Collective Behavior in Dynamics on Networks

By

Jordan Andrew Snyder

Dissertation

Submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Applied Mathematics

in the

Office of Graduate Studies

of the

University of California

Davis

Approved:

Raissa M. D'Souza, Chair

James P. Crutchfield

Timothy J. Lewis

Committee in Charge

2019

Contents

List of Figures

Abstract

Collective Behavior in Dynamics on Networks

This dissertation addresses three problems concerning collective behavior in dynamics on networks: coupling in a heterogeneous population of oscillators under global forcing, low-dimensional effective laws of motion for oscillator networks, and degree-targeting the spread of cascades on modular, degree-heterogeneous networks. We refer to these problems as *coupled entrainment*, *coarse graining*, and *degree-targeting cascades*, respectively.

From coupled entrainment, we learn that there is not just one way to measure order in a population of dynamical units, just as there is no unique way to achieve order in a population of dynamical units. Different mechanisms an give rise to different types of order. And, these types can be both competitive and synergistic with one another.

From coarse graining, we find that under certain special assumptions, it is possible to leverage the presence of collective behavior to simplify a dynamical model of a system. From degree-targeting cascades, we learn that it can be possible to average over a highly heterogeneous population and arrive at an accurate representation with very few degrees of freedom. Given that reduced model, we can assess in detail the interplay of network structure and a seeding policy in determining the spread of an activation cascade.

In all cases, analytical insight comes from the existence of a low-dimensional object occupying a privileged position in a high-dimensional state space, and the existence of that low-dimensional object is crucially related to symmetries (broadly construed) of the underlying high-dimensional system.

In coupled entrainment, the key calculation is a linear stability analysis of a particular fixed point of an infinite-dimensional dynamical system. The fixed point in question is already part of a special family of states in which an oscillator with a given natural frequency is pinned to a given phase. The linearization in question has an eigenvalue $-1/\pi$ with infinite degeneracy and a *single* eigenmode with eigenvalue $-1/\pi + K/2$ that passes through zero as the parameter K is increased. Both stages of simplification relied on the delicate symmetry built into the natural frequency distribution and the forcing function.

Specifically, the choice of natural frequencies, together with the forcing function, brought about a fixed point in which the phases of all oscillators were spread around the unit circle in a perfectly symmetrical way.

In coarse-graining, existence of a low-dimensional attractor in a large (but not infinite) oscillator system is the central theme. Dimension reduction analyses that are valid in the $N \to \infty$ limit guide our search in finite dimensions, and we find reasonable and numerically computable conditions under which such dimension reduction is possible. Here we find that identifying which degrees of freedom admit a reasonable model relies not only on knowing which phases form cohesive groups, but also which oscillators are equivalent to which others in terms of their couplings to other oscillators. The result is joint insight into mesoscale synchronization in the Kuramoto model, on the one hand, and general constraints on discovering low-dimensional dynamical models, on the other.

Finally, in studying cascades, the underlying dynamics are discrete, which differentiates them from the oscillators considered previously. Nonetheless, we find that passing to probabilities of activation leads us to a continuous model that can be drastically compressed due to statistical regularities (or symmetries) of the network that hosts the dynamics. In this case, the symmetries are that nodes in the same module are statistically equivalent in terms of their probability to be connected to an active node.

Acknowledgments

It's often said that it takes a village to raise a child. It's less often said, but just as true, that it takes a village to do a PhD.

First and foremost, thanks go to my parents, Lois and Jerry, for laying the foundation and supporting me every step of the way; and to my brother Daniel, for catalyzing an excitement for math that continues to this day.

Thanks to my professors at RPI. In particular, thanks to Jim Napolitano for setting the bar as an educator and for his much-needed encouragement towards intellectual maturity and self-direction.

Thanks to my Los Alamos "A-team": Anatoly Zlotnik, Andrey Lokhov, Andrew Sornborger, and Aric Hagberg. My experiences with all of you at LANL helped me find my legs as a researcher leading a project and gave me crucial perspective on scientific life outside of a university environment.

Thanks to my thesis committee; to Jim Crutchfield, for serving as a role model of sustained commitment to big ideas; to Tim Lewis, for helping me see how to establish my identity as an applied mathematician; and finally, to my advisor, Raissa D'Souza, for her consistent commitment to my intellectual and professional growth, and for believing in me even (and especially) when I seemed determined not to believe in myself.

Thanks to my coauthors on the work contained herein:

- Chapter 2 was done in collaboration with Anatoly Zlotnik and Aric Hagberg
- Chapter 3 was done in collaboration with Anatoly Zlotnik and Andrey Lokhov
- Chapter 4 was done in collaboration with Weiran Cai and Raissa D'Souza

Thanks to the Santa Fe Institute and the broader SFI community, especially the students, staff, and faculty who helped make the 2018 Complex Systems Summer School a reality. The connections and conversations I've had via SFI have been a major part of my inspiration to pursue an academic career.

Last, but not least, thanks and love to my Davis family. You've all made the last six years not just bearable, but a true joy.

- Be good
- Don't panic
- Stay hydrated

Chapter 1

Introduction

1.1 What's the Big Idea?

This is where I announce my grand vision, as I see it today.

My central intellectual interest is to understand how the world got to be the way it is. An important part of that is to figure out what are the basic building blocks of matter and energy, and what relations they have. Rather beyond the scope of that question is the question of how those pieces come together to make atoms, molecules, cells, planets, humans, sailboats, fire extinguishers, and conscious thought (not necessarily in that order). Even if we know exactly the fundamental laws of the universe, we could hardly claim to "understand" everything that comes from them in any truly satisfying sense.

That's why I'm interested in collective behavior. Taking seriously the idea that a collection of simple-in-isolation units interacting with one another may produce a new type of object opens the door to quantitatively and intuitively understanding how the complexity of the human-scale world emerges from the relative simplicity of smaller-scale building blocks when you put enough of them in one place.

To make progress in understanding collective behavior requires looking deeply at concrete instantiations of systems that exhibit it. The process of posing and answering such concrete questions has been the central challenge of my PhD research. This dissertation presents three iterations of that process, each in a slightly different direction and with

different conclusions. I will now describe each of them briefly in the context of collective behavior.

1.2 Stability of entrainment of a continuum of coupled oscillators

For at least the last 40 years (starting with Kuramoto [41]), synchronization has been an extraordinarily fruitful test-bed for fundamental studies of the mechanisms underlying collective behavior. In short, synchronization refers to a case in which a system of oscillators interacts in such a way as to cause their phases to align. Perhaps the most famous result is due to Kuramoto, and gives the value of the coupling strength when synchronization begins in terms of the heterogeneity of oscillators' natural frequencies. Formally,

$$K_c = \frac{2}{\pi g(0)} \tag{1.1}$$

where $g(\omega)$ is the probability density function from which the oscillators' natural frequencies ω are independently drawn, assumed to be unimodal and symmetric about $\omega = 0$.

Kuramoto's result predicts phase coherence will arise due to the oscillators' all-to-all pairwise interactions and in spite of the oscillators' heterogeneous frequencies. Intuitively, we can think of this process as being *bottom-up*, or *self-organized*.

This is, however, only one route to coherence. Another way, pioneered by Adler [2] and exploited by Zlotnik and coauthors [79] to produce coherent motion in a heterogeneous population, is to apply a common driving signal at a fixed target frequency. If an oscillator's frequency is not too far from the driving frequency, then the oscillator will lock to the driving frequency and have an average phase offset related to the frequency detuning (difference between natural and driving frequencies). Likewise, if all oscillators' frequencies are not too far from the driving frequency, then the common driving signal will cause the entire population to move at a single frequency.

Clearly, this mechanism is of a different character than that of coupling; the oscillators are doing the same thing because they are each made to do that same thing. In other words, coherence due to common driving may be characterized as *top-down*, or *centralized*.

These mechanisms result in coherent motion in a population of oscillators, but there are tell-tale differences. In particular, it is possible to have a situation where, under the influence of common driving, every oscillator attains the same frequency but the phases remain spread uniformly about the unit circle. In other words, while the system is perfectly ordered in terms of frequencies, it is maximally disordered in terms of phases. Mutual coupling attenuates phase disorder, and so this situation is a natural setting to contrast bottom-up and top-down mechanisms of collective behavior.

1.3 Coarse-graining for coupled oscillators

The essence of this project is to take seriously the concept of new entities emerging in the dynamics of interacting individuals. Again we use the setting of coupled oscillators, because they offer great analytical understanding and are known to exhibit rich behavior of the sort that we'd like to understand.

Pioneering work by Ott and Antonsen [52] established that in the thermodynamic limit, populations of oscillators behave in such a way that their global average evolves in time according to an autonomous differential equation. In other words, the intricate dynamics of all the individual oscillators averages out and, to predict the future behavior of the global mean field, one only needs to know the global mean field itself (in addition to some easily calculable constant parameters). Their results extend to the case when coupling is not all-to-all, but shows a group structure; in this case, a similar equation holds, but with one mean field for each group.

This fact has been exploited by, e.g., Skardal and coauthors [65] to study multiple stages of synchronization in group-structured systems. Also of note, qualitatively similar phenomena have been observed in dynamics on modular networks, such as the Ising model and discrete spreading processes [29].

In this work, we take a complementary approach. Rather than using statistical arguments that become exact in the thermodynamic limit, we consider the problem of identifying coarse-grained variables, and differential equations that govern them, directly from finite-dimensional time series data. Calculations suggest that equations of the same

form as those derived in [52] should work, and we validate this conclusion with numerical experiments.

1.4 Cascading extinctions in mutualistic networks

This project arose from a desire to understand collective behavior in terms of the properties of a network that connects a population of individuals. One of the most obvious real-world examples of systems that show interesting collective behavior due to a nontrivial interaction structure are ecosystems, where collective behavior can take the form of drastic events such as cascading extinctions. One particular class of ecosystems that have been fruitfully understood from a network perspective is mutualistic ecosystems: those consisting of multiple types of individuals that interact in a mutually beneficial way. A prototypical example of a mutualistic partnership is that between a plant and its pollinator; examples also appear in other contexts, such as between clothing designers and manufacturers [61].

Two network-structural tropes have been observed among many mutualistic networks: nestedness [11] and modularity [51]. Nestedness is the property that the neighbors of any given node tend to be a subset of the neighbors of any node with higher degree. To a first approximation, nestedness is best explained by a broad degree distribution [37]. Modularity is the property that the nodes can be divided up into groups such that most of the links connect nodes in the same group, with comparably few links connecting nodes in different groups [49].

The goal of this project is to understand the joint impact of nestedness and modularity on spreading-type processes on a network, such as cascading extinctions. To do this, we study the linear threshold model, introduced by Watts [71], in which nodes are either active or inactive, and become active if at least some fraction of their neighbors are active. Once active, a node does not become inactive. In an ecological interpretation, the "active" or absorbing state is extinction, and a species goes extinct if greater than some fraction of its mutualistic partner species are extinct.

A low-dimensional approximation of these dynamics on a modular network with an

arbitrary degree distribution in each module was obtained by Gleeson in [29]. This approximation rests on the assumption that the network is large and locally tree-like, and that a certain fraction of the nodes in each module are selected uniformly at random to be initially active. We extend Gleeson's framework to model initially-active nodes which are selected with a probability that depends on their degree. This generalization still allows for uniform seed selection, or, for instance, targeting maximum- or minimum-degree nodes. Remarkably, our extended formulation does not require keeping track of any more dynamical variables than Gleeson's original framework, although the number of parameters is larger.

We then use this analytical framework for degree-targeted seeding to explore the effect of degree targeting on cascade spreading in a modular network. Nematzadeh and coauthors [48] have shown that in a two-module network wherein a cascade is initiated in one module, a certain fraction of inter-module links is required to spread the cascade to the whole network. However, too high a proportion of inter-module links renders the cascade unable to cover even the entire first module. They refer to this phenomenon as the existence of "optimal modularity" for information diffusion. We extend the results of [48] and show that by seeding the highest-degree nodes, the range of optimal modularity is substantially widened, and that the strength of that effect increases with degree heterogeneity.

Chapter 2

Stability of entrainment of a continuum of coupled oscillators

This chapter corresponds closely to the article:

> Snyder, J., Zlotnik, A., & Hagberg, A. (2017). Stability of entrainment of a continuum of coupled oscillators. Chaos: An Interdisciplinary Journal of Nonlinear Science, 27(10), 103108. https://doi.org/10.1063/1.4994567

2.1 Abstract

At least two well-understood approaches can be applied to impose coherent behavior in a diverse population of dynamical systems: the "top-down" approach of applying a common driving signal, and the "bottom-up" approach of imposing pairwise coupling. While these approaches yield similar behaviors, their precise characteristics can put them in opposition. In this chapter we study a situation that highlights both the synergy and tension that can exist between driving and coupling in collections of oscillators.

2.2 Introduction

Coupled oscillators are a paradigmatic example of a system that exhibits collective behavior, owing to their rich phenomenology and relative tractability in certain limits and special cases. Notably, coupled oscillator systems support coherent behavior by multiple mechanisms; we think especially of the bottom-up process of synchronization due to mu-

tual coupling, and the top-down process of entrainment to an external drive. We examine the design of the entrainment process for an uncountably infinite collection of coupled phase oscillators that are all subject to the same periodic driving signal. In the absence of coupling, an appropriately designed input can result in each oscillator attaining the frequency of the driving signal, with a phase offset determined by its natural frequency. We consider a special case of interacting oscillators in which the coupling tends to destabilize the phase configuration to which the driving signal would send the collection in the absence of coupling. In this setting we derive stability results that characterize the trade-off between the effects of driving and coupling, and compare these results to the well-known Kuramoto model of a collection of free-running oscillators.

There is a long history of studying the emergence of coherent motion in oscillators using phase model representations, dating back to Winfree [74, 75] and Kuramoto [41, 43]. The canonical examples that were posed in these early studies have been widely examined in subsequent work [12], because they exhibit a rich phenomenology while admitting beautiful mathematical descriptions within an extensive range of analytical settings. While original studies focused on mutual entrainment [41], in which coherent motion arises purely from interactions between individual units, more recent studies have investigated the effect of externally-imposed coherence in the form of external driving.

A pioneering study of forced, coupled oscillations was performed by Sakaguchi [63], who considered an infinite, heterogeneous, population of oscillators subject to global sinusoidal coupling and uniform sinusoidal forcing. By deriving a self-consistent equation for the order parameter measuring phase alignment, Sakaguchi was able to predict transitions between regimes of incoherence, mutual entrainment, and forced entrainment. These transitions were investigated in more detail using a linear stability approach [4]. The detailed nature of the bifurcations remained elusive, and suggested an underlying two-dimensional structure which had yet to be exploited. This two-dimensional structure was indeed discovered by Ott and Antonsen in their seminal work [52], which uncovered a particular low-dimensional manifold that captures much of the asymptotic behavior of a wide family of models of coupled phase oscillators; in particular, the forced Kuramoto

model was identified as a possible application of this dimension reduction. Subsequent work has shown that, under certain mild conditions, this manifold is globally attractive [53, 54]. This reduction represents an enormous simplification, because in many cases it permits closed-form evolution equations for the synchrony order parameter directly. Building on the framework established by Ott and Antonsen [52], Childs and Strogatz [15] were able study the dynamics of the forced Kuramoto model on the two-dimensional attractive manifold, and found a complete picture of a system's bifurcation structure. It should be noted that studies of the effect of forcing in this context have almost exclusively considered a sinusoidal forcing function, due to the analytical tractability of the resulting model.

Beyond characterizing the phenomenology of natural and engineered complex oscillating systems, emerging applications in neural systems, electrochemistry, and power grid engineering require new capabilities to control and manipulate the behavior of such phenomena. Indeed, the ability to control a system is the ultimate validation of our understanding of its behavior. For oscillating systems, a general picture of frequency modulation by external forcing was first laid out in 1946 by Adler [2], who derived equations describing the amount by which an external drive signal can shift an oscillator's frequency and amplitude. The idea of "injection locking" has since been of major importance in many fields of engineering [60, 9]. One prediction made by Adler was that an oscillator driven at a frequency different from its own may lock to the driving frequency, and exhibit a phase shift relative to the drive signal which is determined by its natural frequency. For the simple case considered originally, this function is sinusoidal. However, for a general forcing function and a general phase oscillator higher harmonics may be present, as seen in experiments and derived analytically [35, 78]. The general framework of using periodic forcing signals to control the entrainment of nonlinear oscillating systems was exploited to explore energy- and time-optimal control strategies for entrainment of one or more phase oscillators [26, 77, 76]. The effect of coupling on the efficacy of these control strategies remains unexplored.

A challenge in specifying the forcing input to control collections of coupled oscillators

is that they are underactuated; the entire collection of similar dynamical systems with possibly complicated individual behavior must be controlled using a small number of inputs. To overcome this challenge we observe that the entrained or coherent state of a controlled collection of oscillators is characterized not only by synchronization to a forcing frequency, but also by the distribution of subsystems on the neighborhood of a nominal periodic orbit. For a finite collection, it is possible to construct a forcing signal to achieve precise control of the relative phases of an ensemble of structurally similar oscillators with slight heterogeneity in frequencies [79]. With the understanding that such "phase-selective control" is possible for small, finite collections, we examine how the mathematical framework can be extended to continuum systems. Further, we examine the effect of coupling between subsystems, which tends to drive phase differences to zero.

In this chapter we explore a continuum approximation of a very large collection of coupled oscillators subject to a common periodic (but non-sinusoidal) forcing, so that both coupling and forcing influence the collective behavior. Specifically, we consider a situation in which the forcing drives individual phases to be maximally different (in a certain precise sense), while the coupling tends to align the phases. To quantify the trade-off between these two effects, we compute, as a function of the coupling strength, the asymptotic stability of a fixed point in which the phases show no global alignment. By finding the critical coupling strength above which this fixed point is unstable, we demonstrate that mutual synchronization of entrained coupled oscillators occurs *before* mutual synchronization of unforced coupled oscillators, despite the imposed diversity of phases. Moreover, numerical experiments confirm that the external forcing has facilitated phase alignment which is *greater* than that in the unforced case. Our results demonstrate that measuring only phase alignment is bound to miss important information about the global organization of a population of oscillators.

2.3 Preliminaries

2.3.1 Entrainment of a single oscillator

Periodic motion abounds in the natural and engineered world, and a dynamical system with a limit cycle solution provides a general mathematical model of phenomena that exhibit such rhythms. The technique of phase reduction is a powerful method of analyzing limit-cycle behavior that was pioneered by Arthur Winfree [74]. The basic idea is that each point on a given periodic orbit can be assigned a "phase", in such a way that the phase advances with time at a constant rate. If the periodic orbit in question is a stable limit cycle, then any point sufficiently close to the limit cycle will eventually approach it. In this way, it is possible to define phase for points that are near, but not exactly on, the limit cycle.

Using a perturbation approach, it is then possible to approximate, in the phase picture, the response of a near-limit-cycle trajectory to an external force [41, 64]. The result is a differential equation of the form

$$\dot{\psi} = \omega + Z(\psi)u\,, \tag{2.1}$$

where $\psi \in [0, 2\pi)$ is the phase, " $\dot{}$ " denotes the derivative with respect to time, $\omega \in \mathbb{R}$ is the natural frequency of the limit cycle, $u = u(t)$ is an external forcing, and $Z(\psi)$ is known as the *phase response curve*, or PRC. The PRC determines the change in phase resulting from an infinitesimal external force applied at a given phase on the limit cycle [47, 27].

From dynamics of the form (2.1), we can observe various phenomena, including *entrainment* (also known as *injection locking*). Entrainment is a situation in which an oscillator is driven at a frequency different from its own natural frequency, and subsequently moves at the frequency of the driving signal.

