages $\mathbf{y}$ and $\overline{\mathbf{y}}$ under $\mathbf{G}|_U$ and $\overline{\mathbf{G}}|_U$ respectively.

Lemma B.4. *Let $\mathbf{G}, \overline{\mathbf{G}} : \mathbb{T}^N \rightarrow \mathbb{T}^N$ be two maps such that the matrices $\mathbf{D}_{\mathbf{x}}\mathbf{G}$ and $\mathbf{D}_{\mathbf{x}}\overline{\mathbf{G}}$ satisfy the assumptions of Lemma B.1 for all $\mathbf{x} \in \mathbb{T}^N$. Suppose that there is $k \in [N]$ such that $G^{(\ell)} = \overline{G}^{(\ell)}$ for all $\ell \neq k$ and $d_{C^1}(G^{(k)}, \overline{G}^{(k)}) < \xi$. Then*

$$d_{C^0}((G^{-1})^{(\ell)}, (\overline{G}^{-1})^{(\ell)}) \leq \mathcal{O}(\varepsilon) \frac{\varepsilon^j \Delta^{j-1/2}}{c^j}, \quad \forall \ell \in \mathcal{L}_j, \quad j \geq 1$$

where $\{\mathcal{L}_j\}_j$ are the layers associated to the node k as in Definition 4.1.

Proof. We know that $\mathbf{G}$ and $\overline{\mathbf{G}}$ differ for the image of just one of the coordinates. Suppose that this is $k \in [N]$, then $d_{C^1}(G^{(k)}, \overline{G}^{(k)}) < \xi$. Take $\mathbf{x} \in \mathbf{G}(U) \cap \overline{\mathbf{G}}(U)$ and let $\mathbf{y} = \mathbf{G}^{-1}(\mathbf{x})$, $\overline{\mathbf{y}} = \overline{\mathbf{G}}^{-1}(\mathbf{x})$. Recalling that $G^{(\ell)} = \overline{G}^{(\ell)}$ for $\ell \neq k$, it is clear that $\mathbf{y}, \overline{\mathbf{y}} \in \Gamma = \cap_{\ell \neq k} (G^{(\ell)})^{-1}(x_\ell)$ which is a one-dimensional submanifolds in $\mathbb{T}^N$. A unit tangent vector $\mathbf{v}_{\mathbf{z}}$ to this curve at $\mathbf{z} \in \Gamma$ belongs to $\text{Span} \left(\{\nabla_{\mathbf{z}} G^{(\ell)}\}_{\ell \neq k} \right)^\perp$. The gradients whose span gives the orthogonal to $\mathbf{v}_{\mathbf{z}}$, are the rows of the matrix $\mathbf{D}_{\mathbf{z}}\mathbf{G}$ apart from the k -th one. Since this matrix satisfies the assumptions of Lemma B.1, this allows to estimate the entries of $\mathbf{v}_{\mathbf{z}}$ in the following way:

$$|v_{z,\ell}| \leq \frac{\varepsilon^j \Delta^{j-1/2}}{c^j}, \quad \forall \ell \in \mathcal{L}_j^{(k)}, \quad j \geq 1.$$

Furthermore

$$d(\mathbf{G}(\overline{\mathbf{y}}), \mathbf{G}(\mathbf{y})) = d(\mathbf{G}(\overline{\mathbf{y}}), \mathbf{x}) = d(\mathbf{G}(\overline{\mathbf{y}}), \overline{\mathbf{G}}(\overline{\mathbf{y}})) < \xi.$$

Since $\mathbf{G}$ is a local diffeomorphism on a compact manifold,

$$d(\overline{\mathbf{y}}, \mathbf{y}) < \mathcal{O}(\xi).$$

Since by the mean value theorem $\mathbf{y}$ and $\overline{\mathbf{y}}$ lay in the direction of $\mathbf{v}_{\mathbf{z}}$ for some $\mathbf{z} \in \Gamma$,

one has that

$$d(\bar{y}_\ell, y_\ell) \leq \mathcal{O}(\xi)|v_{z,\ell}| \leq \mathcal{O}(\xi) \frac{\varepsilon^j \Delta^{j-1/2}}{c^j}, \quad \forall \ell \in \mathcal{L}_j^{(k)}, \quad j \geq 1.$$