To see how entrainment can arise, consider a driving signal $u(t) = v(\Omega t)$, where v has a period of 2π, so that Ω denotes the (angular) frequency of the driving signal. Inserting this into (2.1) yields

$$\dot{\psi} = \omega + Z(\psi)v(\Omega t). \tag{2.2}$$

If Ω is not too far from ω, we suppose that ψ will behave as Ωt, plus a slowly-varying phase offset. We formalize this supposition by making the change of coordinates $\psi = \Omega t + \phi$, where ϕ now represents the phase offset. In the ϕ coordinate system, the dynamics now read

$$\dot{\phi} = \Delta\omega + Z(\phi + \Omega t)v(\Omega t), \tag{2.3}$$

where we have introduced the *frequency detuning* $\Delta\omega \equiv \omega - \Omega$. Finally it is possible to approximate (2.3) [34] by the time-averaged system

$$\dot{\varphi} = \Delta\omega + \Lambda_v(\varphi), \tag{2.4}$$

where we have introduced the *interaction function*

$$\Lambda_v(\varphi) = \int_0^{2\pi} Z(\varphi + \theta)v(\theta)d\theta, \tag{2.5}$$

in the sense that there exists a change of variables $\varphi = \phi + h(\varphi, \phi)$ that maps solutions of (2.3) to those of (2.4).

The equation (2.4) is an especially simple one-dimensional, autonomous ODE. If the equation $\Delta\omega + \Lambda_v(\varphi) = 0$ has a solution φ^*, then $\varphi(t) = \varphi^*$ is a fixed point solution of (2.4); if, moreover, $\mathrm{d}\Lambda_v/\mathrm{d}\varphi(\varphi^*) < 0$, then this fixed point is asymptotically stable. If this is the case, then we say that the oscillator can be *entrained* by the driving signal.

2.3.2 Entrainment of an ensemble of oscillators

We next consider the case of forcing of many similar, but non-identical, oscillators by a single driving signal [79]. By similar but non-identical, we mean that the oscillator all have a common PRC, but heterogeneous natural frequencies ω. The phenomenon that we wish to highlight in this context is frequency alignment without phase alignment. Note that we do not yet consider any coupling between oscillators.

Suppose we have a system of N phase oscillators of the type described above, all governed by the same PRC, $Z(\psi)$, but having distinct natural frequencies, $\{\omega_i\}$. If these oscillators are simultaneously subject to the same periodic force $v(\Omega t)$, then the system can be described by the equations

$$\dot{\varphi}_i = \Delta\omega_i + \Lambda_v(\varphi_i), \qquad i = 1, \ldots, N\,, \tag{2.6}$$

where $\Delta\omega_i = \omega_i - \Omega$.

First, note that if (2.6) has an asymptotically stable fixed point solution φ_i^* for every i, then the population can display *frequency alignment*; as $t \to \infty$, the phase of the i^{th} oscillator will approach a constant offset of the driving phase Ωt. However, we can immediately see that the phases themselves will, in general, be different, since the solution to the fixed point equation $\Delta\omega_i + \Lambda_v(\varphi_i) = 0$ depends on the frequency detuning $\Delta\omega_i$. This potential phase diversity is displayed in Fig. 2.1 for an example interaction function. In what follows, we consider what happens if the phases of the oscillators are coupled to one another.

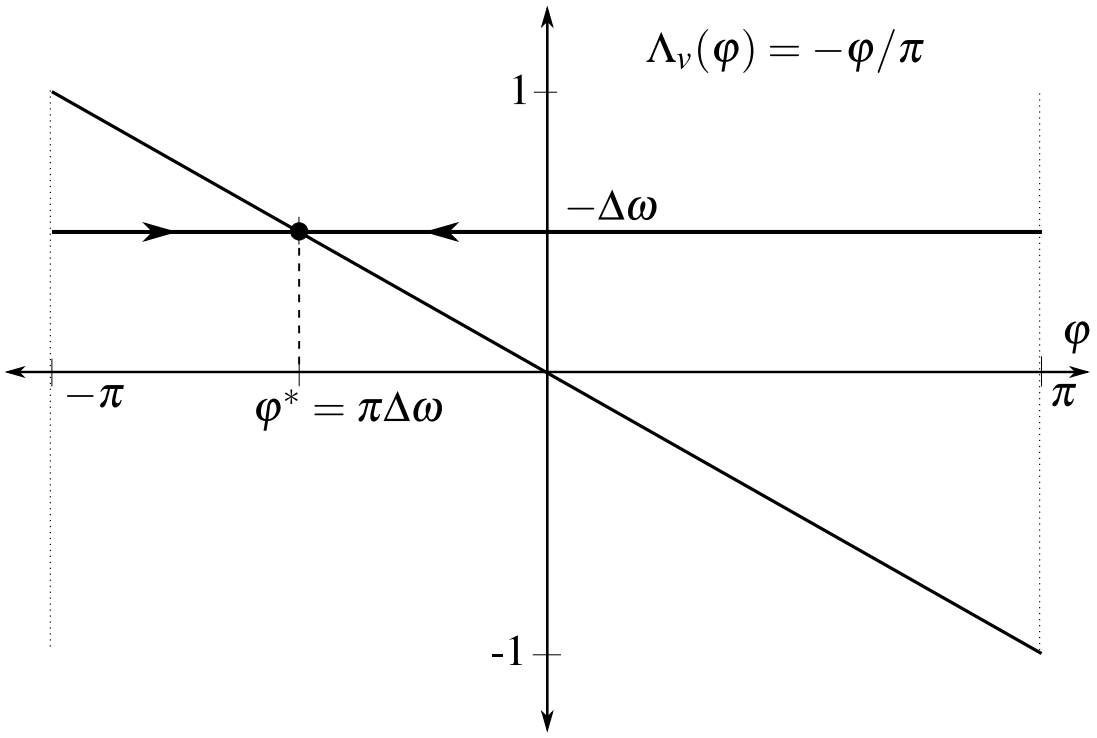

Figure 2.1. An example of frequency locking with phase diversity. Diagonal black line represents a possible interaction function, here a sawtooth function on the circle $[-\pi, \pi]$. Bold horizontal line is the phase line for an oscillator of frequency detuning $\Delta\omega$, with arrows indicating direction of flow. Vertical dashed line indicates the position of the asymptotically stable fixed point, equal to π times the frequency detuning.

2.3.3 The Kuramoto Model

To frame our study of phase coupling, we discuss some standard methods and results relating to synchronization of phase oscillators. Kuramoto introduced a model of the

form

$$\dot{\varphi}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\varphi_j - \varphi_i)\,, \tag{2.7}$$

which was derived as a "simplest" model for a collection of self-sustained linearly-coupled oscillators [41]. Here $\{\varphi_i\}$ are the phases of N oscillators, $\{\omega_i\}$ their natural frequencies (which we allow to take any real values), and $K > 0$ is the strength of coupling.

This ODE can be instructively rewritten in the form

$$\dot{\varphi}_i = \omega_i + KR\sin(\Phi - \varphi_i)\,, \tag{2.8}$$

where we have used the *synchrony* $R \in [0, 1]$, and the *average phase* $\Phi \in [0, 2\pi)$, first introduced by Kuramoto [41] and defined by the formula

$$R\mathrm{e}^{i\Phi} = \frac{1}{N}\sum_{j=1}^{N}\mathrm{e}^{i\varphi_j}. \tag{2.9}$$

In this sense, this form of coupling is *mean-field* in character, as each phase feels a force determined by an average over the entire population.

The key features of this model are

1. The oscillators have differing intrinsic frequency: $\omega_i \neq \omega_j$,
2. The coupling tends to drive phases towards the mean (provided $K > 0$, which we assume throughout).

These two features are at odds with each other, and they undergo a trade-off at a critical value of the coupling strength, $K = K_c^{\mathrm{unf}}$ (where we use the superscript "unf" to emphasize that this is the critical coupling strength in the *unforced* case). If $K < K_c^{\mathrm{unf}}$ the population of oscillators does not show global alignment towards any particular phase, while for $K > K_c^{\mathrm{unf}}$, this situation breaks down and a subset of the oscillators attains the same frequency and group together in phase, establishing a preferred direction and a nonzero value of the synchrony R.

To make these statements precise it is useful to consider a *mean-field approximation*. We suppose that the population of oscillators is large enough that averaging over this

population is well approximated by averaging over a probability distribution that describes the behavior of a typical oscillator. General background on the mean field Kuramoto model can be found in various review articles [66, 1].

The main result we quote from the extensive body of literature on the Kuramoto model is that in the limit of $N \to \infty$, if the oscillators' natural frequencies are drawn at random from a probability distribution having density $g(\omega)$, unimodal and symmetric about zero, then the critical coupling strength described above is given by

$$K_c^{\text{unf}} = \frac{2}{\pi g(0)}. \tag{2.10}$$

The expression (2.10) can be taken as a precise quantification of the trade-off between intrinsic disorder ($g(0)$) and coupling (K). The possibly surprising fact that K_c^{unf} depends *only* on the value of g at the center of the distribution, and no other features of this distribution, is because the first oscillators to synchronize are those at the center of the distribution. The rest of the density g then determines the growth of R with $K > K_c^{\text{unf}}$.

In what follows, we will define a new model, show that it exhibits behavior that is qualitatively similar to that of the Kuramoto model, and find the location of the corresponding critical point. The expression (2.10) will serve as reference to interpret our results.

2.4 Model for Forcing of Coupled Oscillations

2.4.1 Finite N

We now formulate a model of a population of oscillators that exhibits both frequency alignment by broadcast forcing and phase alignment by attractive coupling. Many similar models have been developed [63, 46, 15, 52, 4], and our present formulation aims to augment the rich existing literature.

In general, we can consider each oscillator to respond to external forcing according to one phase response curve, and to respond to forcing from its neighboring oscillators according to another phase response curve. That is,

$$\dot{\psi_i} = \omega_i + Z_e(\psi_i)u(t) + \frac{K}{N}\sum_{j=1}^{N} Z_c(\psi_i)f(\psi_j) \tag{2.11}$$

where Z_e is the PRC for external forcing, Z_c is the PRC for coupling, and $f()$ describes the force an oscillator exerts on its neighbors as a function of its phase. The prefector K/N allows us to adjust the coupling strength K in a way that allows comparison between different values of N.

Assuming, as before, that $u(t) = v(\Omega t)$ with v having period 2π, we move into a rotating reference frame with frequency Ω and average over one period of the driving signal, obtaining the averaged equations

$$\dot{\varphi}_i = \Delta\omega_i + \Lambda_v(\varphi_i) + \frac{K}{N}\sum_{j=1}^{N} g(\varphi_j - \varphi_i) \tag{2.12}$$

where $\varphi_i, \Delta\omega_i$, and Λ_v are defined as before (2.4) and

$$g(\Delta\varphi) = (2\pi)^{-1}\int_0^{2\pi} Z_c(\theta + \Delta\varphi) f(\theta) d\theta$$

.

Clearly, many different systems may be defined in this form given appropriate choices for Z_e, Z_c, v, and f. In order to exhibit the qualitative features of phase dispersion caused by external forcing combined with phase alignment caused by coupling, while retaining tractability, we assume that Z_c and f are such that $g(\Delta\varphi) = \sin(\Delta\varphi)$.

Hence, we take a model of the form

$$\dot{\varphi}_i = \Delta\omega_i + \Lambda_v(\varphi_i) + \frac{K}{N}\sum_{j=1}^{N} \sin(\varphi_j - \varphi_i), \tag{2.13}$$

which can also be written

$$\dot{\varphi}_i = \Delta\omega_i + \Lambda_v(\varphi_i) + KR\sin(\Phi - \varphi_i), \tag{2.14}$$

with R and Φ defined as in (2.9).

As a first step, we choose $\{\Delta\omega_i\}$ and Λ_v such that all oscillators can be entrained individually, but the resulting phase offsets are as far as possible from alignment. This can be achieved by setting

$$\Delta\omega_i = \frac{2i}{N} - 1\,, \tag{2.15}$$

and

$$\Lambda_v(\varphi) = \frac{-\varphi}{\pi}, \qquad \varphi \in (-\pi, \pi]. \tag{2.16}$$

We refer to the function defined in (2.16) as the *sawtooth interaction function*, or just *sawtooth*, as it has a sawtooth shape when plotted on $\mathbb{R}$ (see Fig. 2.1).

The standard unforced Kuramoto model with this choice of natural frequencies has been recently studied by Ottino-Löffler and Strogatz [55], who found the asymptotic behavior of the locking threshold as $N \to \infty$, in agreement with results in the thermodynamic limit obtained earlier by Pazó [57]. These results will serve as a reference to put our findings in context. For now, we return to the forced case.

In the absence of coupling ($K = 0$), the i^{th} oscillator will be driven to a phase offset φ_i^* defined by

$$\Delta\omega_i + \Lambda_v(\varphi_i^*) = 0 \implies \varphi_i^* = \pi\Delta\omega_i = \frac{2\pi i}{N} - \pi. \tag{2.17}$$

A straightforward calculation shows that for this phase configuration, the synchrony is $R = 0$. For this reason, we refer to this fixed point as the *desynchronized state*. Another term used to describe such a state is *splay state*. The point $\varphi^* = (\varphi_i^*) \in (-\pi, \pi]^N$ is a fixed point of the dynamics (2.13) for any value of coupling strength K.

In this respect, the situation is similar to the *incoherent state* discussed for the Kuramoto model in Section 2.3.3, with the key difference that in this case, all oscillators have attained identical frequency locking to the forcing input. We proceed to study the asymptotic stability of this fixed point as a function of K, and obtain a critical coupling strength K_c analogous to K_c^{unf} as defined in (2.10).

2.4.2 The $N \to \infty$ limit

Next we introduce a thermodynamic limit of the model (2.13), and the fixed point corresponding to that defined in (2.17).

We replace our population of oscillators, formerly a collection of N individual oscillators with natural frequencies evenly spaced from -1 to 1, by a continuum of oscillators with natural frequencies distributed uniformly on $[-1, 1]$.

Because our state of interest for finite N is such that each oscillator's phase is fixed at a value determined by its natural frequency, we describe the state of our infinite system

by a function $\varphi(\omega)$ that gives the phase of any oscillator having natural frequency ω. As the system evolves the whole function $\varphi(\omega)$ will change in time, but for visual clarity we omit writing the time-dependence explicitly when discussing fixed points. This sort of formulation is used, for example, by Mirollo and Strogatz [46], except that oscillators are indexed by their "pinning phase" rather than their natural frequency. We describe their work in more detail in Section 2.8.

To determine fixed points, we must establish the dynamics in the appropriate continuum setting. The intrinsic dynamics and effects of forcing remain the same, so we only need to concern ourselves with the coupling term. For finite N, we simply had an average over the population, and in the infinite setting, we use a mean-field approach to say that averaging over the infinite population is equivalent to averaging over the distribution of natural frequencies[66]. Our infinite-dimensional dynamics are

$$\partial_t \varphi(\omega) = \omega + \Lambda_v(\varphi(\omega)) + K \int_{\mathbb{R}} g(\omega') \sin(\varphi(\omega') - \varphi(\omega)) \mathrm{d}\omega' \,, \tag{2.18}$$

where g is the density of the distribution of natural frequencies. These dynamics can be rewritten in the form

$$\partial_t \varphi(\omega) = \omega + \Lambda_v(\varphi(\omega)) + KR \sin(\Phi - \varphi(\omega)) \,, \tag{2.19}$$

where R and Φ are the synchrony and average phase, defined for the infinite system as

$$R\mathrm{e}^{i\Phi} = \int_{\mathbb{R}} g(\omega) \mathrm{e}^{i\varphi(\omega)} \mathrm{d}\omega. \tag{2.20}$$

Using the sawtooth interaction function introduced above (see (2.16)), and $g(\omega) = 1/2$ for $\omega \in [-1, 1]$ and 0 elsewhere, the fixed point condition for φ now reads

$$0 = \omega - \frac{\varphi(\omega)}{\pi} + K \int_{-1}^{1} \frac{1}{2} \sin(\varphi(\omega') - \varphi(\omega)) \mathrm{d}\omega' \,, \tag{2.21}$$

A straightforward calculation shows that the function $\varphi(\omega) = \pi\omega$ satisfies the condition (2.21). Note that this is precisely the infinite-N analog of the finite-N fixed point defined in (2.17). In what follows, we perform a linear stability analysis, finding the coupling strength K_c at which this state becomes unstable.

2.5 Stability Analysis of the Entrainment Phase Distribution

2.5.1 Finite N

We now analyze the stability of the fixed point $\varphi^* = (\varphi_i^*) \in (-\pi, \pi]^N$ as defined in equation (2.17).

Asymptotic stability of φ^* is controlled by the spectrum $\sigma(\mathbf{J})$ of the Jacobian $\mathbf{J}$ of the right-hand side of (2.13) with respect to φ, evaluated at φ^*. If every element of $\sigma(\mathbf{J})$ has negative real part, then φ^* is an asymptotically stable fixed point, and if any element of $\sigma(\mathbf{J})$ has positive real part, then φ^* is unstable[67]. The matrix elements of $\mathbf{J}$ are

$$\begin{aligned} J_{ij} = &\left(\Lambda_v'(\varphi_i^*) - \frac{K}{N} \sum_{k \neq i} \cos(\varphi_i^* - \varphi_k^*) \right) \delta_{ij} \\ &+ (1 - \delta_{ij}) \frac{K}{N} \cos\left(\varphi_i^* - \varphi_j^*\right) , \end{aligned} \tag{2.22}$$

where δ_{ij} is the Kronecker delta.

Next we calculate the spectrum of $\mathbf{J}$. By symmetry of the phase configuration, the sum in the diagonal term is independent of i, and can be computed by noticing

$$\begin{aligned} 0 &= \sum_{k=1}^{N} \cos(\varphi_i^* - \varphi_k^*) \\ &= \cos(\varphi_i^* - \varphi_i^*) + \sum_{k \neq i} \cos(\varphi_i^* - \varphi_k^*) , \end{aligned} \tag{2.23}$$

where the first equality in (2.23) follows from symmetry ($R = 0$). Hence

$$\sum_{k \neq i} \cos(\varphi_i^* - \varphi_k^*) = -\cos(\varphi_i^* - \varphi_i^*) = -1 . \tag{2.24}$$

This allows us to write the matrix entries in a simpler form, which will facilitate calculation of eigenvalues,

$$\begin{aligned} J_{ij} &= \left(\Lambda_v'(\varphi_i^*) + \frac{K}{N} \right) \delta_{ij} + (1 - \delta_{ij}) \frac{K}{N} \cos\left(\varphi_i^* - \varphi_j^*\right) \\ &= \frac{-1}{\pi} \delta_{ij} + \frac{K}{N} \cos(\varphi_i^* - \varphi_j^*) . \end{aligned} \tag{2.25}$$

Note that we have used $\Lambda_v'(\varphi_i^*) = -1/\pi$ for all i, and $\delta_{ij} + (1-\delta_{ij})\cos(\varphi_i^* - \varphi_j^*) = \cos(\varphi_i^* - \varphi_j^*)$ for all i, j. We can write this in matrix form as

$$\mathbf{J} = K\mathbf{C} - \frac{1}{\pi}\mathbf{I}\,, \tag{2.26}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{C}$ is the matrix with entries $C_{ij} = N^{-1}\cos(\varphi_i^* - \varphi_j^*)$. This form makes it clear that to find the eigenvalues of $\mathbf{J}$ for arbitrary values of K, it suffices to find the eigenvalues of $\mathbf{C}$. To do this, we can write the action of $\mathbf{C}$ on an arbitrary vector x as

$$\begin{aligned}(\mathbf{C}x)_i &= \sum_{j=1}^{N} C_{ij}x_j = N^{-1}\sum_{j=1}^{N}\cos(\varphi_i^* - \varphi_j^*)x_j \\ &= \cos(\varphi_i^*)\left[N^{-1}\sum_j \cos(\varphi_j^*)x_j\right] \\ &\quad + \sin(\varphi_i^*)\left[N^{-1}\sum_j \sin(\varphi_j^*)x_j\right]\,,\end{aligned} \tag{2.27}$$

where we have used the sum angle identity for cosine. The range of $\mathbf{C}$ is spanned by the vectors $e^1 = (\cos(\varphi_i^*))_{i=1}^N$ and $e^2 = (\sin(\varphi_i^*))_{i=1}^N$. Each of these is in fact an eigenvector with eigenvalue $1/2$, which follows from

$$\begin{aligned}(\mathbf{C}e^1)_i &= \cos(\varphi_i^*)\left[N^{-1}\sum_j \cos(\varphi_j^*)\cos(\varphi_j^*)\right] + \sin(\varphi_i^*)\left[N^{-1}\sum_j \sin(\varphi_j^*)\cos(\varphi_j^*)\right] \\ &= \cos(\varphi_i^*)\left[N^{-1}\sum_j \cos^2(\varphi_j^*)\right] \\ &= \cos(\varphi_i^*)\left[N^{-1}\sum_j \frac{1+\cos(2\varphi_j^*)}{2}\right] \\ &= \frac{1}{2}\cos(\varphi_i^*) = \frac{1}{2}(e^1)_i\,,\end{aligned} \tag{2.28}$$

and similarly for e^2. Hence e^1 and e^2 are eigenvectors of $\mathbf{C}$ with eigenvalue $1/2$, and all other eigenvalues of $\mathbf{C}$ are zero.

Finally, we can find the eigenvalues of $\mathbf{J}$ for arbitrary K. Notice that

$$\begin{aligned}
\lambda \in \sigma(\mathbf{J}) &\iff \det(\mathbf{J} - \lambda \mathbf{I}) = 0 \\
&\iff \det\left(K\mathbf{C} - \left(\frac{1}{\pi} + \lambda\right)\mathbf{I}\right) = 0 \\
&\iff \det\left(\mathbf{C} - K^{-1}\left(\frac{1}{\pi} + \lambda\right)\mathbf{I}\right) = 0 \\
&\iff K^{-1}\left(\frac{1}{\pi} + \lambda\right) \in \sigma(\mathbf{C})\,. \qquad (2.29)
\end{aligned}$$

The eigenvalues λ of $\mathbf{J}$ are of the form $\lambda = -1/\pi + K\mu$, for $\mu \in \sigma(\mathbf{C}) = \{0, 1/2\}$. In other words,

$$\sigma(\mathbf{J}) = \left\{\frac{-1}{\pi}, \frac{-1}{\pi} + \frac{K}{2}\right\}\,. \qquad (2.30)$$

so the desynchronized state has a critical point $K_c = 2/\pi$ and is linearly stable when $K < 2/\pi$, and linearly unstable for $K > 2/\pi$.

2.5.2 The $N \to \infty$ limit

Finally we will perform a linear stability analysis of the desynchronized fixed point of the infinite-N model (2.18).

To obtain a linearization of the dynamics near the fixed point $\varphi^*(\omega) = \pi\omega$, we consider an infinitesimal perturbation,

$$\varphi(\omega) = \varphi^*(\omega) + \epsilon\eta(\omega)\,, \qquad (2.31)$$

where $\eta\colon [-1, 1] \to \mathbb{R}$ is a function which we take to be bounded and measurable.

First, inserting the form $\varphi(\omega) = \varphi^*(\omega) + \epsilon\eta(\omega)$ into equation (2.18) yields

$$\begin{aligned}
\partial_t(\varphi^* + \epsilon\eta) =& \omega - \frac{\varphi^*(\omega) + \epsilon\eta(\omega)}{\pi} \qquad (2.32) \\
&+ K\int_{-1}^{1} \frac{1}{2}\sin\left(\varphi^*(\omega') - \varphi^*(\omega) + \epsilon\left(\eta(\omega') - \eta(\omega)\right)\right)\mathrm{d}\omega'\,. \qquad (2.33)
\end{aligned}$$

Next we expand the sine function in the integrand around the point $\varphi^*(\omega') - \varphi^*(\omega)$, and

obtain

$$\begin{aligned}\sin(\varphi^*(\omega') - \varphi^*(\omega) + \epsilon(\eta(\omega') - \eta(\omega))) = &\sin(\varphi^*(\omega') - \varphi^*(\omega)) \\ &+ \epsilon\cos(\varphi^*(\omega') - \varphi^*(\omega))(\eta(\omega') - \eta(\omega)) \\ &+ \mathcal{O}(\epsilon^2). \end{aligned} \tag{2.34}$$

From here we can read off the terms of order ϵ^0 from each side of the equation, and get

$$\partial_t \varphi^* = \omega - \frac{\varphi^*(\omega)}{\pi} + K \int_{-1}^{1} \frac{1}{2} \sin\left(\varphi^*(\omega') - \varphi^*(\omega)\right) \mathrm{d}\omega' \,, \tag{2.35}$$

which clearly holds, as each side evaluates to zero for all $\omega \in [-1, 1]$.

Next, we gather terms of order ϵ^1 and obtain (dropping the ϵ factor from all terms)

$$\partial_t \eta = -\frac{\eta(\omega)}{\pi} + K \int_{-1}^{1} \frac{1}{2} \cos\left(\varphi^*(\omega') - \varphi^*(\omega)\right) \left[\eta(\omega') - \eta(\omega)\right] \mathrm{d}\omega'. \tag{2.36}$$

We can in fact simplify the integral above by noticing that

$$\begin{aligned} &\int_{-1}^{1} \frac{1}{2} \cos\left(\varphi^*(\omega') - \varphi^*(\omega)\right) \left[\eta(\omega') - \eta(\omega)\right] \mathrm{d}\omega' \\ &\quad = \int_{-1}^{1} \frac{1}{2} \cos\left(\varphi^*(\omega') - \varphi^*(\omega)\right) \eta(\omega') \mathrm{d}\omega' - \eta(\omega) \int_{-1}^{1} \frac{1}{2} \cos\left(\varphi^*(\omega') - \varphi^*(\omega)\right) \mathrm{d}\omega' \\ &\quad = \int_{-1}^{1} \frac{1}{2} \cos\left(\varphi^*(\omega') - \varphi^*(\omega)\right) \eta(\omega') \mathrm{d}\omega', \end{aligned} \tag{2.37}$$

which follows from the symmetry of the phase configuration φ^*. We then arrive at the linearized dynamics

$$\partial_t \eta(\omega) = -\frac{1}{\pi}\eta(\omega) + K \int_{-1}^{1} \frac{1}{2} \cos(\varphi^*(\omega') - \varphi^*(\omega)) \eta(\omega') \mathrm{d}\omega'. \tag{2.38}$$

Finally, we demonstrate the diagonalization of the linearized dynamics (2.38) in the Fourier basis. As η is a function on $[-1, 1]$, the appropriate Fourier basis is $\{\mathrm{e}^{\mathrm{i}k\pi\omega} | k \in \mathbb{Z}\}$, so we write

$$\eta(\omega) = \sum_{k \in \mathbb{Z}} c_k(t) \mathrm{e}^{\mathrm{i}k\pi\omega} \,, \tag{2.39}$$

with the understanding that η is real-valued and the coefficients $\{c_k\}$ will obey $\overline{c_k} = c_{-k}$, where "$\bar{\ }$" denotes complex conjugate.

Next, we use $\varphi^*(\omega) = \pi\omega$ and Euler's formula to write

$$\cos\left(\varphi^*(\omega') - \varphi^*(\omega)\right) = \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\pi(\omega'-\omega)} + \mathrm{e}^{-\mathrm{i}\pi(\omega'-\omega)}\right). \tag{2.40}$$

Inserting (2.40) and (2.39) into (2.38) gives

$$\partial_t\eta(\omega) = -\frac{1}{\pi}\eta(\omega) + K\int_{-1}^{1}\frac{1}{4}\sum_{k\in\mathbb{Z}}c_k(t)\mathrm{e}^{\mathrm{i}k\pi\omega'}\left(\mathrm{e}^{\mathrm{i}\pi(\omega'-\omega)} + \mathrm{e}^{-\mathrm{i}\pi(\omega'-\omega)}\right)\mathrm{d}\omega'. \tag{2.41}$$

The only terms of the sum that do not vanish in the integral are those with $k = \pm 1$. For the $k = \pm 1$ terms, the integral evaluates to

$$\int_{-1}^{1}\frac{1}{4}c_1(t)\mathrm{e}^{\mathrm{i}\pi\omega'}\mathrm{e}^{-\mathrm{i}\pi(\omega'-\omega)}\mathrm{d}\omega' = \frac{1}{2}c_1(t)\mathrm{e}^{\mathrm{i}\pi\omega} \tag{2.42}$$

and likewise for $k = -1$. This shows that the coupling term acts on η diagonally in the Fourier basis. Equating Fourier coefficients on each side of (2.41), we obtain

$$k = \pm 1\colon \quad \partial_t c_k(t) = \left(\frac{-1}{\pi} + \frac{K}{2}\right)c_k(t) \tag{2.43}$$

$$k \neq \pm 1\colon \quad \partial_t c_k(t) = \frac{-1}{\pi}c_k(t)\,. \tag{2.44}$$

Hence all Fourier components of the perturbation η except for the first decay exponentially in time with a rate $1/\pi$, while the first Fourier component will grow or shrink with time, depending on the sign of $-1/\pi + K/2$. Specifically, if $K < 2/\pi$, then the first Fourier mode also decays in time, while if $K > 2/\pi$, the first Fourier mode grows in time, and the fixed point φ^* is unstable. Hence we have, as in the finite-N case, the critical coupling strength $K_c = 2/\pi$.

2.5.3 Interpretation

In both the finite- and infinite-dimensional versions of our model, we have found that nonzero synchrony spontaneously develops as the coupling strength K exceeds $K_c = 2/\pi$. We contrast this result with that for the corresponding unforced model,

$$\dot{\varphi}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\varphi_j - \varphi_i)\,, \tag{2.45}$$

where $\omega_i = 2\pi i/N - \pi$. While the standard result (2.10) does not directly apply in this case, since the uniform density is not unimodal, it has been established by Pazó [57] that the synchronization transition does in fact occur at $K_c^{\text{unf}} = 4/\pi = 2/(\pi g(0))$, which is twice the value at which the forced model begins to show nonzero synchrony. This result can be considered surprising, given that we have taken a forcing term, (2.16), that was designed specifically to drive the system to a state of zero synchrony.

The situation becomes clearer if we compare the desynchronized state present in the forced model to the incoherent state in the unforced model. The desynchronized state, defined by $\varphi(\omega) = \pi\omega$, has zero synchrony as measured by the order parameter R. It has the property that every oscillator moves at equal frequency. This is in contrast with the incoherent state of the unforced Kuramoto model, in which each oscillator moves at its own natural frequency. Hence, in the sense of frequencies, the desynchronized state is far more organized than the incoherent state, although this fact is missed by the synchrony parameter R, which only measures instantaneous alignment of phases.

To understand the role of frequency alignment in establishing phase alignment, it is instructive to consider again the standard unforced Kuramoto model. As we have already quoted (2.10), the critical coupling strength is $K_c^{\text{unf}} = 2/\pi g(0)$, where g is the density of the distribution of natural frequencies. Intuitively, this expression captures the trade-off between disorder in the natural frequencies and the ordering influence of coupling; the tighter the distribution of natural frequencies, the larger $g(0)$, and the smaller K_c^{unf}. In other words, the coupling strength must be large enough to overcome the diversity of natural frequencies in order to bring about a preferred phase.

In the desynchronized state of the forced model, the oscillators move with a single frequency. Hence, there is no disorder to be overcome by the coupling. All that keeps the system in the desynchronized state is the forcing, which appears as the eigenvalue $-1/\pi$ in the spectrum of the Jacobian. The second eigenvalue, $-1/\pi + K/2$, directly captures the trade-off between the driving and the coupling, showing that the stability of the entrained state is the only force that needs to be countered by coupling.

2.6 Numerical Simulations

Here we present some numerical studies of the dynamical system (2.13), which confirm the bifurcation at $K = K_c = 2/\pi$ and illuminate the system's behavior away from the bifurcation point. To serve as reference, we also present data from numerical solution of the system in the absence of forcing which is the Kuramoto model with evenly spaced natural frequencies.

As we can see in Fig. 2.2, the synchrony R achieved at any value of the coupling strength K greater than $2/\pi$ is greater in the forced case than in the unforced case. This confirms the conclusion that entrainment by broadcast periodic forcing has brought the system closer to synchrony, as measured by the order parameter R.

In the numerical simulations above, we find a sharp increase in the steady-state value of R as a function of coupling strength K. To obtain a deeper understanding of the nature of this transition and of the $R > 0$ fixed point of (2.13), we perform, for a range of N values, numerical continuation of the $R > 0$ fixed point with respect to the bifurcation parameter K using the numerical continuation software AUTO [23].

To perform numerical continuation with AUTO, it is first necessary to locate an attractor (in our case, a fixed point) on the branch of interest. For each N from 3 to 100, this was accomplished by numerical integration of (2.13) until stationarity with $K = 0.7$. This value of K was chosen as it is greater than $K_c = 2/\pi$, and was observed to lead to an $R > 0$ fixed point in all instances. The AUTO software was then instructed to locate a connected family of fixed points in the joint parameter-state space $\mathbb{R} \times [-\pi, \pi]^N \ni (K, \varphi)$, searching in the negative K direction from the user-supplied fixed point. AUTO equation and constants files, including initial fixed point locations for $3 \leq N \leq 100$, are available upon request.

Representative results of the continuation just described are shown in Fig. 2.3. In particular, we find that for any $N = 3 \ldots 100$, the stable $R > 0$ branch undergoes a saddle-node bifurcation at a coupling strength $K = K_s(N) < K_c = 2/\pi$. The unstable portion of this branch exists for all $K \in [K_s(N), K_c]$, and meets the $R = 0$ branch (i.e. the desynchronized fixed point) transversally, precisely at $K = K_c$. Fig. 2.3 clearly shows,

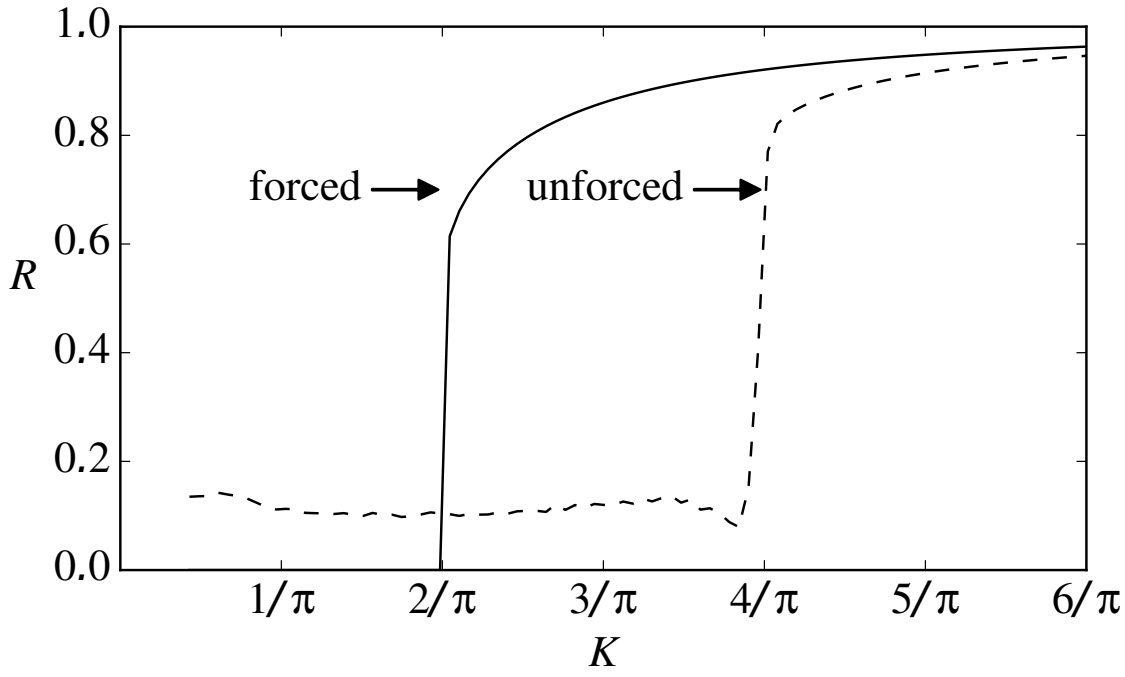

Figure 2.2. Synchronization R vs. coupling strength K. In the unforced case (dashed line) the synchronization threshold is $K_c^{\rm unf} = 4/\pi$. When forcing is added to drive the system to a splay state of equally distributed phase angles it synchronizes at a lower coupling strength $K_c = 2/\pi$. The data were generated from a simulation of $N = 50$ oscillators, starting from random initial conditions. For the forced case, integration was carried out until the system was determined to be at a fixed point. For the unforced case, integration was carried out until the system was determined to be in a statistically steady state.

for $N = 5, 10, 20$, the existence of a bistable region $[K_s(N), K_c]$, implying that hysteresis is possible upon slow variation of K.

Moreover, we find that the shape of the bifurcation diagram in the bistable region obeys a strong regularity across different values of N. In particular, the width of the bistable region, namely $K_c - K_s(N)$, follows a power-law scaling with N, with exponent -1.67. Additionally, the value of R at the saddle-node point, which we denote $R_s(N)$, is observed to approach a value $R_c = 1/2$ from below, according to a power-law with exponent -1.29 (see Fig. 2.4). We expect, therefore, that the infinite-N system will exhibit a jump bifurcation at $K = 2/\pi$ with a height (as measured by R) of $1/2$, but without a hysteresis loop.

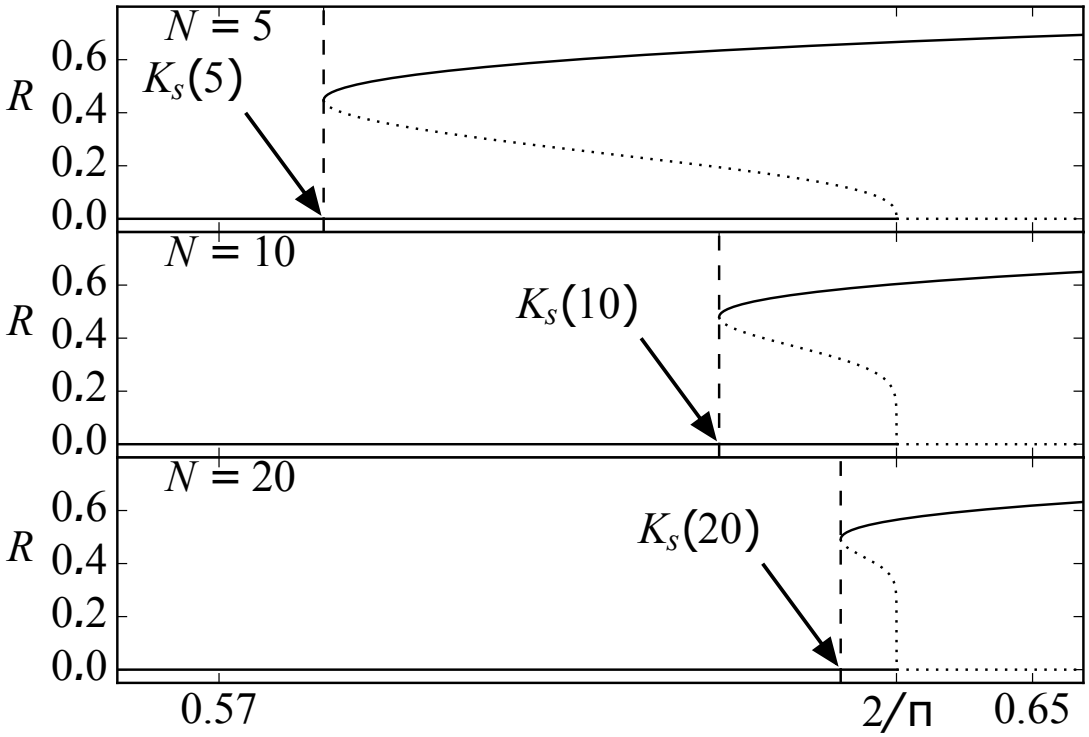

Figure 2.3. Bifurcation diagrams for the finite-N system showing the bistable region as it depends on N. Solid lines indicate stable fixed points; dotted lines, unstable. Data generated using AUTO software[70].