□

B.4 Derivatives Manipulations

Suppose that $\psi \in C^1(\mathbb{T}^N, \mathbb{R})$, and $G \in C^1(\mathbb{T}^N, \mathbb{T}^N)$ then:

$$\begin{aligned} \partial_i \psi \circ G &= (\nabla(\psi \circ G) DG^{-1})_i \\ &= \sum_j \partial_j(\psi \circ G)(DG^{-1})_{ij} \\ &= \sum_j \partial_j(\psi \circ G(DG^{-1})_{ij}) - \psi \circ G \partial_j(DG^{-1})_{ij} \end{aligned} \quad (\text{B.10})$$

Transfer Operator

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

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

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 *nonsingular* if F_*m is absolutely continuous with respect to m and we write it $F_*m \ll m$. If F is nonsingular, given a measure $\mu \ll m$ then also $F_*\mu \ll m$. This means that one can define an operator

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

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

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

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

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

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

Disintegration of Measures

In this appendix we revise definitions and results on measure disintegration. What follows can be found in [CP97] and [Sim12].

Definition D.1 (Measure Disintegration). Suppose that $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are two measurable spaces, and $T : X \rightarrow Y$ is a measurable function between them. If μ and ν are two σ -finite measures on X and Y respectively, a disintegration of μ with respect to T and ν is given by the collection of measures $\{\mu_y\}_{y \in Y}$ that satisfies the following conditions:

1. μ_y is a σ -finite measure on $(X, \mathcal{A})$ supported on $T^{-1}(y)$ ¹ for ν -almost all $y \in Y$;
2. for all measurable functions $\varphi : X \rightarrow \mathbb{R}^+$, $y \mapsto \mu_y(\varphi)$ is measurable and

$$\mu(\varphi) = \nu(\mu_y(\varphi)).$$

Remark D.1.

- In general a disintegration is not unique.
- If $T_*\mu$ is σ -finite, we sometimes denote a disintegration of μ with respect to T and $\nu := T_*\mu$ just as a disintegration with respect to T without specifying the measure ν .

The following is an existence result for disintegrations of Radon measures on metric spaces and can be deduced by results in [Bill13].

Theorem D.1. *Suppose that X is a metric space, $\mathcal{A}$ its Borel σ -algebra and μ a Radon measure². If $\mathcal{B}$ is countably generated and contains all the singletons $\{y\}$ $y \in Y$, then for any measurable $T : X \rightarrow Y$ there exists a disintegration $\{\mu_y\}_{y \in Y}$ of*

¹ $\mu_y(\{T = y\}^c) = 0$

² This means that $\mu(K) < \infty$ for all K compact, and for any $B \in \mathcal{A}$

$$\mu(B) = \sup_{\substack{K \subset B \\ K \text{ compact}}} \mu(K).$$

μ with respect to T . Furthermore if $\{\mu'_y\}_{y \in Y}$ is any other such disintegration, then $(T_*\mu)(\{y \in Y : \mu_y \neq \mu'_y\}) = 0$.

We now list some useful facts about disintegrations.

Proposition D.1. *Suppose, under the hypotheses of the previous theorem, that $\eta \ll \mu$ has finite density ρ .*

1. *If $\{\eta_y\}_{y \in Y}$ is a disintegration of η with respect to T , then $\eta_y \ll \mu_y$ and $\frac{d\eta_y}{d\mu_y} = \rho$;*
2. *η_y is a probability measure for all $y \in Y$ if and only if $\{\eta_y\}_{y \in Y}$ is a disintegration with respect to T (and $\nu = T_*\eta$).*

Definition D.2. Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measurable spaces, and $T : X \rightarrow Y$ is a measurable function. If the weak limit

$$\mu_y = \lim_{\varepsilon \rightarrow 0^+} \frac{\mu|_{T^{-1}B_\varepsilon(y)}}{\nu(B_\varepsilon(y))} \quad (\text{D.1})$$

exists, then it is called the *topological disintegration* of μ on $T^{-1}(y)$.

Under some assumptions, Eq. (D.1) gives the expression for the disintegrated measure.