The situation is similar to that investigated by Pazó [57], who found the locking threshold for the (unforced) Kuramoto model with evenly spaced natural frequencies. In contrast with the typically considered case in which the density g of the natural frequency distribution has $g''(0) < 0$, leading to a continuous synchronization transition [66], the uniform distribution has $g''(0) = 0$, and the transition is discontinuous. Precise results for the height of the jump, R_c^{unf}, and the scaling of $R - R_c^{\text{unf}}$ for $K > K_c^{\text{unf}}$ were derived using a self-consistent approach[57].

Correspondingly, Pazó found, in the finite-N system, a phenomenon of global frequency alignment for K below the infinite-N critical point, K_c^{unf}. Specifically, it happens that as coupling strength is increased, oscillators with nearby frequencies lock to each other, forming clumps, which then merge as K is further increased. The final merge

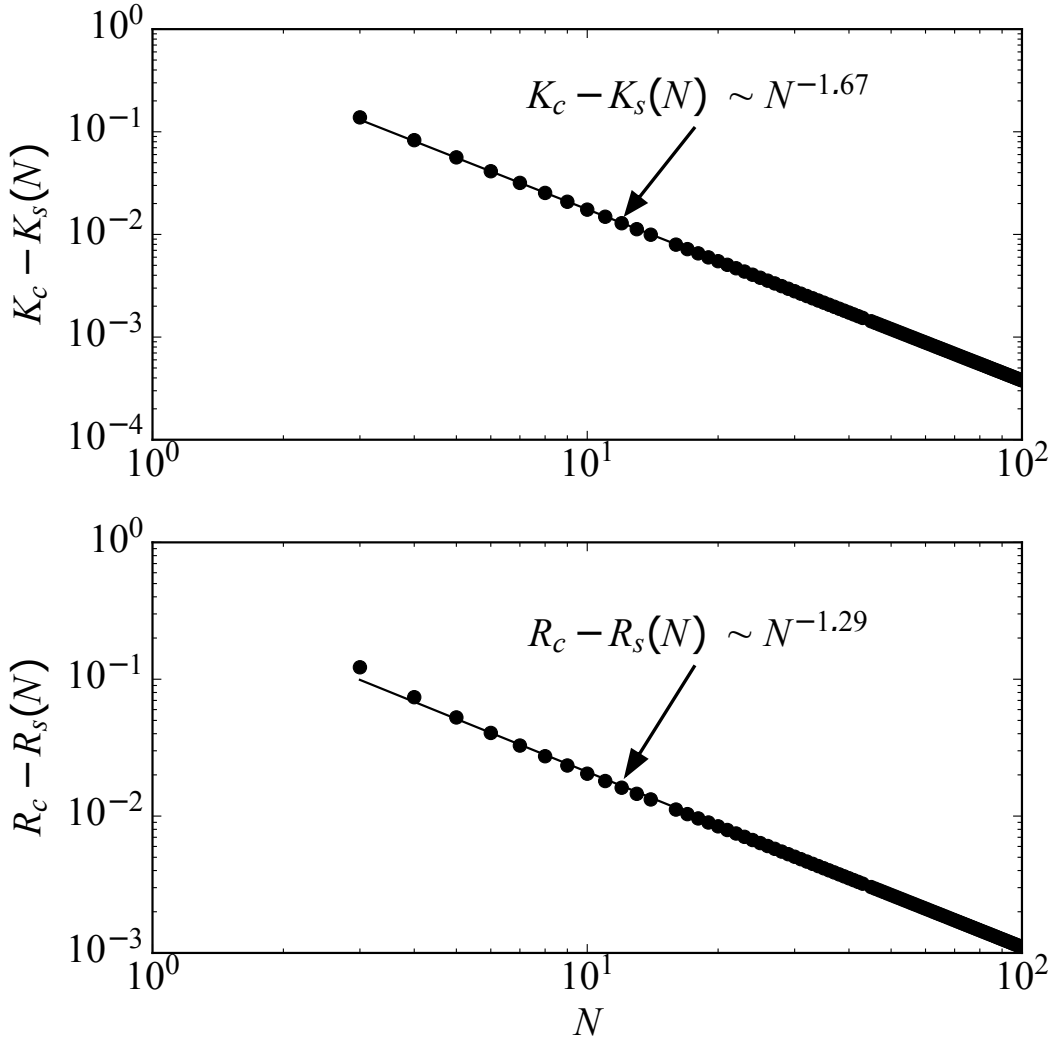

Figure 2.4. (upper) The extent of the stable $R > 0$ branch below $K_c = 2/\pi$ as a function of N. The data show approximately a power law scaling $N^{-1.67}$ for N between 3 and 100. (lower) The difference between the value of synchrony R at the saddle-node point and a numerically estimated critical value of $R_c = 1/2$ as a function of N. The data show approximately a power law scaling $N^{-1.29}$ for N between 3 and 100. Circles represent data measured from AUTO simulation, solid line is a power law fit.

occurs at $K = K_s(N)$, which approaches K_c^{unf} from below as $N \to \infty$, according to $K_c^{\text{unf}} - K_s(N) \sim N^{-\mu}$ with $\mu \approx 3/2$. We should note that for finite N, the transition in the unforced case is not hysteretic, as it is in the forced case.

2.7 Conclusions

We have explored, using an idealized model, the interplay between two ways in which a population of phase oscillators may be caused to behave coherently: common periodic forcing and attractive coupling. Based on the synchrony order parameter, R, forcing and coupling can appear to be at odds; the forcing drives R towards zero while the coupling drives R towards one. However, as we demonstrate both analytically and numerically, this view is inherently limited, since for K above $K_c = 2/\pi$, the forced system exhibits greater phase alignment than the corresponding unforced system. An intuitive explanation for this mismatch is that the parameter R measures only phase alignment, and is prone to miss the necessary precondition of frequency alignment.

Though we have gained considerable intuition from the results already obtained, there is more work to be done. First, we are still lacking analytical understanding of the upper branch of solutions, which would include an expression for the height of the jump and the scaling of R with K above the jump (see Fig. 2.3).

Another set of questions involves the (in)feasibility of the sawtooth interaction function Λ_v (defined as $\Lambda_v(\varphi) = -\varphi/\pi$ for $\varphi \in (-\pi, \pi]$ (2.16)). A simple argument reveals that for any integrable forcing waveform $v \in L^2(0, 2\pi)$, the corresponding interaction function Λ_v will be continuous on S^1, a condition which the sawtooth does not satisfy. It remains unexplored to what degree the results presented here may be approximated by interaction functions that approximate a sawtooth. One could investigate the scaling of dynamical properties with the energy of the input signal used.

Techniques for analysis and control of entrainment processes can be used to examine and even manipulate numerous processes in biology [68]. In addition to developing an initial mathematical framework for characterizing stability of coherence phase structures in a continuum of interacting oscillators, our work presents a potential path towards ad-

dressing a compelling biological application. Specifically, although some disagreement about the nature and phenomenology of epilepsy exists in the neuroscience literature [36], studies in animal models have indicated that control of synchronization of neural dynamics can mitigate epileptiform activity [31]. It is understood that neural stimulation is an underactuated system because one or a few electrodes are used to control the mean field of a very large collection of interacting neurons, which for practical purposes may be approximated by a continuum [16]. The ability to characterize the stability of phase decoherence in continuum models of general coupled oscillators could determine the possibility of developing effective desynchronizing stimuli for treatment of epilepsy. The criterion that is derived in Section 2.5.3 and validated by numerical experiments in Section 2.6 could in principle be tested experimentally [35].

2.8 The relationship to Previous Work

Finally, we discuss the relationship of the present model to previous work on models of globally coupled oscillators subject to common forcing. The existing literature has focused almost exclusively on sinusoidal forcing [63, 46, 52, 15, 4], owing to the analytical progress that this assumption allows. The two bodies of work most closely related to that presented above are due to Ott-Antonsen [52] and Mirollo-Strogatz [46].

2.8.1 Ott-Antonsen Ansatz

One model of a population of coupled oscillators subject to common forcing was discussed by Ott and Antonsen [52] as a possible application of the powerful dimension reduction known as the Ott-Antonsen (OA) ansatz. It is therefore natural to suppose that the system we consider here can be fully understood in that framework. However, while it is the case that the OA ansatz can describe the fixed point that we consider, the dynamics away from the fixed point do not leave the OA manifold invariant, precisely because the sawtooth forcing function we consider is not sinusoidal. We now demonstrate this fact.

The OA ansatz concerns a probability density formulation of the mean-field Kuramoto model in which the system is described by a probability density $f(\omega, \theta, t)$ that gives the probability that a randomly selected oscillator has natural frequency ω and phase θ at

time t. The density f then evolves according to a Fokker-Planck equation,

$$\partial_t f(\omega, \theta, t) + \partial_\theta \left(v(\omega, \theta, t) f(\omega, \theta, t) \right) = 0 \,, \tag{2.46}$$

where $v(\omega, \theta, t)$ gives the velocity of an oscillator having natural frequency ω and phase θ at time t (analogous to the right-hand side of (2.18)).

It is convenient to express f as a Fourier series in θ,

$$f(\omega, \theta, t) = \frac{g(\omega)}{2\pi} \left(1 + \sum_{n=1}^{\infty} f_n(\omega, t) \exp(\mathrm{i}n\theta) + \text{c.c.} \right) , \tag{2.47}$$

where c.c. denotes complex conjugate. The OA ansatz is the assumption that the Fourier coefficients $f_n(\omega, t)$ have the special form

$$f_n(\omega, t) = \left(\alpha(\omega, t) \right)^n \,, \tag{2.48}$$

for some *fixed*, independent of n, function α. For example, $\alpha \equiv 0$ describes a uniform distribution over phase for every natural frequency, and $|\alpha| \nearrow 1$ gives $f \to \delta(\theta - \psi) g(\omega)/2\pi$ [52].

The truly remarkable result of [52] is that the dynamics, under the mean-field Kuramoto model without external forcing, of a density f satisfying the OA ansatz can be reduced to dynamics of only the function α. This is seen by inserting the Fourier series for f into the dynamics (2.46), and separating terms by harmonics $\exp(\mathrm{i}n\theta)$.

A straightforward calculation shows that the n^{th} equation is

$$n\alpha^{n-1} \frac{\partial \alpha}{\partial t} + \mathrm{i}n\omega\alpha^n + \frac{K}{2} n\alpha^{n-1} \left(r\alpha^2 - r^* \right) = 0. \tag{2.49}$$

Every term contains a factor of $n\alpha^{n-1}$, and dividing it out leaves

$$\frac{\partial \alpha}{\partial t} + \frac{K}{2}(r\alpha^2 - r^*) - \mathrm{i}\omega\alpha = 0. \tag{2.50}$$

This single equation gives dynamics for α directly; the set of densities described by the OA ansatz is invariant under the dynamics of the Kuramoto model.

This result holds for the Kuramoto model without forcing, as well as the Kuramoto model with *sinusoidal* forcing. The reason for this is that in writing the equation satisfied

at an arbitrary harmonic $\exp(\mathrm{i}n\theta)$, a sinusoidal forcing function introduces terms from the $\exp(\mathrm{i}(n-1)\theta)$ and $\exp(\mathrm{i}(n+1)\theta)$ harmonics, but no further.

Now, we can consider our forced model in the same way. The forcing term $\Lambda_v(\theta) = -\theta/\pi, \theta \in [-\pi, \pi)$ appears in the Fokker-Planck equation as

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \theta}\left\{\left[\omega - \frac{\theta}{\pi} + \frac{K}{2i}\left(r\mathrm{e}^{\mathrm{i}\theta} - r^*\mathrm{e}^{-\mathrm{i}\theta}\right)\right] f\right\} = 0. \tag{2.51}$$

If we attempt the same procedure as before, inserting the ansatz (2.48) into (2.51), the equation at the n^{th} harmonic is (after dividing through by α^{n-1})

$$n\frac{\partial \alpha}{\partial t} + \mathrm{i}n\omega\alpha + \left(\frac{-1}{\pi} - \frac{in\theta}{\pi}\right)\alpha + \frac{K}{2}n\left(r\alpha^2 - r^*\right) = 0\,, \tag{2.52}$$

from which the n-dependence cannot be removed, precisely because of the $-1/\pi$ term that comes from the forcing function. In other words, there is no single equation for α whose truth guarantees that (2.48) gives a solution to (2.51). Hence, we *cannot* conclude that the set of distributions of the form (2.48) is invariant under the dynamics (2.51).

However, it does happen that the desynchronized fixed point, $f(\omega, \theta, t) = \delta(\theta - \pi\omega)$, fits the form (2.48), with $\alpha(\omega, t) = \exp(-i\pi\omega)$.

2.8.2 Mirollo-Strogatz random pinning model

Another system much more closely similar to ours is the "random pinning" model studied by Mirollo and Strogatz [46]. The random pinning model consists of a system of N spins, with each one pinned by an anonymous driving force to a particular (randomly chosen) phase. In explicit terms, the dynamics are

$$\dot{\varphi}_i = \sin(\alpha_i - \varphi_i) + \frac{K}{N}\sum_{j=1}^{N}\sin(\varphi_j - \varphi_i)\,, \tag{2.53}$$

where $\{\alpha_i\}$ are random quantities sampled from the uniform distribution on the unit circle. The only difference between this equation and the one that we study is the term $\omega_i - \varphi_i/\pi$ is replaced by $\sin(\alpha_i - \varphi_i)$. It remains the case that in the absence of coupling, each oscillator evolves according to an autonomous ODE on the unit circle with one stable fixed point, and that the state in which each oscillator is at its individual fixed point has $R \approx 0$.

The authors proceed to present a continuum formulation of the dynamics (2.53) that is of the same form as (2.18); where we represent phase as a function of natural frequency, they represent phase as a function of pinning phase α. Owing to the regularity of the sine function, it is possible to obtain precise analytical results on the existence, number, and stability of fixed points. Our formulation is not amenable to the same analysis, for the reason that the sawtooth forcing function we consider has infinitely many Fourier modes.

Chapter 3

Coarse-Graining for Coupled Oscillators

Numerical simulations of complex multi-scale phenomena are fundamental to modern science. We investigate the following question: given a simulation of a rhythmic dynamical process, does there exist a good lower-dimensional representation? If so, finding such a representation may enable accelerated computational simulation and provide fundamental insight into the dynamics of interest. In this study, we infer coarse-grained equations of motion that describe a heterogeneous population of oscillators with a modular coupling structure. We choose this system because it is known to exhibit a transition from high- to low-dimensional behavior, where the latter is well-described by equations of a known form. We explore phase transitions in clustering and model selection through numerical simulations, and evaluate generalizations by systematically discarding several of the simplifying assumptions involved.

3.1 Introduction

The Kuramoto model of coupled oscillators [42] has served since its introduction as a simple model of collective behavior. Many generalizations have been proposed, one class of which involves supposing that the oscillators are not coupled all-to-all or on a regular lattice, but according to a network with a nontrivial structure. Many results have been obtained that connect the coupling network's structure with the oscillators' dynamics,

either in steady state or during the transient leading up to it [7, 30, 5, 22]. A particularly intuitive and interesting class of results supports the notion that if the coupling network exhibits a strongly modular structure (see e.g. [49] for a discussion of modularity), then oscillators in the same community will tend to align with each other more readily than with the rest of the network [7, 6, 5, 22, 30]. This property makes the Kuramoto model an ideal example for investigating coarse-graining of nonlinear dynamics.

The study of the Kuramoto model on modular networks goes back at least to Lumer and Huberman [45], who considered coupling according to a hierarchical lattice. They found that depending on parameter settings, it is possible to observe a sequence of local synchronization transitions that nonetheless may not result in global synchronization in the infinite-size limit. Building on prior work of Daido in the setting of regular lattices [20], their analysis proceeds by replacing successively larger and larger clusters of oscillators by a single "imaginary giant cluster oscillator", reminiscent of renormalization group analysis [73]. A rigorous and thorough renormalization group analysis of Lumer and Huberman's model was recently performed by Garlaschelli and collaborators [28], and renormalization for coupled oscillators on a 1-D lattice has been investigated by Kogan et. al. [39].

More recent connecting modular network structure with the dynamics of synchronization is due to Oh et. al. [50], who studied the synchronization transition in a modified Kuramoto model on various empirical and synthetic modular network topologies, finding that the location and scaling properties of the transition depend on the manner in which modules are linked.

Another line of work connecting oscillator dynamics to the structure of the networks that couple them examines the case of *identical* oscillators [7, 6, 5, 22, 30]. In this case, there is a single global attractor, corresponding to complete synchronization (provided the coupling graph is connected). Rather than study this relatively uninteresting attractor, one can instead consider the transient. By measuring the correlation (over an ensemble of initial conditions) between all pairs of oscillators as a function of time, it is possible to track the manner in which oscillators come to synchrony. The oscillators begin uncorrelated, while in the end, all oscillators are synchronized and hence all perfectly correlated with

each other; therefore, time serves as a natural resolution parameter for decomposing the network into mutually-correlated groups.

Finally, the seminal work of Ott and Antonsen [52] demonstrates rigorously that a group-structured population of oscillators can be represented (in the $N \to \infty$ limit) by a closed ODE with one degree of freedom per group. This framework has been leveraged by Skardal and Restrepo [65] to study in detail the phenomena of synchronization within and between modules, finding several interesting bifurcations.

In this paper, we informally define coarse-graining as as the reduction of a detailed (micro-scale) dynamical system model to a lower-dimensional system that reproduces the overall (macro-scale) behavior of the original. We define explicit procedures for the inference of a set of mappings that provide a formal correspondence between each coarse-grained variable and a subset of detailed variables whose collective dynamics it is used to represent. Conceptually, this is similar to the aim of renormalization group methods in statistical physics [73]. Here, we seek a theoretically and computationally verifiable approach for dimensional reduction of Kuramoto oscillators with modular coupling to the natural coarse dynamical representation.

3.2 Background

The Kuramoto model is the ordinary differential equation (ODE) system

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^{N} K_{ij} \sin(\theta_j - \theta_i), \qquad i = 1 \dots N, \tag{3.1}$$

where $\theta_i \in S^1$ is the *phase* of the i^{th} oscillator, $\omega_i \in \mathbb{R}$ is its *natural frequency*, N is the total number of oscillators, and $K_{ij} \in \mathbb{R}_{\geq 0}$ is the *coupling matrix* that defines which oscillators influence each other.

We can equivalently formulate the Kuramoto model in terms of *complex phases* $y_i =$

$\exp(\mathrm{i}\theta_i)$. By the chain rule, we have

$$\begin{aligned}
\dot{y}_i &= \mathrm{i}y_i\dot{\theta}_i \\
&= \mathrm{i}y_i\left(\omega_i + \sum_{j=1}^{N} K_{ij}\sin(\theta_j - \theta_i)\right) \\
&= \mathrm{i}y_i\left(\omega_i + \sum_{j=1}^{N} K_{ij}\frac{1}{2\mathrm{i}}\left(e^{\mathrm{i}(\theta_j-\theta_i)} - e^{-\mathrm{i}(\theta_j-\theta_i)}\right)\right) \\
&= \mathrm{i}y_i\left(\omega_i + \sum_{j=1}^{N} K_{ij}\frac{1}{2\mathrm{i}}\left(y_j y_i^* - y_j^* y_i\right)\right) \\
&= \mathrm{i}\omega_i y_i + \frac{1}{2}\sum_{j=1}^{N} K_{ij}(y_j - y_j^* y_i^2) \qquad (3.2)
\end{aligned}$$

where we have used the fact that $y_i^* y_i = 1$.

Perhaps the best-studied case of the Kuramoto model is that of *mean-field* coupling, where $K_{ij} = K/N$ for all i, j. In this case it is well known that in the limit $N \to \infty$, the system (3.1) exhibits a phase transition with respect to K: there exists K_c such that for $K < K_c$, the oscillators behave mostly independently, while for $K > K_c$ a subset of oscillators spontaneously lock to a single frequency.

Synchronization in the Kuramoto model is typically quantified by the order parameter,

$$z := \frac{1}{N}\sum_{j=1}^{N} \mathrm{e}^{\mathrm{i}\theta_j} = R\mathrm{e}^{\mathrm{i}\Theta}, \qquad (3.3)$$

where $R \in [0, 1]$ is the *synchrony* and $\Theta \in [0, 2\pi)$ is the *average phase*. If all phases are equal then $R = 1$, and if the phases are spread evenly over the unit circle, then $R \approx 0$. Thus R is a natural measure of synchronization.

The dynamics of z depend on the dynamics of all θ_i, but it is natural to suppose that in some limit there exists a *closed* equation for the dynamics of z. Indeed there is, as demonstrated by Ott and Antonsen [52]. Assuming that w_i are Cauchy, distributed, i.e., $\omega_i \sim g(\omega)$ where g is the Cauchy probability density function with mode 0 and width 1, and that the coupling is mean-field, then the evolution of z is described in terms of R and

Θ as

$$\frac{\mathrm{d}R}{\mathrm{d}t} + \left(1 - \frac{K}{2}\right) R + \frac{K}{2} R^3 = 0, \tag{3.4}$$

$$\frac{\mathrm{d}\Theta}{\mathrm{d}t} = 0. \tag{3.5}$$

The form (3.4) shows clearly that $R = 0$ is always a solution, but undergoes a pitchfork bifurcation at $K_c = 2$, when a new solution $R = \sqrt{1 - 2/K}$ appears, representing partial synchrony that becomes global synchrony (i.e. $R \to 1$) as $K \to \infty$.

Interestingly, the same analysis carries over to the case where oscillators are not coupled all-to-all, but are divided into subsets such that the strength of coupling between any two oscillators depends on the subsets to which they belong. Following [65, 10], we represent such a modular system as

$$\dot{\theta}_i^\sigma = \omega_i^\sigma + \sum_{\sigma'=1}^{C} \frac{N_\sigma}{N} \frac{K^{\sigma\sigma'}}{N_{\sigma'}} \sum_{j=1}^{N_{\sigma'}} \sin(\theta_j^{\sigma'} - \theta_i^\sigma) \tag{3.6}$$

where module σ (of C total) is of size N_σ, and $N = \sum_\sigma N_\sigma$ denotes the total number of oscillators. If oscillators in the same module are coupled much more strongly than those in different modules, we may expect that modules synchronize internally and the system is describable at a larger scale. Generalizing from the mean-field case, we suppose that the appropriate collective variables are the *cluster order parameters* $\{z_\sigma\}$, defined by

$$z_\sigma = \frac{1}{N_\sigma} \sum_{i=1}^{N_\sigma} e^{\mathrm{i}\theta_i^\sigma}. \tag{3.7}$$

If $\{\omega_i^\sigma\}$ are distributed according to a Cauchy distribution with mode Ω_σ and width δ_σ, then the Ott-Antonsen ansatz predicts that the cluster order parameters will obey the system of C complex ODEs

$$\dot{z}_\sigma = \mathrm{i}(\Omega_\sigma + \mathrm{i}\delta_\sigma) z_\sigma + \frac{1}{2} \sum_{\sigma'=1}^{C} \frac{N_{\sigma'}}{N} K^{\sigma\sigma'} \left(z_{\sigma'} - z_{\sigma'}^* z_\sigma^2\right) \tag{3.8}$$

which is mathematically of the same form as the original Kuramoto model written in terms of complex phases (3.2). The consequences of this equation for the existence of mesoscale synchronization were discussed in detail in [65]. For now, we note that (3.8)

reduces exactly to the Kuramoto model in the case that $\delta_\sigma = 0$ and $|z_\sigma| = 1$ for all σ. In this sense, this natural emergence of (3.8) implies that, in the Kuramoto model, groups of oscillators can behave collectively as a single oscillator. In other words, there is a renormalization procedure for coupled oscillator systems that remains within the same model class (in particular, the one defined by (3.8)).

Here we take a complementary approach, and develop a process to infer the parameters of the equation (3.8) given time-series measurements of the system (3.6), so that the resulting system of form (3.8) can accurately reproduce the dynamics of the cluster order parameters of (3.6).

The question of existence of a good coarse-grained model is nuanced, so we now unpack it into two related, but distinct, questions.

The first question is, does the value of $z \in \mathbb{C}^C$ exactly determine the value of $\dot{z}$? And if so, what is the functional relationship $\dot{z} = f(z)$? One straightforward way to address this question is to attempt to fit a function to the data $\{(z(t), \dot{z}(t) | t \in [0, T]\}$. In general this is an infinite-dimensional optimization problem, but can be restricted to a finite-dimensional problem by choosing a finite set of basis functions. If we succeed in finding f such that $\|\dot{z} - f(z)\|$ is small, then we conclude that the dynamics of z (i.e. $\dot{z}$) can be accurately computed from only the value of z.

Assuming that answering the first question gives us a functional relationship $\dot{z} = f(z)$, the next question is: how well is it possible to predict the future of z based on its present value? This question is much harder to answer in general, because even if the dynamics of z are perfectly deterministic, they may be highly sensitive to initial conditions and hence practically impossible to forecast beyond a short time. We therefore make a distinction between *training error*, the quality of the fit $\|\dot{z} - f(z)\|$, and *prediction error*, $\|z - \hat{z}\|$ where $\mathrm{d}\hat{z}/\mathrm{d}t = f(\hat{z})$ and $\hat{z}(t_0) = z(t_0)$. Because prediction error is much harder to bound in general, we focus first on the training error as an indicator of the existence of a coarse-grained model.

Finally, notice that we have implicitly assumed so far that the oscillators are partitioned by the coupling network, and that this partition is the appropriate one to consider

when seeking a coarse-grained model. That is, each oscillator is identified by two indices, σ giving its cluster and i giving its position within the cluster. However, we may want to relax this assumption and allow the data to tell us how to partition the oscillators. To this end we introduce some notation.

Let $\{\theta_i | i = 1, \ldots, N\}$ be a collection of phases that evolve in time according to (3.1), and let $P = (P_1, \ldots, P_C)$ be a *partition* of $\{1, \ldots, N\}$. Formally,

$$\bigcup_{\sigma=1}^{C} P_\sigma = \{1, \ldots, N\} \text{ and } P_\sigma \cap P_{\sigma'} = \emptyset \quad \forall \sigma \neq \sigma'. \tag{3.9}$$

In words, P is a set of subsets of nodes that are mutually disjoint and cover the whole network. We can then define the coarse-grained variables according to this partition by

$$z_\sigma = \frac{1}{|P_\sigma|} \sum_{i \in P_\sigma} \mathrm{e}^{\mathrm{i}\theta_i} \tag{3.10}$$

or, using angle brackets as shorthand for averaging, $z_\sigma = \langle \exp(\mathrm{i}\theta_i) \rangle_{i \in P_\sigma}$.

Note that we can recover (3.6) by setting

$$P_\sigma = \left\{ i \in \{1, \ldots, N\} \middle| \sum_{\sigma' < \sigma} N_{\sigma'} < i \leq N - \sum_{\sigma' > \sigma} N_{\sigma'} \right\} \tag{3.11}$$

and

$$K_{ij} = \frac{N_\sigma}{N} \frac{K^{\sigma\sigma'}}{N_{\sigma'}} \qquad \forall i \in P_\sigma, j \in P_{\sigma'} \tag{3.12}$$

3.3 Inference of Modular Parameters

We now describe a procedure for inferring the parameters of the coarse-grained model (3.8) assuming that the modular structure of (3.6) is known. Given a solution $\{\theta_i^\sigma(t)\}$ of the system (3.6), we can obtain a coarse-grained time series $\{z_\sigma(t)\}$ according to (3.7). If these coarse-grained variables evolve according to (3.8), then it should be possible to recover the effective natural frequencies and coupling parameters using a least-squares fit. Observe first that the right-hand side of (3.8) is a linear combination of terms that can be measured directly from data. To clarify this, we re-write (3.8) as

$$\dot{z}_\sigma = \tilde{\omega}_\sigma z_\sigma + \sum_{\sigma'=1}^{C} B_{\sigma\sigma'} \left(z_{\sigma'} - z_{\sigma'}^* z_\sigma^2 \right), \tag{3.13}$$

where $\tilde{\omega}_\sigma = \mathrm{i}(\Omega_\sigma + \mathrm{i}\delta_\sigma)$ and $B_{\sigma\sigma'} = N_{\sigma'} K^{\sigma\sigma'}/2N$.

Let $\boldsymbol{z}_\sigma = (z_\sigma(t_1), \ldots, z_\sigma(t_n))^T$ and $\dot{\boldsymbol{z}}_\sigma = (\dot{z}_\sigma(t_1), \ldots, \dot{z}_\sigma(t_n))^T$ denote the time-series of observations of the coarse-grained variable z_σ and its derivative, respectively. Then the parameters $\boldsymbol{p}_\sigma = (\tilde{\omega}_\sigma, B_{\sigma 1}, \ldots, B_{\sigma C})^T$ can be inferred by solving the linear system $\dot{\boldsymbol{z}}_\sigma = \boldsymbol{G}^\sigma \boldsymbol{p}_\sigma$, where the $n \times C + 1$ matrix $\boldsymbol{G}^\sigma$ contains the values of the terms of which the entries of $\boldsymbol{p}_\sigma$ are coefficients. Explicitly, we can write $\boldsymbol{G}^\sigma$ as

$$\boldsymbol{G}^\sigma = \left[\begin{array}{c|c|c|c} \boldsymbol{z}_\sigma & \boldsymbol{z}_1 - \boldsymbol{z}_1^* \boldsymbol{z}_\sigma^2 & \ldots & \boldsymbol{z}_C - \boldsymbol{z}_C^* \boldsymbol{z}_\sigma^2 \end{array} \right] \tag{3.14}$$

where complex conjugate, product, and squaring of the vectors $\boldsymbol{z}_\sigma$ are understood to be taken element-wise.

Many more time-series observations are required than the number of modules in the model ($n \gg C$), so the linear system $\dot{\boldsymbol{z}}_\sigma = \boldsymbol{G}^\sigma \boldsymbol{p}_\sigma$ will be overdetermined. Minimizing the squared residual of this system for each $\sigma = 1, \ldots, C$ gives a complete set of parameters for the ODE (3.13). Given those parameters, we can compute the training fit $\|\dot{z} - f(z)\|$, where f is the right-hand side of (3.13).

We can further compute a prediction error. Assuming we have access to the true trajectory for $t > t_n$ (if, say, we only use part of the trajectory for training), then we can use that true trajectory as a basis for comparison. To generate a prediction, we simply integrate the ODE $\mathrm{d}\hat{z}/\mathrm{d}t = f(\hat{z})$, where again f is the right-hand side of (3.13) with inferred parameter values, with the initial condition $\hat{z}(0) = z(t_n)$. The overall procedure – coarse-graining, inference, and prediction – is illustrated in Figure 3.1. Based on the results of [52, 65], we expect to see a good fit if N (the number of oscillators) is very large and the coupling is as described in (3.6). We wish to understand the suitability of (3.13) as a coarse-grained model for finite N.

3.4 Numerical Experiments

To make the situation concrete, we now present results of numerical experiments on a system of $N = 15$ oscillators, organized into three modules of five nodes each. For each oscillator we sample a natural frequency from a standard normal distribution (i.i.d.), and fix them through all experiments. The coupling between nodes is such that each pair of

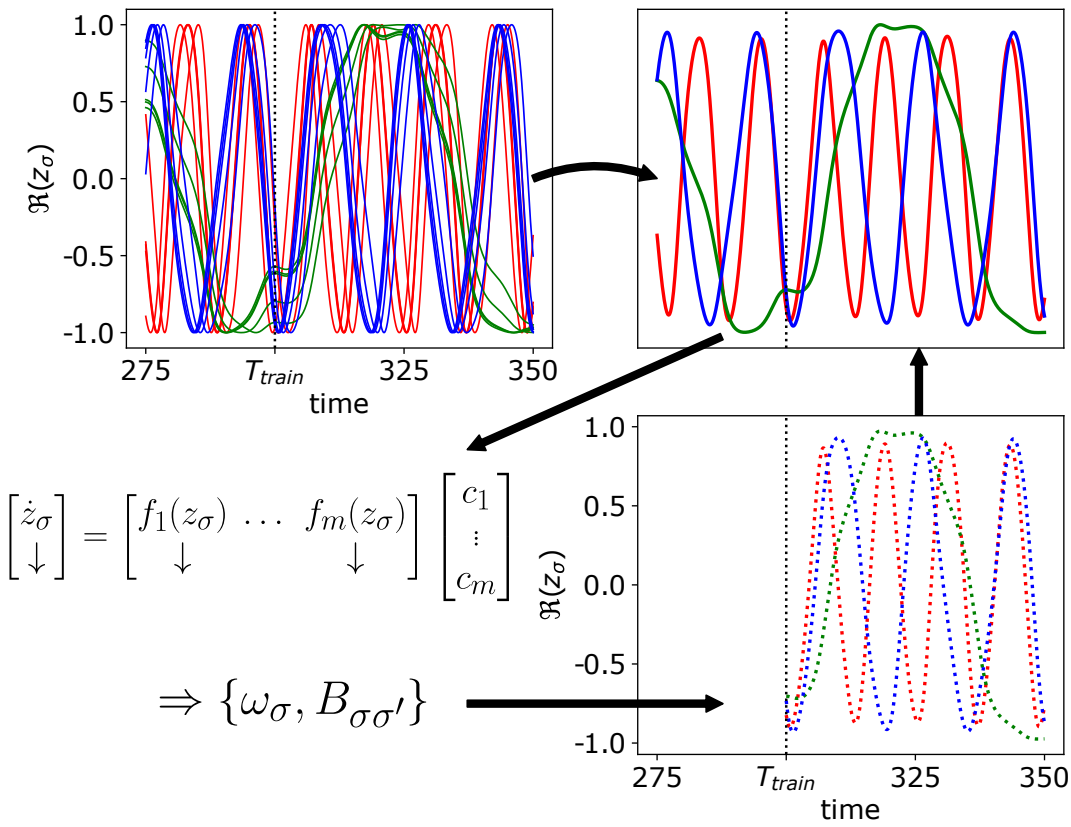

Figure 3.1. Schematic depicting the inference and validation procedure described in the text. At top left are the time series of each of fifteen oscillators, color coded by which of the three modules they belong to. At top right, we see the time series of the corresponding cluster order parameters. The time series of cluster order parameters are then used to construct a least-squares problem (bottom left) whose solution gives a set of parameters defining an instance of (3.13). Finally, we can use the inferred ODE to predict the future evolution of the cluster order parameters for $t > T_{\text{train}}$.

nodes in the same module is coupled with strength $K_{ij} = K_{\text{in}}/N$ and pairs of nodes in different modules are coupled with strength $K_{ij} = K_{\text{out}}/N$. We then examine the behavior of the system for $K_{\text{in}} \in [0, 10]$ and $K_{\text{out}} \in [0, 2.5]$.

First, we can check whether or not the modules are internally synchronized. To do this we compute the cluster order parameters z_σ according to (3.7), and compute its magnitude, averaged over time after an initial relaxation period. We also compute the

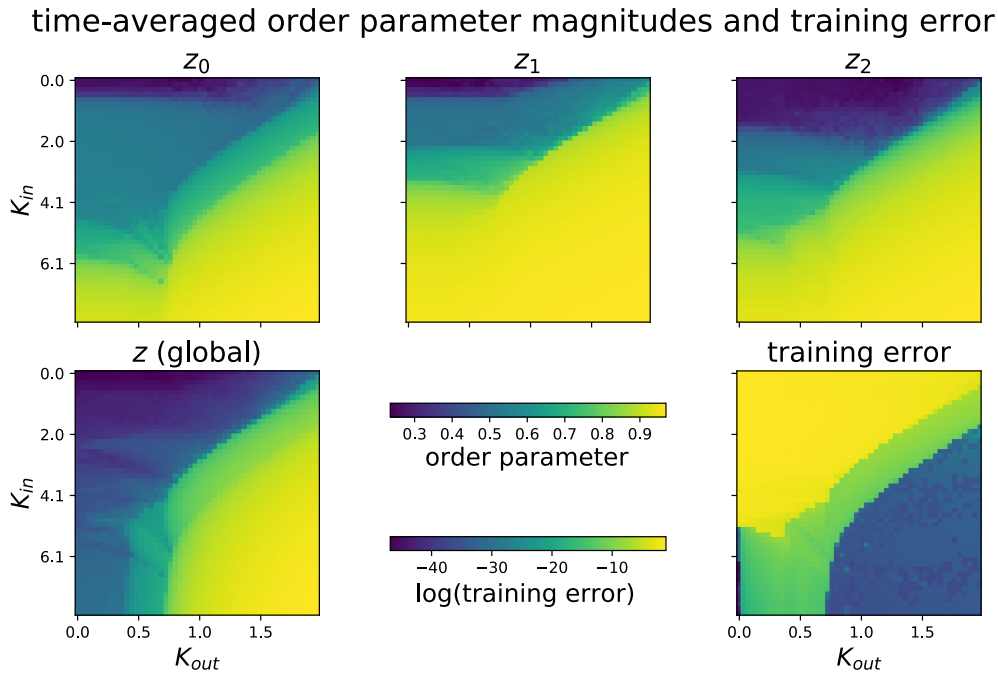

Figure 3.2. An overview of the behavior of the fifteen-oscillator system described in the text as a function of coupling parameters $(K_{\text{in}}, K_{\text{out}})$. Top row: magnitude of the order parameter within each module. Bottom left: magnitude of the global order parameter. Bottom right: natural logarithm of the training error, i.e. the difference between the observed value of $\dot{z}$ and the inferred right-hand side.

time-averaged magnitude of the global order parameter (3.3)in the same way. Unsurprisingly, we find that if K_{in} and K_{out} are both large enough, the whole system is synchronized, as indicated by a large value of the global order parameter. More interestingly, we also find that for some values of $(K_{\text{in}}, K_{\text{out}})$, each module is internally synchronized, while the system as a whole is not. Finally there are other cases where not all modules are internally synchronized, including the low-coupling regime where the oscillators move essentially independently. The situation is summarized in Figure 3.2.

It is reasonable to suppose that if each module is internally synchronized, then there should be a good model of the dynamics at the level of one (complex) degree of freedom per module. To test this hypothesis, we take the time series $\{\theta_i(t) | i = 1, \ldots, 15\}$ and coarse-grain them according to the structural partition of the network,

$$P = \{(1, \ldots, 5), (6, \ldots, 10), (11, \ldots, 15)\},$$

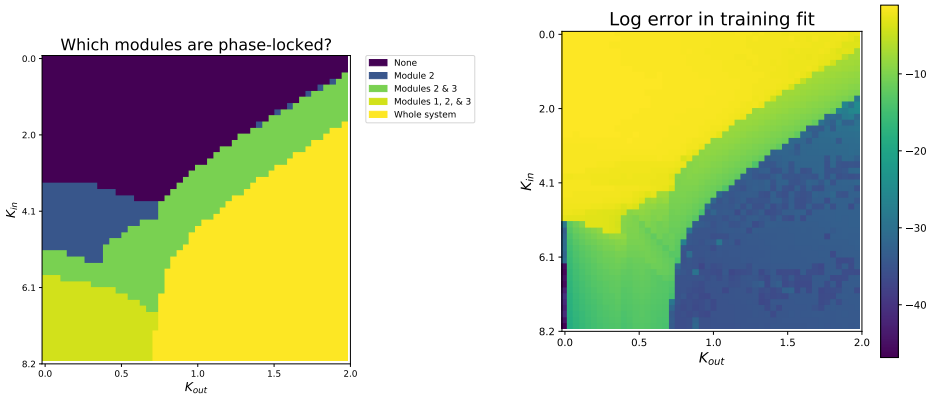

Figure 3.3. Heatmaps showing the relationship between local phase-locking and training fit of a coarse-grained model. Left: diagram in the $(K_{\text{in}}, K_{\text{out}})$ plane of which modules are internally phase-locked. Right: natural logarithm of the training error, i.e. the difference between the observed value of $\dot{z}$ and the inferred right-hand side.

obtaining the coarse-grained time series $\{z_\sigma(t)|\sigma = 1, 2, 3\}$. From these coarse-grained time series we compute a centered-difference approximation of the derivative, $\dot{z}_\sigma(t) = (z_\sigma(t + \Delta t) - z_\sigma(t - \Delta t))/2\Delta t$. We then perform the fitting procedure described above to find parameters $\tilde{\omega}_\sigma, B_{\sigma\sigma'}$ that most nearly satisfy the equation (3.13). The residual (in logarithm) is displayed in the bottom right of Figure 3.2 as a function of $(K_{\text{in}}, K_{\text{out}})$, and repeated at the right of Figure 3.3.