Theorem D.2 ([Sim12]). *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two locally compact, locally finite metric measure spaces and suppose furthermore that Y has a Riemannian structure. Let $T : X \rightarrow Y$ be a measurable function and suppose that $T_*\mu \ll \nu$. Then for ν -almost any $y \in Y$, μ_y as in (D.1) exists and $\{\mu_y\}_{y \in Y}$ is a disintegration of the measure μ with respect to T and ν as in Definition D.1.*

Bibliography

- [AADF11] M. Aguiar, P. Ashwin, A. Dias, and M. Field. Dynamics of coupled cell networks: synchrony, heteroclinic cycles and inflation. *Journal of Nonlinear Science*, 21(2):271–323, 2011. (Cited on pages 8 and 26.)
- [AB02] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002. (Cited on page 1.)
- [AS04] Daniel M Abrams and Steven H Strogatz. Chimera states for coupled oscillators. *Physical Review Letters*, 93(17):174102, 2004. (Cited on page 119.)
- [BDEI⁺98] Viviane Baladi, Mirko Degli Esposti, Stefano Isola, Esa Järvenpää, and Antti Kupiainen. The spectrum of weakly coupled map lattices. *Journal de mathématiques pures et appliquées*, 77(6):539–584, 1998. (Cited on pages 25 and 27.)
- [Ber93] Walter Bergweiler. Iteration of meromorphic functions. *Bulletin of the American Mathematical Society (N.S.)*, 29(2):151–188, 1993. (Cited on page 78.)
- [BF17] Christian Bick and Michael Field. Asynchronous networks and event driven dynamics. *Nonlinearity*, 30(2):558, 2017. (Cited on page 18.)
- [BG01] Pietro-Luciano Buono and Martin Golubitsky. Models of central pattern generators for quadruped locomotion i. primary gaits. *Journal of mathematical biology*, 42(4):291–326, 2001. (Cited on page 26.)
- [BGP⁺09] Paolo Bonifazi, Miri Goldin, Michel A Picardo, Isabel Jorquera, A Cattani, Gregory Bianconi, Alfonso Represa, Yehezkel Ben-Ari, and Rosa Cossart. Gabaergic hub neurons orchestrate synchrony in developing hippocampal networks. *Science*, 326(5958):1419–1424, 2009. (Cited on pages 1 and 19.)
- [Bil13] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013. (Cited on page 147.)
- [Bir57] Garrett Birkhoff. Extensions of Jentzsch’s theorem. *Transactions of the American Mathematical Society*, 85:219–227, 1957. (Cited on page 50.)

- [BK95] J Bricmont and A Kupiainen. Coupled analytic maps. *Nonlinearity*, 8(3):379, 1995. (Cited on page 25.)
- [BKZ09] Jean-Baptiste Bardet, Gerhard Keller, and Roland Zweimüller. Stochastically stable globally coupled maps with bistable thermodynamic limit. *Communications in Mathematical Physics*, 292(1):237–270, 2009. (Cited on pages 22 and 24.)
- [Bla11] ML Blank. Self-consistent mappings and systems of interacting particles. In *Doklady Mathematics*, volume 83, pages 49–52. Springer, 2011. (Cited on page 24.)
- [Bol01] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001. (Cited on pages 9, 16 and 130.)
- [BP02] Mauricio Barahona and Louis M Pecora. Synchronization in small-world systems. *Physical Review Letters*, 89(5):054101, 2002. (Cited on pages 26 and 83.)
- [BR01] Viviane Baladi and Hans Henrik Rugh. Floquet spectrum of weakly coupled map lattices. *Communications in Mathematical Physics*, 220(3):561–582, 2001. (Cited on pages 25 and 27.)
- [BRS⁺12] Murilo S Baptista, Hai-Peng Ren, Jochen CM Swarts, Rodrigo Carareto, Henk Nijmeijer, and Celso Grebogi. Collective almost synchronisation in complex networks. *PloS one*, 7(11):e48118, 2012. (Cited on page 1.)
- [BS88] L A Bunimovich and Ya G Sinai. Spacetime chaos in coupled map lattices. *Nonlinearity*, 1(4):491, 1988. (Cited on pages 25 and 27.)
- [Bus73] P. J. Bushell. Hilbert’s metric and positive contraction mappings in a Banach space. *Archive for Rational Mechanics and Analysis*, 52:330–338, 1973. (Cited on page 50.)
- [CF05] Jean-René Chazottes and Bastien Fernandez. *Dynamics of coupled map lattices and of related spatially extended systems*, volume 671. Springer, 2005. (Cited on page 27.)
- [CL06] Fan Chung and Linyuan Lu. *Complex graphs and networks*, volume 107 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by

- the American Mathematical Society, Providence, RI, 2006. (Cited on pages 17 and 130.)
- [CLP16] Vaughn Climenhaga, Stefano Luzzatto, and Yakov Pesin. The geometric approach for constructing Sinai–Ruelle–Bowen measures. *Journal of Statistical Physics*, pages 1–27, 2016. (Cited on page 43.)
- [CP97] Joseph T Chang and David Pollard. Conditioning as disintegration. *Statistica Neerlandica*, 51(3):287–317, 1997. (Cited on page 147.)
- [dMvS93] Welington de Melo and Sebastian van Strien. *One-dimensional dynamics*, volume 25 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1993. (Cited on pages 35, 36, 71 and 76.)
- [DSL16] Jacopo De Simoi and Carlangelo Liverani. Statistical properties of mostly contracting fast-slow partially hyperbolic systems. *Inventiones Mathematicae*, 206(1):147–227, 2016. (Cited on page 27.)
- [DWG⁺05] Luis Garcia Dominguez, Richard A Wennberg, William Gaetz, Douglas Cheyne, O Carter Snead, and Jose Luis Perez Velazquez. Enhanced synchrony in epileptiform activity? local versus distant phase synchronization in generalized seizures. *Journal of neuroscience*, 25(35):8077–8084, 2005. (Cited on page 19.)
- [DZ09] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, 2009. (Cited on page 42.)
- [EM14] Jan R. Engelbrecht and Renato Mirollo. Classification of attractors for systems of identical coupled Kuramoto oscillators. *Chaos*, 24(1):013114, 10, 2014. (Cited on page 26.)
- [Fer14] Bastien Fernandez. Breaking of ergodicity in expanding systems of globally coupled piecewise affine circle maps. *Journal of Statistical Physics*, 154(4):999–1029, 2014. (Cited on pages 21, 22 and 23.)
- [Fie73] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Math. J.*, 23(98):298–305, 1973. (Cited on page 82.)
- [Fie07] Michael J Field. *Dynamics and symmetry*, volume 3. World Scientific, 2007. (Cited on page 26.)

- [FR00] Torsten Fischer and Hans Henrik Rugh. Transfer operators for coupled analytic maps. *Ergodic Theory and Dynamical Systems*, 20(01):109–143, 2000. (Cited on page 25.)
- [GGMA07] Jesús Gómez-Gardenes, Yamir Moreno, and Alex Arenas. Paths to synchronization on complex networks. *Physical Review Letters*, 98(3):034101, 2007. (Cited on page 1.)
- [GM13] Georg A. Gottwald and Ian Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proceedings of The Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, 469(2156):20130201, 16, 2013. (Cited on page 27.)
- [GS03] Martin Golubitsky and Ian Stewart. *The symmetry perspective: from equilibrium to chaos in phase space and physical space*, volume 200. Springer Science & Business Media, 2003. (Cited on page 26.)
- [GS06] Martin Golubitsky and Ian Stewart. Nonlinear dynamics of networks: the groupoid formalism. *Bulletin of the American Mathematical Society (N.S.)*, 43(3):305–364, 2006. (Cited on pages 8 and 26.)
- [GSBC98] Martin Golubitsky, Ian Stewart, Pietro-Luciano Buono, and J. J. Collins. A modular network for legged locomotion. *Physica D: Nonlinear Phenomena*, 115(1-2):56–72, 1998. (Cited on pages 26 and 118.)
- [Izh07] Eugene M Izhikevich. *Dynamical systems in neuroscience*. MIT press, 2007. (Cited on pages 3 and 21.)
- [Jär97] Esa Järvenpää. A SRB-measure for globally coupled circle maps. *Nonlinearity*, 10(6):1435, 1997. (Cited on page 23.)
- [JDCJ⁺13] Premysl Jiruska, Marco De Curtis, John GR Jefferys, Catherine A Schevon, Steven J Schiff, and Kaspar Schindler. Synchronization and desynchronization in epilepsy: controversies and hypotheses. *The Journal of physiology*, 591(4):787–797, 2013. (Cited on page 19.)
- [Jus95] Wolfram Just. Globally coupled maps: phase transitions and synchronization. *Physica D: Nonlinear Phenomena*, 81(4):317–340, 1995. (Cited on pages 22 and 25.)
- [KAB⁺14] Mikko Kivelä, Alex Arenas, Marc Barthélemy, James P Gleeson, Yamir Moreno, and Mason A Porter. Multilayer networks. *Journal of complex networks*, 2(3):203–271, 2014. (Cited on page 18.)