Interestingly, the heatmap of training error as a function of $(K_{\text{in}}, K_{\text{out}})$ reveals a strong separation of parameter space into distinct regions. Notably, the separation of parameter space appears much more discrete in terms of training error than in terms of the magnitude of the order parameters. This discrepancy suggests that the critical condition determining the quality of fit is discrete in nature; for instance, whether or not the modules are internally phase-locked. Formally, we use the notion of *phase cohesiveness*, following [24]. Let the set $\overline{\text{Arc}}_n(\gamma) \subseteq \mathbb{T}^n$ be defined as the closed set of angle arrays $(\theta_1, \ldots, \theta_n)$ such that all θ_i are contained within an arc of angle γ. Given a partition P, we say that the set of phases $(\theta_i(t))_{i \in P_\sigma}$ in module σ is *phase cohesive* if it is in $\overline{\text{Arc}}_{|P_\sigma|}(\gamma)$ for some $\gamma < 2\pi$, for all t (possibly after discarding an initial transient). The left panel of Figure 3.3 shows, as a function of $(K_{\text{in}}, K_{\text{out}})$, which modules are phase cohesive and whether the whole system is phase cohesive, with angle $\gamma = \pi$.

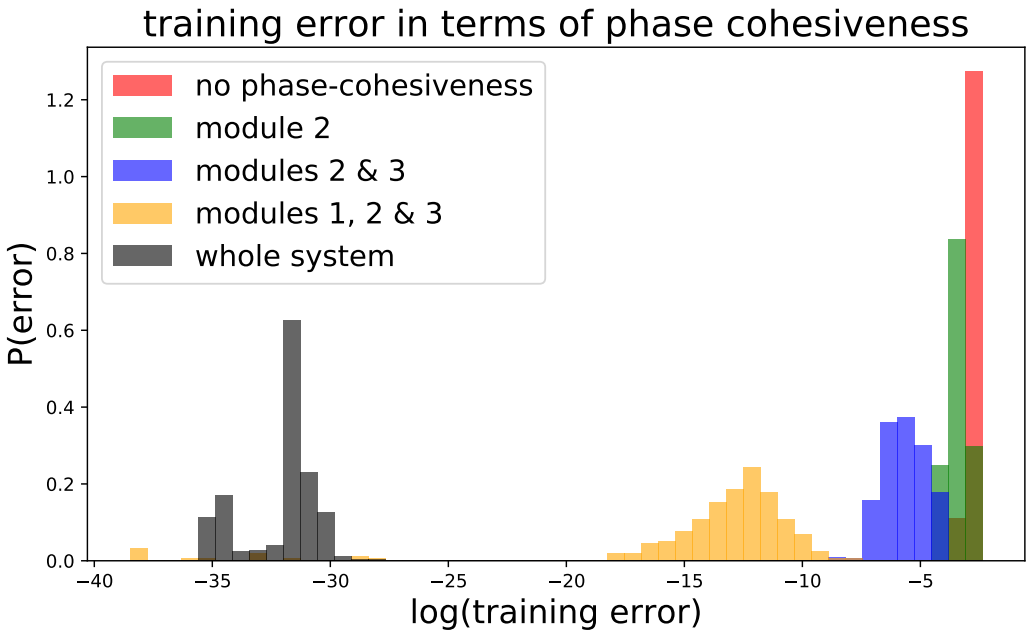

Figure 3.4. Histogram of training errors, coded by which modules are phase cohesive. For each of the 50×50 numerical experiments, we determine whether or not each of the three modules is phase cohesive with angle π. We observe five different cases: i) none of the modules is phase cohesive, ii) module 2 is phase cohesive, iii) modules 2 and 3 are phase cohesive, iv) modules 1, 2, and 3 are all phase cohesive, and v) the whole system is phase cohesive. The distribution of training errors for each of these five cases is shown here in a different color.

First, we see that it is exactly the cases when the whole system is phase cohesive that training error is especially small, on the order of $\exp(-35)$. Conversely, we find a poor fit (large training error, on the order $\exp(-1)$) for cases where none of the modules is not phase cohesive. Finally, we observe quite a good fit for cases where each of the three modules is phase cohesive, and a slightly poorer fit for cases where modules two and three are phase cohesive while module one is not. A detailed breakdown is given in Figure 3.4, where we visualize the distribution of training error as a function of which modules are phase cohesive. This observation suggests that phase cohesiveness is the critical condition that determines whether or not a given partition of oscillators supports a coarse-grained model.

3.5 Finding a good partition

Next we describe a method to use phase cohesiveness to find a partition of the nodes of an oscillator network such that the corresponding coarse-grained variables evolve according to an equation of the form (3.8).

Given a trajectory $\{\theta_i(t)\}$ of the ODE (3.1), let $P^{\text{dyn}} = (P_1, \ldots, P_C)$ be the coarsest partition such that $(\theta_i(t))_{i \in P_\sigma}$ are phase cohesive for all σ and all t (possibly after removing an initial transient). Then, given P^{dyn}, we coarse-grain according to (3.10) to obtain the coarse-grained trajectory $\{z_\sigma(t)\}$.

In practice, it is prohibitively expensive to enumerate all partitions of N nodes to find P^{dyn} that satisfies the condition described above, so we use a heuristic. Notice that if two oscillators are in the same phase cohesive cluster, then their average frequency will be nearly equal over a long time. So given a trajectory $\{\theta_i(t)\}$, we compute a long-term average frequency $\overline{\omega}_i = (\theta_i(t_0 + T) - \theta_i(t_0))/T$, and cluster nodes initially by their value of $\overline{\omega}$. We can then check if the resulting clusters are in fact phase cohesive, and if they are not, perform a finer clustering according to $\overline{\omega}$ and repeat until we have phase cohesive clusters.

For each of the 50×50 numerical experiments on the fifteen-node system, we perform the dynamical clustering procedure described above, coarse-grain the system according to the resulting partition, and seek an ODE model of the form (3.13) to describe the dynamics of the coarse-grained variables. A comparison of the training error obtained using the dynamical partition to that using the structural partition is presented in Figure 3.5. For the structural partition, we have already noted that there are many cases where there is no good ODE describing the coarse-grained variables, hence a tall peak at the far right. This peak is eliminated when we use dynamical partitioning, indicating that we can always find a moderately well-fitting coarse grained model if we allow the partition to adapt to the data.

To further investigate the strengths and limitations of coarse-graining according to a dynamically determined partition, we can break down the cases in terms of properties of the partition. Our primary conclusion is that the worst-fitting cases are those in which

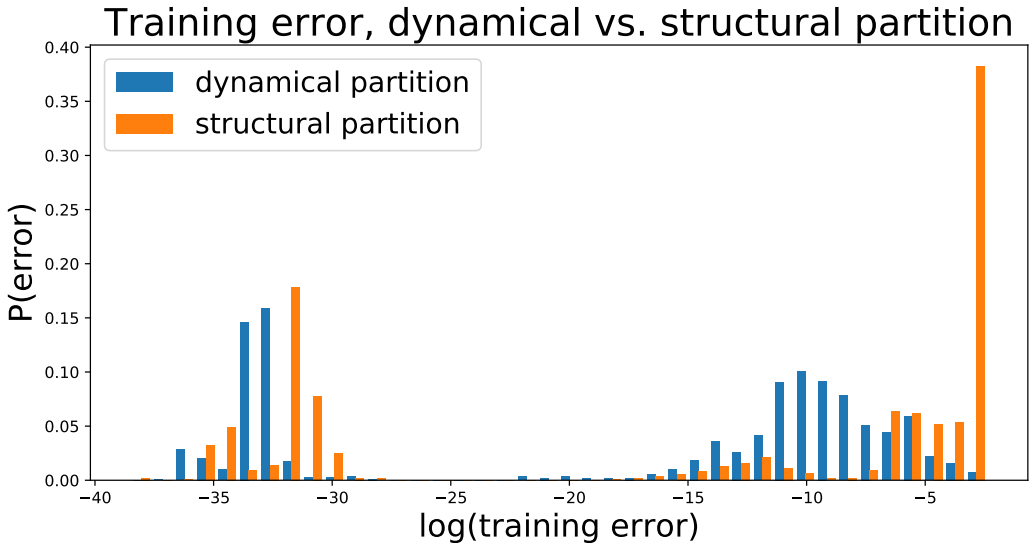

Figure 3.5. Comparison of training error for the coarse-grained model obtained by coarse graining according to the structural partition (three modules of five nodes each, orange) to the dynamical partition (variable number of modules, found as the coarsest possible partitioning into phase cohesive groups, blue).

the system splits into two clusters, one of which is a singleton. The situation is shown in Figure 3.6.

To see in detail what about such a case makes a coarse model not fit correctly, we select the case with the worst performance of all, which corresponds to parameters $K_{\text{in}} = 6.73$, $K_{\text{out}} = 0.92$. In this case we find that the system decomposes into two clusters, one containing a single oscillator. To visualize the actual performance of the fitting procedure, we plot the actual time derivative of each coarse-grained variable alongside the inferred right-hand side $f(z)$. The result is shown in Figure 3.7. We can intuitively understand the failure of this fit as follows. The values of K_{in} and K_{out} clearly indicate that every oscillator is much more strongly affected by other oscillators in its own module than by oscillators in different modules. Hence we know that the singleton oscillator is much more strongly affected by four of the fourteen phase cohesive oscillators than by the remaining ten. This means that its dynamics may be affected by the other group in ways that are not appropriately summarized by the simple average of their phases.

In other words, the coarse-grained model may fail to work because the coupling be-

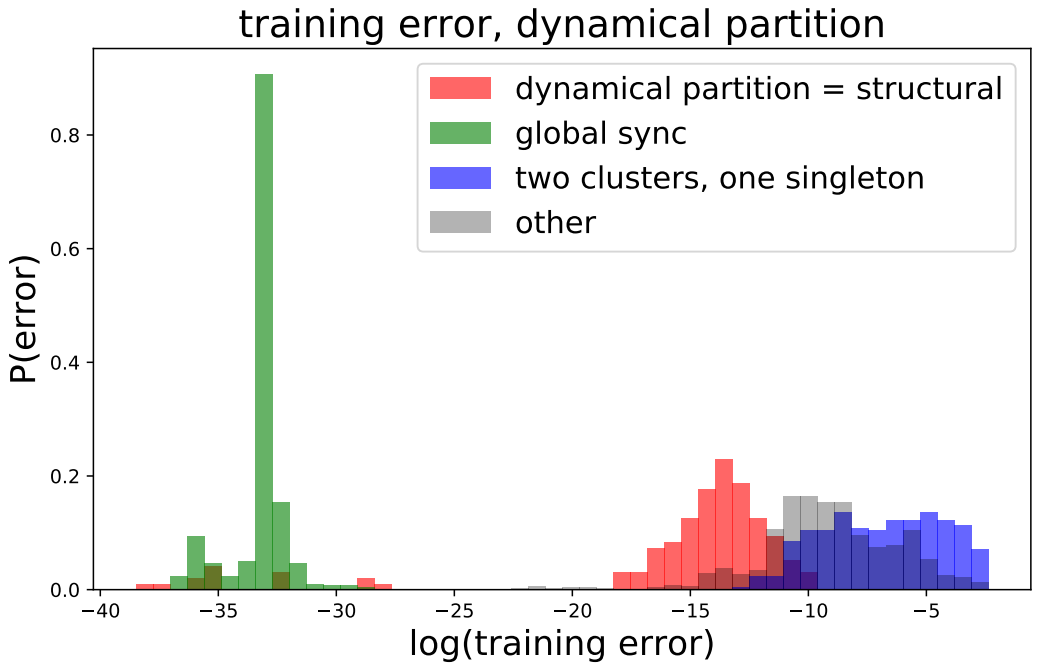

Figure 3.6. Training fit for a coarse-grained model obtained according to the phase cohesive partition, coded by properties of the partition. In blue are cases where the system decomposes into two clusters, one of which contains only a single node. We see clearly that these cases comprise the most extreme errors.

tween oscillators may not directly determined by their cluster membership. To see if this is actually the case, we consider the ensemble of dynamics-based partitions in a different way. For each partition obtained by imposing phase cohesiveness, we determine whether or not it is a refinement of the structural partition $\{(1,\ldots,5),(6,\ldots,10),(11,\ldots,15)\}$; that is, whether or not each phase cohesive cluster is a subset of one of the structural modules. Breaking the cases down in this way gives us Figure 3.8.

Finally, we consider the worst case scenario among the cases where the dynamical partition refines the structural partition. That case is the one where $K_{\text{in}} = 5.92$ and $K_{\text{out}} = 0.26$, and the system decomposes into five modules. The actual and fitted right-hand sides are shown in Figure 3.9. Notice the fit is much better than in Figure 3.7.

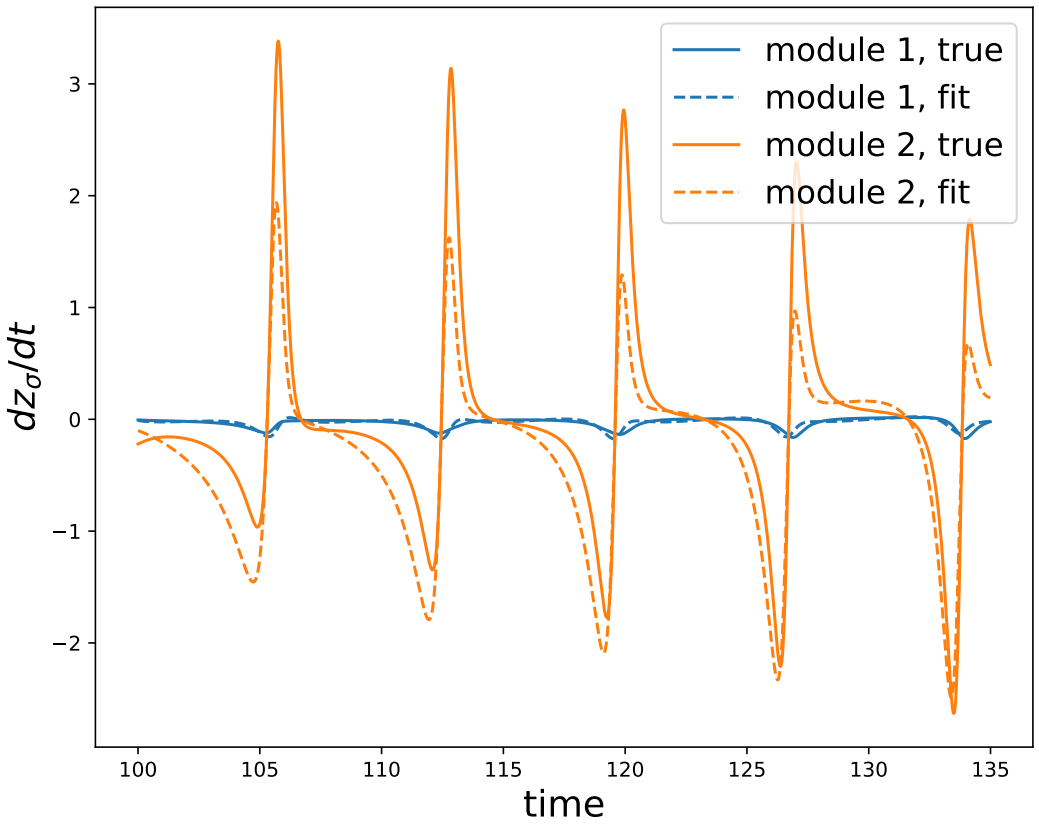

Figure 3.7. Visualization of training fit in the worst case observed in the 50×50 numerical experiments on the fifteen-node network. In this case one cluster consists of a single oscillator. In orange is the time derivative of the coarse-grained variable corresponding to the singleton cluster, and in blue is the time derivative of the coarse-grained variable corresponding to the 14-node cluster. Solid lines are ground truth, and dashed lines are the inferred right-hand side evaluated on the coarse-grained time series. The large difference in the sizes of the clusters induces a separation of scales in the derivative, which we can clearly see.

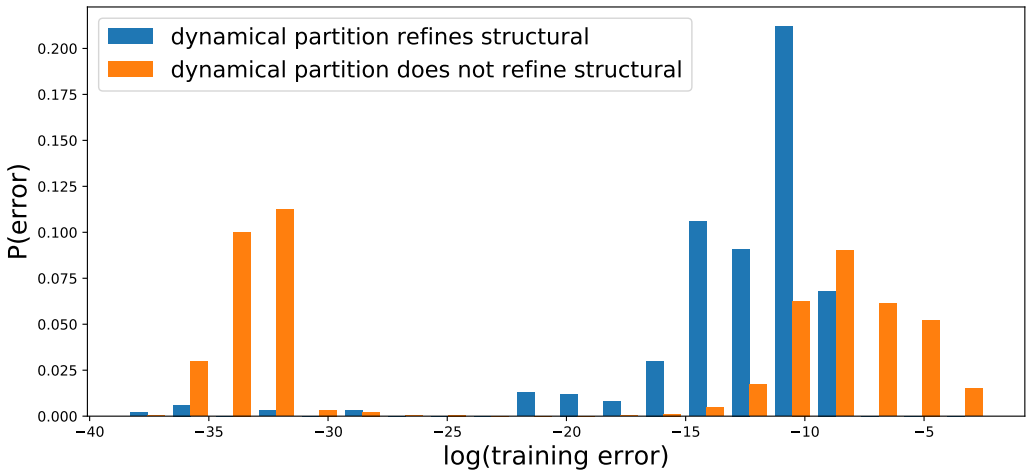

Figure 3.8. Histogram of training errors for the coarse-grained model corresponding to the dynamics-based partition, coded by whether or not the resulting partition refines the structural one. Notice that the maximum error in the case that the dynamical partition refines the structural one is far below the overall maximum.

3.6 Prediction

Here we evaluate the coarse-grained ODEs obtained as described above in terms of their ability to predict future evolution of the true system. As mentioned earlier, the problem of prediction is nontrivial even if we do know the exact dynamics that govern a system, because those dynamics may be highly sensitive to initial conditions. This issue is only amplified if the degrees of freedom evolve only approximately according to the inferred dynamics.

Give a trajectory $\{\theta_i(t)\}$ for $t \in [0, T_f]$ of the fine-grained system (3.6), and a partition P of the nodes, we can apply (3.10) to obtain the coarse-grained trajectory $\{z_\sigma(t)\}$ for each module σ. We then use the least-squares fit described in Section 3.3 to infer parameters $\{\omega_\sigma\}$, $\{B_{\sigma\sigma'}\}$ based on an initial portion of the coarse-grained trajectory, $\{z_\sigma(t)\}$ for $t \in [0, T_{\text{train}}]$. Using the inferred parameters and the initial condition $\{z_\sigma(T_{\text{train}})\}$, we integrate (3.8) to obtain a predicted trajectory $\{\hat{z}_\sigma(t)\}$ for $t \in [T_{\text{train}}, T_f]$. We can then compare the predicted trajectory $\{\hat{z}_\sigma(t)\}$ with the coarse-grained trajectory $\{z_\sigma(t)\}$ for $t \in [T_{\text{train}}, T_f]$, as shown in Figure 3.10.