- [Kan84] Kunihiro Kaneko. Period-doubling of kink-antikink patterns, quasiperiodicity in antiferro-like structures and spatial intermittency in coupled logistic lattice towards a prelude of a “field theory of chaos”. *Progress of Theoretical Physics*, 72(3):480–486, 1984. (Cited on page 25.)
- [Kan90] Kunihiro Kaneko. Globally coupled chaos violates the law of large numbers but not the central-limit theorem. *Physical Review Letters*, 65(12):1391, 1990. (Cited on pages 22 and 24.)
- [Kan92] Kunihiro Kaneko. Overview of coupled map lattices. *Chaos*, 2(3):279–282, 1992. (Cited on page 27.)
- [Kan95] Kunihiro Kaneko. Remarks on the mean field dynamics of networks of chaotic elements. *Physica D: Nonlinear Phenomena*, 86(1-2):158–170, 1995. (Cited on page 22.)
- [Kel00] Gerhard Keller. An ergodic theoretic approach to mean field coupled maps. In *Fractal Geometry and Stochastics II*, pages 183–208. Springer, 2000. (Cited on pages 22 and 23.)
- [KH95] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza. (Cited on page 64.)
- [KK92] Gerhard Keller and Martin Künzle. Transfer operators for coupled map lattices. *Ergodic Theory and Dynamical Systems*, 12(02):297–318, 1992. (Cited on page 25.)
- [KL99] Gerhard Keller and Carlangelo Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 28(1):141–152, 1999. (Cited on page 43.)
- [KL04] Gerhard Keller and Carlangelo Liverani. Coupled map lattices without cluster expansion. *Discrete and Continuous Dynamical Systems*, 11:325–336, 2004. (Cited on pages 25 and 27.)
- [KL05] Gerhard Keller and Carlangelo Liverani. 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*, pages 115–151. Springer, 2005. (Cited on pages 25 and 27.)

- [KL06] Gerhard Keller and Carlangelo Liverani. Uniqueness of the SRB measure for piecewise expanding weakly coupled map lattices in any dimension. *Communications in Mathematical Physics*, 262(1):33–50, 2006. (Cited on pages 4, 11, 25, 27, 29, 94 and 97.)
- [KSvS07] O. Kozlovski, W. Shen, and S. van Strien. Density of hyperbolicity in dimension one. *Annals of Mathematics (2)*, 166(1):145–182, 2007. (Cited on pages 27, 77 and 119.)
- [KT11] Kunihiko Kaneko and Ichiro Tsuda. *Complex systems: chaos and beyond: a constructive approach with applications in life sciences*. Springer Science & Business Media, 2011. (Cited on page 25.)
- [Kur84] Y. Kuramoto. *Chemical oscillations, waves, and turbulence*, volume 19 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 1984. (Cited on pages 21 and 26.)
- [KY10] José Koiller and Lai-Sang Young. Coupled map networks. *Nonlinearity*, 23(5):1121, 2010. (Cited on pages 25 and 26.)
- [Lov12] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Society Providence, 2012. (Cited on page 14.)
- [MMAN13] Adilson E Motter, Seth A Myers, Marian Anghel, and Takashi Nishikawa. Spontaneous synchrony in power-grid networks. *Nature Physics*, 9(3):191–197, 2013. (Cited on page 1.)
- [Mn87] Ricardo Mañé. *Ergodic theory and differentiable dynamics*, volume 8 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1987. Translated from the Portuguese by Silvio Levy. (Cited on page 65.)
- [Moh91] Bojan Mohar. Eigenvalues, diameter, and mean distance in graphs. *Graphs and Combinatorics*, 7(1):53–64, 1991. (Cited on page 81.)
- [Moh92] Bojan Mohar. Laplace eigenvalues of graphs—a survey. *Discrete Math.*, 109(1-3):171–183, 1992. Algebraic graph theory (Leibnitz, 1989). (Cited on page 82.)
- [NK98] Naoko Nakagawa and Komatsu. Dominant collective motion in globally coupled tent maps. *Physical Review E*, 57:1570, 1998. (Cited on pages 22 and 23.)