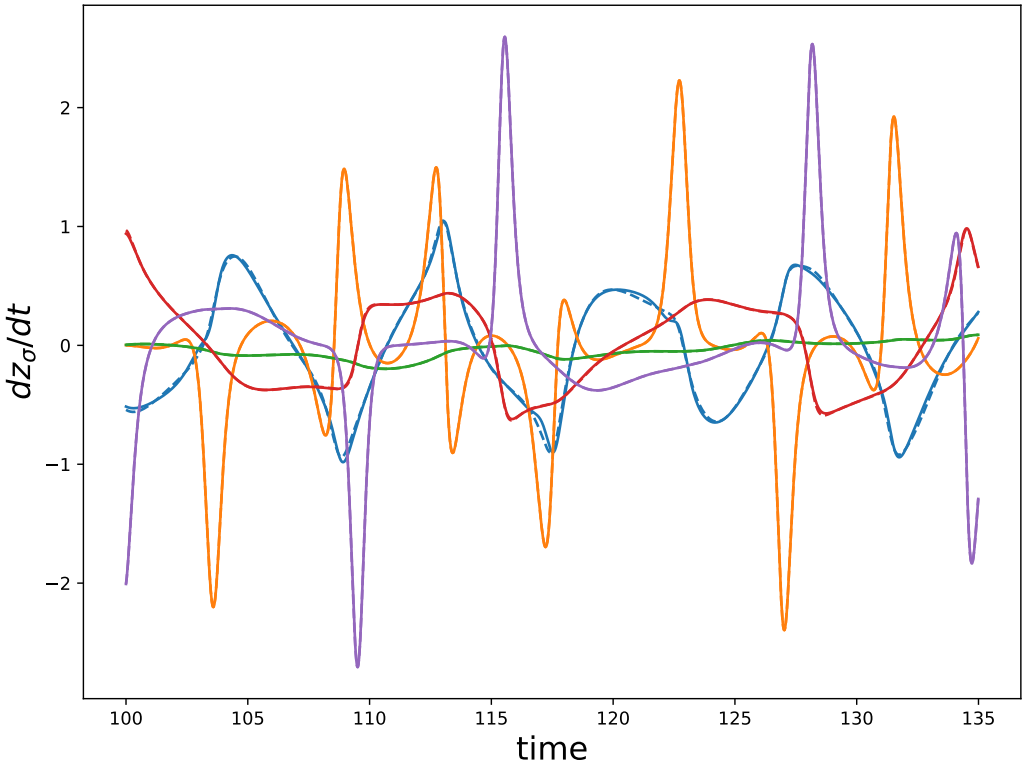

Figure 3.9. The worst case scenario for fitting a coarse-grained right-hand side in the case where the phase cohesive partition is a refinement of the structural partition.

3.7 Discussion

In this study, we present numerical experiments that support the widely-held intuition that the process of modular synchronization in the Kuramoto system exhibits similarities to the renormalization procedure in statistical physics. In particular, we have demonstrated that in a modularly synchronized Kuramoto oscillator system, groups of mutually synchronized oscillators may be treated each as meta-oscillators, thereby yielding dynamics that belong to a family that strictly generalizes the Kuramoto model.

As shown in Figure 3.3, even the presence of a strongly modular structure in the coupling network does not guarantee that the actual meso-scale synchronization will follow the same divisions. In contrast, the identities of the mutually synchronized clusters de-

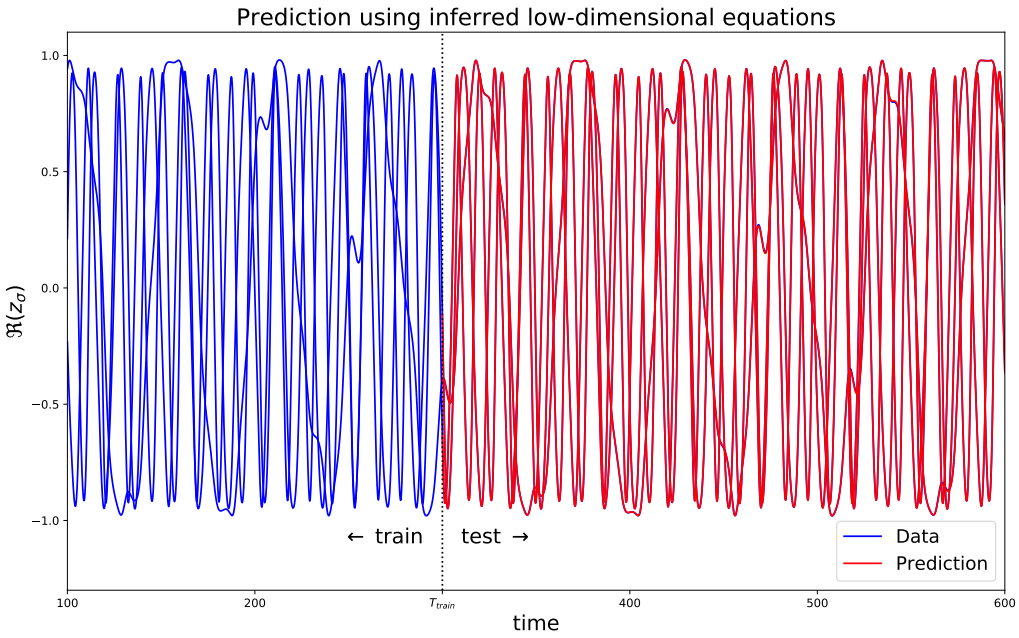

Figure 3.10. Example of the validation procedure described in the text. Blue curves are trajectories of the three cluster order parameters in a 15-node network consisting of three modules of five nodes each. Parameters of equation 3.8 were inferred from these data for $t \in [0, T_{\text{train}}] = [0, 300]$ and the resulting equations were integrated for $t \in [T_{\text{train}}, T_f] = [300, 1000]$ (red curves). Note nearly-exact overlap.

pends jointly on the coupling structure and the natural frequencies, in a way that can be understood through the procedure outlined in 3.5. This level of detail is easily glossed over in mean-field type treatments. The method that we employ can be extended to study the emergence of increasingly complex collective behavior in networks of interacting rhythmic dynamical systems. For instance, our approach can be used to elucidate the interplay of dynamical heterogeneity (i.e., distinct natural frequencies) and the coupling structure of a complex oscillator network.

3.8 Future Work

We now discuss several directions in which this work could be deepened and extended

3.8.1 Closure methods

It is possible to write exact evolution equations for the coarse-variables if you allow explicit time-dependence through the residuals $x_i(t) = z_\sigma(t) - \exp(\mathrm{i}\theta_i(t))$. Explicitly, we have

$$\begin{aligned}\dot{z}_\sigma &= A_\sigma + B_\sigma z_\sigma + D_\sigma z_\sigma^2 + \\ &+ \sum_{\sigma'=1}^{C} \left[E_{\sigma\sigma'}(z_{\sigma'} - z_{\sigma'}^* z_\sigma^2) + F_{\sigma\sigma'} z_{\sigma'}^* + G_{\sigma\sigma'} z_\sigma z_{\sigma'}^* \right]. \end{aligned} \tag{3.15}$$

The terms highlighted in blue are those that are present in the original OA equations.

The functional form of the coefficients is given by

$$A_\sigma = \left\langle \mathrm{i}\omega_i x_i + \sum_{\sigma'=1}^{C} \sum_{j \in P_{\sigma'}} K_{ij}(x_j - x_j^* x_i^2) \right\rangle_{i \in P_\sigma} \tag{3.16}$$

$$B_\sigma = \left\langle \mathrm{i}\omega_i + \sum_{\sigma'=1}^{C} \sum_{j \in P_{\sigma'}} K_{ij}(-2x_j^* x_i) \right\rangle_{i \in P_\sigma} \tag{3.17}$$

$$D_\sigma = \left\langle \sum_{\sigma'=1}^{C} \sum_{j \in P_{\sigma'}} K_{ij}(-x_j^*) \right\rangle_{i \in P_\sigma} \tag{3.18}$$

$$E_{\sigma\sigma'} = \left\langle \sum_{j \in P_{\sigma'}} K_{ij} \right\rangle_{i \in P_\sigma} \tag{3.19}$$

$$F_{\sigma\sigma'} = \left\langle \sum_{j \in P_{\sigma'}} K_{ij}(-x_i^2) \right\rangle_{i \in P_\sigma} \tag{3.20}$$

$$G_{\sigma\sigma'} = \left\langle \sum_{j \in P_{\sigma'}} K_{ij}(-2x_i) \right\rangle_{i \in P_\sigma} \tag{3.21}$$

where

$$\langle a_i \rangle_{i \in P_\sigma} \coloneqq \frac{1}{|P_\sigma|} \sum_{i \in P_\sigma} a_i. \tag{3.22}$$

If one could obtain an accurate expression for x_i in terms of microscopic parameters ω_i, K_{ij}, and the coarse variables z_σ, then the explicit time-dependence could be removed, leaving an autonomous ODE for the coarse variables $\{z_\sigma\}$. A promising approach could be to build on the work of Gottwald concerning collective coordinates [32], in which the basic idea is that within a phase-locked cluster of oscillators, the residual phase of an oscillator is directly proportional to it natural frequency relative to the cluster average

frequency. The associated constant of proportionality serves as a "collective coordinate", allowing an expression involving all phase variables to be expressed in terms of the cluster average phase and the collective coordinate.

Another possible way to remove the explicit time-dependence is essentially take a more data-driven approach to the above procedure. The time-dependence of the coefficients (A)–(G) in the ODE governing the coarse variables can be computed directly from the fine-grained time series, so one could conceivably infer a functional relationship between those coefficients and the coarse-grained variables using a method analogous to the one we used to infer a functional relationship between $\dot{z}$ and z. If such a function fit is successful, its substitution into the (3.15) would render that ODE autonomous, albeit with potentially more intricate dependence on z_σ than in just the terms that are written in (3.15) directly.

3.8.2 Data-driven function discovery

Stripping away yet another layer of theory, we can remain largely agnostic about the functional form of the equations that should govern $\dot{z}$ and infer them using a procedure such as SINDy [13]. Given the enormous complexity of the space of functions $f\colon \mathbb{C}^C \to \mathbb{C}^C$, it is prudent to impose certain restrictions on the function dictionary we choose.

First, the OA ansatz and its associated dimension reduction shows that the reduced model exhibits on-site and pairwise terms, and no higher-order (i.e. three-or-more-particle) terms. Moreover, each coupling term is of the same form. Therefore it is reasonable to suppose that our function basis should contain a set of onsite terms and a set of coupling terms, and that these terms should be replicated for each node and edge, respectively.

Next, it is well known that coarse-graining of dynamics can induce memory effects [80], so it is reasonable to expect that a coarse-grained model including memory or nonlocality (in the form of polyadic coupling terms) would be effective. Data-driven discovery of coarse-grained dynamical models including memory is discussed in [56]. Polyadic coupling terms make no conceptual difference, but combinatorial explosion may make their inclusion computationally costly.

Finally we remark on some physical principles that inform our methods for inferring coarse-grained dynamical models. Notice that in (3.8), the coefficient of the linear term has negative real part (since $\delta_\sigma \geq 0$ is the width of a Cauchy distribution, and so cannot be negative), and the coupling matrix $N_\sigma K^{\sigma\sigma'}/2N$ is real. These conditions together imply that the C-fold product of the unit disk is invariant; in other words, if all complex phases z_σ initially have magnitude not greater than 1, then they will continue to have magnitude not greater than 1 for all time. This condition breaks down if we allow the coupling matrix to have an imaginary part or the coefficient of the linear term to have positive real part, and so we impose those constraints during our inference procedure.

Those constraints can be thought of in a more general context. We know that the coarse-grained variables z_σ should always remain bounded in the unit disk, because they are averages of quantities on the unit circle. Knowing this, we can perform *constrained* inference of the right-hand side of an ODE governing those variables, with the constraint being that a function is feasible only if the corresponding ODE leaves the appropriate set in state space invariant.

Such a constraint may appear unruly. Notice, however, that we only need to impose that the vector field never points outward at the boundary. If we alternately write the system in terms of amplitudes and phases, that condition is simply $\dot{r}_\sigma \leq 0$ whenever $r_\sigma = 1$ and $r_{\sigma'} \leq 1$ for all σ'. Moreover, this constraint is convex! That means that if we construct a basis of functions $f_1, \ldots, f_L$ that all obey the disk-invariance condition, then any convex combination of those basis functions will obey it as well. Given such a basis, we can fit $\dot{z} = a_1 f_1(z) + \cdots + a_L f_L(z)$ with $a_i \geq 0$, and rest assured that the resulting right-hand side will obey the constraint that the coarse variables z remain bounded in the appropriate way.

Chapter 4

Cascading Extinction Events in Mutualistic Ecosystems

This chapter comprises work done in close collaboration with Weiran Cai and Raissa D'Souza.

4.1 Introduction

A large class of systems exhibiting collective behavior includes those consisting of many similar individuals whose pairwise interactions form the edges of a complex network. Examples from the engineered world include the worldwide web [3], critical infrastructure systems (electricity, water, natural gas) [40, 25], transportation networks (road, air, sea), supply chains [38], and organizational hierarchies (government, business) [59]. In the natural world, waterways form physical networks that connect the physical landscape in new and nontrivial ways [69]; chemical species are linked by relations of reactant and product [19]; living things depend on biophysical networks to transport materials critical to life (mycelia, trees, vasculature) [72]; many organisms arrange themselves in myriad social structures [58]; and finally the many species that exist in any given environment interact with one another [33].

It is this last example that we take as our inspiration for the present line of work. Specifically, we mention *mutualistic* ecosystems. A mutualistic ecosystem is one in which multiple types of species interact in a mutually beneficial manner (for example, plants

and the animals that pollinate them). The data of which species exist and which pairwise relations they hold can be called a *mutualistic network*. Note that mutualism is not strictly constrained to ecological contexts; mutualistic networks have been identified in socio-technical contexts as well, and have been observed to share certain structural similarities with their ecological counterparts [61]. The two structural features that have been most prominently observed in mutualistic networks are *nestedness* [8, 11] and *modularity* [49, 51]. In short, nestedness is the tendency for the neighbors of a given node to be a subset of the set of neighbors of any higher-degree node, and modularity is the property that the network can be divided into modules such that most of the links are between nodes in the same module as each other.

One can ask any number of questions about a mutualistic ecosystem and how its operation can be understood by measuring its network architecture. We focus especially on the phenomenon of cascading extinction events. By *cascading*, we mean that one extinction event may precipitate others, and so on, possibly leading to the collapse of the entire ecosystem.

Our central aim is to understand the effect of nestedness (as manifested by degree heterogeneity) and modularity on the behavior of cascades of extinction events on networks. In what follows, we will formulate a mathematical model that captures certain qualitative features of this phenomenon and analyze its behavior.

4.2 Technical survey

There are many levels at which to model species extinction. A common choice is to write a stochastic differential equation (SDE) that governs the abundance (say, biomass in kg) of each species as a function of time. In this framework it is possible to model species extinction either as a function of an initial set of extinct species, or as a function of changing environmental conditions, and to observe that extinction events are preceded by generic signals that indicate proximity to a tipping point [21].

One can also model an ecosystem at a more coarse level, assigning to each species a binary variable whether that species is present or absent. While this is a significant

simplification, the system still supports rich phenomenology. One interesting line of work in this direction has been developed by Colin Campbell and coauthors [14], who model the formation of mutualistic networks and study their sensitivity to species removal.

4.3 Dynamics

For our context, we choose to use a version of the threshold model introduced by Watts [71] to study diffusion of information through a social network. In Watts' original work, the system is comprised of N individuals connected by a network with (unweighted and undirected) adjacency matrix $A = A_{ij} \in \{0, 1\}$. Each node i is assigned a state $u_i(t) \in \{0, 1\}$ that evolves in discrete time $t \in \mathbb{N}$; the state 0 is called "inactive" while the state 1 is called "active". In the context of information diffusion, an active node knows some piece of information and an inactive node does not. Some proportion of nodes are selected at random to be active at time $t = 0$ while all the others remain inactive. The state of each node then updates in time according to the rule

$$u_i(t+1) = \begin{cases} 1 & \sum_j A_{ij} u_j(t) > \theta k_i \text{ and } u_i(t) = 1 \\ 0 & \text{else} \end{cases} \tag{4.1}$$

where $k_i = \sum_j A_{ij}$ is the *degree* of node i and $\theta \in [0, 1]$ is the *threshold*.

In the language of information diffusion, an actor learns a piece of information if it is known by at least some fraction of its neighbors, and no actor ever forgets that piece of information. In the language of population dynamics, an inactive species is present, and an active species is extinct; a species becomes extinct (active) if at least some fraction of its neighbor (i.e. mutualistic partner) species is extinct (active). It may be initially confusing that the "active" state corresponds to extinction; the reason for that choice is that extinction is the absorbing state (i.e. a species cannot come back from extinction in our model).

4.3.1 Deriving linear threshold dynamics from population ODEs

Here we show a calculation, starting from Lotka-Volterra equations for population dynamics, that motivate the binary-state dynamics described above as a faithful approximation of extinction behavior.

To begin, we follow [21] and others and suppose that the abundances of animal species $\{A_k\}$ and of plant species $\{P_i\}$ evolve in time as

$$\begin{aligned}\frac{\mathrm{d}P_i}{\mathrm{d}t} &= P_i\left(\alpha_i^P - \sum_{j=1}^{N_P}\beta_{ij}^P P_j + \frac{\sum_{k=1}^{N_A}\gamma_{ik}^P A_k}{1+h\sum_{k=1}^{N_A}\gamma_{ik}^P A_k}\right)\\ \frac{\mathrm{d}A_k}{\mathrm{d}t} &= A_k\left(\alpha_k^A - \sum_{l=1}^{N_A}\beta_{kl}^A A_l + \frac{\sum_{i=1}^{N_P}\gamma_{ki}^A P_i}{1+h\sum_{i=1}^{N_P}\gamma_{ki}^A P_i}\right)\end{aligned} \tag{4.2}$$

where α represents a species' intrinsic rate of growth (possibly negative), β are matrices describing competition among plants and animals, γ are matrices describing the mutually beneficial relations between plants and animals, and h is a parameter known as the handling time that controls the saturation of the mutualistic term.

We assume uniform growth rate ($\alpha_i^P \equiv \alpha_k^A \equiv 0$) and uniform competition ($\beta_{ij}^G \equiv 1/N_G$, $G \in \{A, P\}$, and $\beta_{ii}^P = \beta_{kk}^A = 1$). We assume further that mutualistic strength is inversely proportional to node degree, which we write as $\gamma_{ki}^A = \gamma_0 M_{ki}/k_k^A$, $\gamma_{ik}^P = \gamma_0 M_{ik}/k_i^P$, where $M \in \{0,1\}^{N_A \times N_P}$ is the bipartite adjacency matrix of the mutualistic network, so that $M_{ki} = 1$ if animal k and plant i interact, and zero otherwise. The degree is then defined as $k_k^A = \sum_j M_{kj}$ and $k_i^P = \sum_l M_{li}$, for animals and plants respectively.