- [NMLH03] Takashi Nishikawa, Adilson E Motter, Ying-Cheng Lai, and Frank C Hoppensteadt. Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize? *Physical review letters*, 91(1):014101, 2003. (Cited on page 1.)
- [NRS16] Eddie Nijholt, Bob Rink, and Jan Sanders. Graph fibrations and symmetries of network dynamics. *Journal of Differential Equations*, 261(9):4861–4896, 2016. (Cited on page 8.)
- [Per10] Tiago Pereira. Hub synchronization in scale-free networks. *Physical Review E*, 82(3):036201, 2010. (Cited on page 1.)
- [PERV14] Tiago Pereira, Jaap Eldering, Martin Rasmussen, and Alexei Veneziani. Towards a theory for diffusive coupling functions allowing persistent synchronization. *Nonlinearity*, 27(3):501–525, 2014. (Cited on pages 26 and 78.)
- [Pes97] Yakov B. Pesin. *Dimension theory in dynamical systems*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1997. Contemporary views and applications. (Cited on page 76.)
- [PG14] Mason A Porter and James P Gleeson. Dynamical systems on networks: a tutorial. *arXiv preprint arXiv:1403.7663*, 2014. (Cited on page 21.)
- [PSV01] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200, 2001. (Cited on page 1.)
- [PvST17] Tiago Pereira, Sebastian van Strien, and Matteo Tanzi. Heterogeneously coupled maps: hub dynamics and emergence across connectivity layers. *arXiv preprint arXiv:1704.06163*, 2017. (Cited on pages 34, 36, 37 and 38.)
- [QXZ09] Min Qian, Jian-Sheng Xie, and Shu Zhu. *Smooth ergodic theory for endomorphisms*. Springer, 2009. (Cited on page 64.)
- [RGvS15] Lasse Rempe-Gillen and Sebastian van Strien. Density of hyperbolicity for classes of real transcendental entire functions and circle maps. *Duke Mathematical Journal*, 164(6):1079–1137, 2015. (Cited on pages 6, 78 and 119.)

- [RS15] Bob Rink and Jan Sanders. Coupled cell networks: semigroups, Lie algebras and normal forms. *Transactions of the American Mathematical Society*, 367(5):3509–3548, 2015. (Cited on pages 18 and 26.)
- [Rug02] Hans Henrik Rugh. Coupled maps and analytic function spaces. In *Annales scientifiques de l'Ecole normale supérieure*, volume 35, pages 489–535, 2002. (Cited on page 25.)
- [SB05] F Sélley and P Bálint. Mean-field coupling of identical expanding circle maps. arXiv preprint arXiv:1512.05556, 2005. (Not cited.)
- [SB16] Fanni Sélley and Péter Bálint. Mean-field coupling of identical expanding circle maps. *Journal of Statistical Physics*, 164(4):858–889, 2016. (Cited on pages 22, 23, 24 and 26.)
- [Sch04] Matthias Schmitt. Spectral theory for nonanalytic coupled map lattices. *Nonlinearity*, 17(2):671, 2004. (Cited on page 25.)
- [Sél16] Fanni M Sélley. Symmetry breaking in a globally coupled map of four sites. *arXiv preprint arXiv:1612.01310*, 2016. (Cited on pages 22 and 23.)
- [Sim12] David Simmons. Conditional measures and conditional expectation; Rohlin’s disintegration theorem. *Discrete Contin. Dyn. Syst.*, 32(7):2565–2582, 2012. (Cited on pages 147 and 148.)
- [SSTC01] L. P. Shilnikov, A.L. Shilnikov, D.V. Turaev, and L.O. Chua. *Methods of qualitative theory in nonlinear dynamics*, volume 5. World Scientific, 2001. (Cited on page 3.)
- [Str00] Steven H Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1):1–20, 2000. (Cited on pages 21, 22 and 26.)
- [SVM07] J. A. Sanders, F. Verhulst, and J. Murdock. *Averaging methods in nonlinear dynamical systems*, volume 59 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2007. (Cited on page 27.)
- [Tsu01] Masato Tsujii. Fat solenoidal attractors. *Nonlinearity*, 14(5):1011–1027, 2001. (Cited on pages 39, 53 and 119.)
- [US06] Peter J Uhlhaas and Wolf Singer. Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron*, 52(1):155–168, 2006. (Cited on page 19.)

-
- [Via97] Marcelo Viana. Stochastic dynamics of deterministic systems. Lecture Notes XXI Braz. Math. Colloq. IMPA Rio de Janeiro, 1997. (Cited on pages [28](#), [30](#), [50](#) and [51](#).)
- [WAH88] CO Weiss, NB Abraham, and U Hübner. Homoclinic and heteroclinic chaos in a single-mode laser. *Physical Review Letters*, 61(14):1587, 1988. (Cited on page [3](#).)
- [Wil12] Amie Wilkinson. Smooth ergodic theory. In *Mathematics of complexity and dynamical systems. Vols. 1–3*, pages 1533–1547. Springer, New York, 2012. (Cited on page [64](#).)