Under these assumptions, the equations simplify to

$$\begin{aligned}\frac{\mathrm{d}P_i}{\mathrm{d}t} &= P_i\left(\alpha - P_i - \frac{\beta}{N_P}\sum_{j=1}^{N_P}P_j + \frac{\frac{\gamma_0}{k_i^P}\sum_{k=1}^{N_A}A_{ik}A_k}{1+h\frac{\gamma_0}{k_i^P}\sum_{k=1}^{N_A}A_{ik}A_k}\right)\\ \frac{\mathrm{d}A_k}{\mathrm{d}t} &= A_k\left(\alpha - A_k - \frac{\beta}{N_A}\sum_{l=1}^{N_A}A_l + \frac{\frac{\gamma_0}{k_k^A}\sum_{i=1}^{N_P}A_{ki}P_i}{1+h\frac{\gamma_0}{k_k^A}\sum_{i=1}^{N_P}A_{ki}P_i}\right)\end{aligned} \tag{4.3}$$

Now, consider a focal animal species k, and consider the case that $A_l = \overline{A}$ for all $l \neq k$, and $P_i = \overline{P}$ for some i and $P_i = 0$ for the rest. How many partners of animal species k must remain alive in order for it to persist? In other words, we seek conditions such that there exists a positive equilibrium value of A_k. Letting $x = \gamma_0\overline{P}m_k/k_k^A$, with m_k denoting the number of partners of animal species k that are still alive, we find that animal species k can survive if

$$x > \frac{\beta\overline{A} - \alpha}{1 + \alpha h - \beta h\overline{A}} \tag{4.4}$$

or equivalently if

$$\frac{m_k}{k_k^A} > \frac{1}{\gamma_0 \overline{P}} \frac{\beta\overline{A} - \alpha}{1 + \alpha h - \beta h \overline{A}} \tag{4.5}$$

which is just a fractional threshold condition, provided $1 + \alpha h - \beta h\overline{A} > 0$.

Mathematically, this calculation works because of our assumption, made in previous work [21, 62], that the benefit any given species receives due to a mutualistic interaction is inversely proportional to the number of mutualistic partners (neighbors) that species has. Logically, this same reasoning applied to the discrete setting leads to the fractional threshold model that we now study.

4.3.2 Analytical approximation of cascade dynamics

We now return to the discrete linear threshold model introduced by Watts [71] and defined in (4.1). The basic sort of question we consider about this model is: given an assortment of initially-activated "seed" nodes, and properties of the network, what can we predict about the size or likelihood of a large cascade of activation?

Watts' analysis proceeds by considering the size of the largest "vulnerable cluster" of nodes. A node is considered vulnerable if it can be activated by a single active neighbor. Clearly, then, a connected set of vulnerable nodes can be all activated by only a single activation of any one of them (or of any neighbor of a node in the connected vulnerable cluster), so the size of the vulnerable cluster is strongly informative of the expected size of a cascade of activation.

A thorough and rigorous analysis by Gleeson [29] established a general method for studying dynamics with the "permanently active property" (PAP) on networks with arbitrary degree distributions and either modularity or degree correlations, which we now summarize. The assumptions required for this method to work are that the dynamics should be binary-state (i.e. a node is either active or inactive) and irreversible (i.e. a node that has become active cannot become inactive), and the network on which the dynamics take place should be well-described by a partition $\{C_1, \ldots, C_M\}$ of the nodes $\{1, \ldots, N\}$ into modules, a degree distribution $p_k^{(i)}$ for each module, and a mixing matrix $\mathbf{e} \in \mathbb{R}^{M \times M}$, defined such that e_{rs} is the fraction of links that connect module r to module s.

Assuming further that the network can be modeled as locally tree-like, and that the number of initially active nodes is small, we can write a recursion relation governing the quantities $q_n^{(i)}$, defined as the probability that a node in module i and tree level n is active, conditional on its parent (at level $n+1$) being inactive. With $\rho_0^{(i)}$ representing the fraction of nodes in module i that are initially active, and $F^{(i)}(m,k)$ the probability that a node of degree k in module i with m active neighbors will become active, we have

$$q_{n+1}^{(i)} = \rho_0^{(i)} + (1-\rho_0^{(i)}) \sum_k \frac{k}{z^{(i)}} p_k^{(i)} \sum_{m=0}^{k-1} \binom{k-1}{m} (\overline{q}_n^{(i)})^m (1-\overline{q}_n^{(i)})^{k-1-m} F^{(i)}(m,k) \tag{4.6}$$

$$\coloneqq g^{(i)}(\overline{q}_n^{(i)}) \tag{4.7}$$

where

$$\overline{q}_n^{(i)} \coloneqq \frac{\sum_j e_{ij} q_n^{(j)}}{\sum_j e_{ij}} \tag{4.8}$$

$$z^{(i)} \coloneqq \sum_k k p_k^{(i)}. \tag{4.9}$$

Here $\overline{q}_n^{(i)}$ represents the probability that a randomly chosen neighbor of a node at tree level n in module i, is active, and $z^{(i)}$ is the average degree in module i.

We can iterate this equation to find $q_\infty^{(i)} = \lim_{n\to\infty} q_n^{(i)}$, and then translate this into the asymptotic density of active nodes by the formula

$$\rho_\infty^{(i)} = \rho_0^{(i)} + (1-\rho_0^{(i)}) \sum_k p_k^{(i)} \sum_{m=0}^{k-1} \binom{k}{m} (\overline{q}_\infty^{(i)})^m (1-\overline{q}_\infty^{(i)})^{k-1-m} F^{(i)}(m,k). \tag{4.10}$$

Next, we note that this setup is substantially more general than we need for our purposes. The Watts cascade model is obtained by setting

$$F^{(i)}(m,k) = \begin{cases} 1 & m \geq \theta k \\ 0 & m < \theta k \end{cases} \tag{4.11}$$

In this case, Equation 4.6 can be rewritten as

$$q_{n+1}^{(i)} = \rho_0^{(i)} + (1-\rho_0^{(i)}) \sum_k \frac{k}{z^{(i)}} p_k^{(i)} \sum_{m=\lceil 1/\theta \rceil}^{k-1} \binom{k-1}{m} (\overline{q}_n^{(i)})^m (1-\overline{q}_n^{(i)})^{k-1-m} \tag{4.12}$$

Finally, we demonstrate how to translate the discrete-time update equation Equation 4.6 into an ODE. If we assume that nodes update their state in random order at a rate of f fraction of the nodes per unit time, then we can describe the state of the network at time t by the variables $q^{(i)}(t)$, which obey the ODE

$$\frac{\mathrm{d}q^{(i)}}{\mathrm{d}t} = f\left[g^{(i)}(\overline{q}^{(i)}(t)) - q^{(i)}(t)\right]^{+} \tag{4.13}$$

where $g^{(i)}$ is defined in Equation 4.7 and $[\cdot]^{+}$ denotes the positive part of the quantity in square brackets. The initial condition is $q^{(i)}(0) = \rho_0^{(i)}$.

In [29], Gleeson presents the solution of this ODE in comparison with direct numerical simulation of the corresponding network of $\mathcal{O}(10^5)$ nodes, finding excellent agreement.

4.3.2.1 Special cases

The presentation above is only one of three formulations given in [29]. The other two can be obtained straightforwardly from the one presented here, as we now describe.

First is the case where there is no modular structure. In simple terms, we have $M = 1$, and $\mathbf{e}$ is the 1×1 matrix 1. Dropping all "(i)" superscripts gives exactly the appropriate formulation for a network without modular structure.

Next is the case of a network without a modular structure but with degree correlations. Mathematically, the trick is to consider each value of degree k to define a "module" consisting of all those nodes having degree k. The joint degree distribution $P(k, k')$, defined as the probability that a randomly selected link joins a node of degree k to a node of degree k', then serves directly as the mixing matrix $\mathbf{e}$. Finally, this situation allows for degree-targeted seeding; that is, the initial density $\rho_0^{(k)}$ represents the probability that a node of degree k is active initially.

4.3.3 Degree-targeted seeding

We now present a generalization of the Gleeson's framework, described in subsection 4.3.2, to model cascading dynamics on a modular network in a way that allows for seed nodes to be selected with a probability that depends on their degree.

The generalization is simple. Instead of the initial active density being $\rho_0^{(i)}$, it now depends on degree k, and we write $\rho_{0,k}^{(i)}$ as the probability that a randomly selected node

in module i having degree k is active initially. This generalization allows to model analytically cases where, for example, a cascade is initialized by activating the highest-degree nodes in a given module. The case that initialization is independent of node degree is recovered by setting $\rho_{0,k}^{(i)} = \rho_0^{(i)}$ for all k. The total fraction of nodes initially activated is given by

$$\rho_{0,\mathrm{tot}}^{(i)} = \sum_k p_k^{(i)} \rho_{0,k}^{(i)}. \tag{4.14}$$

We remark that as noted in subsubsection 4.3.2.1, it is already possible to model degree-dependent seeding within the framework that Gleeson presents. However, this is restricted to cases where there is not modular structure, and the solution requires to keep track of infinitely many dynamical variables $\{q_n^{(k)} | k \in \mathbb{N}\}$. As we will show, it is possible to model degree-targeted seeding in a modular network in a way that is only marginally more computationally intensive than the uniform-seeding case. In particular, the number of dynamical variables required is the same, while the number of parameters roughly doubles (from $\mathcal{O}(M^2) + Mk_{\max}$ to $\mathcal{O}(M^2) + 2Mk_{\max}$).

Following the derivation in [29], we assume the network is locally treelike, and denote by $q_{n,k}^{(i)}$ the probability that a node in module i having degree k at level n of the tree is active, conditional on its parent (at level $n+1$) being inactive. Accounting for degree-weighted seeding, the variables $q_{n,k}^{(i)}$ obey the recursion relation

$$q_{n+1,k}^{(i)} = \rho_{0,k}^{(i)} + (1 - \rho_{0,k}^{(i)}) \sum_{m=0}^{k-1} \binom{k-1}{m} \left(\overline{q}_n^{(i)}\right)^m \left(1 - \overline{q}_n^{(i)}\right)^{k-1-m} F^{(i)}(m,k) \tag{4.15}$$

where $\overline{q}_n^{(i)}$ is given by

$$\overline{q}_n^{(i)} = \frac{1}{\sum_j e_{ij}} \sum_j e_{ij} \left[\sum_k \frac{k}{z^{(j)}} p_k^{(j)} q_{n,k}^{(j)} \right]. \tag{4.16}$$

The total density of active nodes can be constructed by

$$\begin{aligned} \rho_n^{(i)} &= \sum_k p_k^{(i)} \left[\rho_{0,k}^{(i)} + (1 - \rho_{0,k}^{(i)}) \sum_{m=0}^{k} \binom{k}{m} \left(\overline{q}_n^{(i)}\right)^m \left(1 - \overline{q}_n^{(i)}\right)^{k-m} F^{(i)}(m,k) \right] \\ &\coloneqq H^{(i)}(\overline{q}) \end{aligned} \tag{4.17}$$

The variables $\overline{q}_n^{(i)}$ represent the probability that a randomly chosen neighbor of a node in module i at tree level n is active, hence the factor $k/z^{(j)}$. With this in mind, we can interpret the update rule Equation 4.15 as follows. Given a node in module i with degree k at tree level n whose parent is not active, it was active from the start with probability $\rho_{0,k}^{(i)}$. With probability $1-\rho_{0,k}^{(i)}$, it was not active from the start, and has a chance to be activated by its $k-1$ children. Each of those $k-1$ children is active with probability $\overline{q}_n^{(i)}$, meaning that the probability that m of them are active at once is $\binom{k-1}{m}\left(\overline{q}_n^{(i)}\right)^m\left(1-\overline{q}_n^{(i)}\right)^{k-1-m}$. If m children are active, then the focal node becomes active with probability $F^{(i)}(m,k)$.

Finally, consider the unconditional probability $\rho_n^{(i)}$ that a node in module i at tree level n is active. If we choose a node at random from module i, it has degree k with probability $p_k^{(i)}$. Given that it has degree k, it was active from the start with probability $\rho_{0,k}^{(i)}$. Otherwise, it has k neighbors, each of which is active with probability $\overline{q}_n^{(i)}$. The total number of active neighbors then follows the binomial distribution seen in Equation 4.17, and the focal node becomes active with probability $F^{(i)}(m,k)$ if it has m active neighbors.

It may appear over-complicated that we must keep track of a variable $q_{n,k}^{(i)}$ for every module i and every degree k. However, the dependence on k is such that we can compose equations (4.15) and (4.16) to obtain a recursive formula for $\overline{q}_n^{(i)}$ directly. We have

$$\overline{q}_{n+1}^{(i)} = \frac{1}{\sum_j e_{ij}}\sum_j e_{ij}\Bigg[\sum_k \frac{k}{z^{(j)}}p_k^{(j)}\Bigg(\rho_{0,k}^{(j)} + (1-\rho_{0,k}^{(j)})\cdot \sum_{m=0}^{k-1}\binom{k-1}{m}\left(\overline{q}_n^{(j)}\right)^m\left(1-\overline{q}_n^{(j)}\right)^{k-1-m}F^{(j)}(m,k)\Bigg)\Bigg] \tag{4.18}$$

$$:= G^{(i)}(\overline{q}_n) \tag{4.19}$$

Note that if $\rho_{0,k}^{(i)} = \rho_0^{(i)}$ for all k, then Equation 4.19 reduces to Equation 4.8 composed with Equation 4.6.

We can further extend this framework to describe the time course of the dynamics obtained by assuming that nodes update asynchronously at a given rate per unit time.

Following the same reasoning as in [29], we have

$$\begin{aligned} \frac{\mathrm{d}\overline{q}^{(i)}}{\mathrm{d}t} &= f\left[G^{(i)}(\overline{q}) - \overline{q}\right]^{+} \\ \frac{\mathrm{d}\overline{\rho}^{(i)}}{\mathrm{d}t} &= f\left[H^{(i)}(\overline{q}) - \rho\right]^{+} \end{aligned} \tag{4.20}$$

Intuitively, the selection of seed nodes should have a notable impact on the spreading dynamics. To validate this, we replicate the numerical experiment performed by Gleeson in [29]. The experiment consists of a network of four modules, connected in a ring. Their degree distributions are Poisson with mean 5.8, Poisson with mean 8, regular with degree 8, and regular with degree 8, respectively. The mixing matrix is

$$e = \frac{1}{29.8}\begin{bmatrix} 5.5 & 0.15 & 0.15 & 0 \\ 0.15 & 7.7 & 0 & 0.15 \\ 0.15 & 0 & 7.7 & 0.15 \\ 0 & 0.15 & 0.15 & 7.7 \end{bmatrix} \tag{4.21}$$

The initial condition is that 1% of the nodes in the first module are active, and the threshold is taken to be $\theta = 0.18$. For these parameter settings, the cascade eventually takes over the whole network, but reaches each of the modules at different times, due to their differing internal characteristics (i.e. degree distributions) and the nature of the links between them (as encoded by the matrix e).

To see the effect of seeding the highest-degree nodes in the network, we recreate calculations for the same network using the same fraction of nodes initially activated (i.e. 1% of the first module), selecting nodes uniformly at random on the one hand, and on the other hand selecting them according to highest degree.

For concreteness, we now explicitly construct $\rho_{0,k}^{(i)}$ for the case where we seed only the highest-degree nodes in the network. Say we have, for each module, a target fraction of nodes that we would like to activate initially, $\rho_{0,\mathrm{tot}}^{(i)} \in [0,1]$ for $i = 1, \ldots, d$. We would like to choose $\rho_{0,k}^{(i)}$ such that the appropriate total number of nodes are activated, but only the nodes of highest possible degree are chosen.

To do this, we need to find, for each module i, a value K^i such that the fraction of nodes in module i having degree greater than or equal to K^i is close to $\rho_{0,\mathrm{tot}}^{(i)}$. Formally,

let

$$K^i = \min \left\{ K \;\middle|\; \sum_{k>K} p_k^{(i)} \leq \rho_{0,\text{tot}}^{(i)} \right\} \tag{4.22}$$

with the understanding that the minimum of an empty set is $+\infty$. Then, define

$$\rho_{0,k}^{(i)} = \begin{cases} 1 & k > K^i \\ \rho_{0,\text{tot}}^{(i)} - \sum_{k>K^i} p_k^{(i)} & k = K^i \\ 0 & k < K^i \end{cases} \tag{4.23}$$

We then construct the ODE defined in Equation 4.20, with threshold $\theta = 0.18$. Results are shown in Figure 4.1. Note excellent agreement between theory and experiment for the "baseline" case, i.e. seed probability independent of degree, as considered by Gleeson; the solid and dashed orange curves nearly perfectly overlap each other. This tells us that our generalization does in fact reduce to the known model when dependence of seed probability on degree becomes trivial. For the degree-weighted case, we also see nearly perfect agreement, indicating that our approach is sound.

Next we consider a slight modification of the above example, to highlight the effect of maximum-degree seeding. In particular, we take the same network as a above, but with threshold $\theta = 0.21$. We choose this value because if seeding is independent of degree, we see no global cascade, while we do see a global cascade if seeding is targeted to the highest-degree nodes. Results are presented in Figure 4.2.

4.4 Application: extending work on optimal modularity

Here we describe extensions to the research presented in [48]. In that work, the authors consider a network composed of two modules, with a fraction $\mu \in [0, 1]$ of the links joining nodes in the same module and the remaining $(1 - \mu)$ fraction of the links joining nodes in different modules. On this network, they consider linear threshold dynamics of the type described above, with seed nodes localized to one of the modules.

The main result of [48] is that for certain values of the seed density, there is an optimal range of values of the modularity parameter μ, such that within this range the cascade

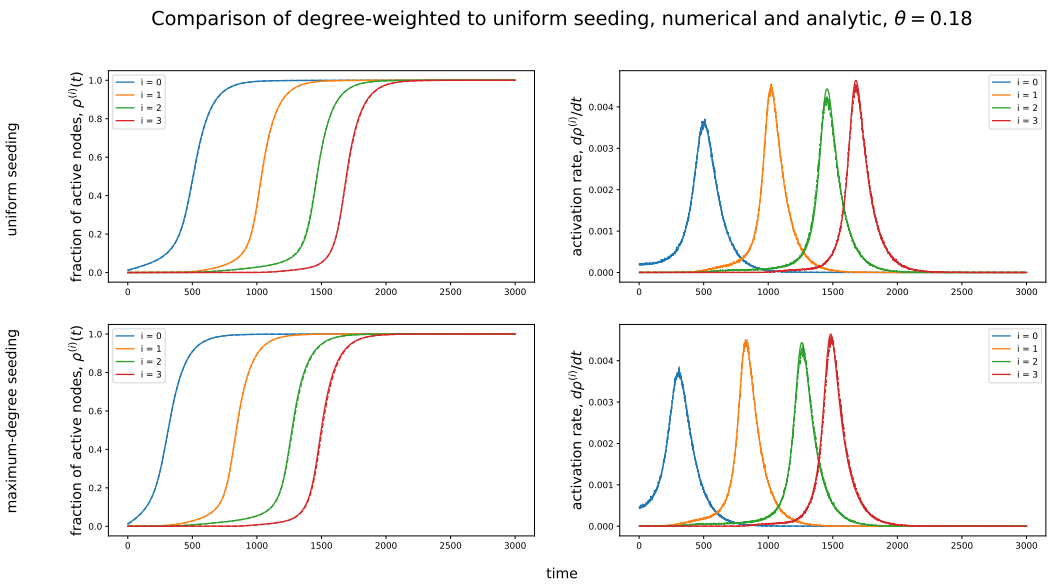

Figure 4.1. Demonstration of the performance of the approximate ODEs obtained by the treelike approximation under the assumption of degree-dependent seeding. Parameters are identical to those used to create Fig. 1 in [29]. Here we display both the total fraction of active nodes (left) and the rate of increase of the number of active nodes (right) in each module, for both uniform (top) and degree-targeted seeding (bottom). Solid curves are solutions of the ODE system Equation 4.20, and dashed curves are the corresponding quantities in a direct numerical simulation of linear threshold dynamics on a network of size $N = 5 \times 10^5$, averaged over 10 realizations of the network. nodes.

covers the whole system, while on either side the cascade remains localized to the module where it began. Their conclusions follow from both direct simulations of dynamics on networks and calculations based on the analytical framework developed by Gleeson [29].

We now generalize their results in two ways. First, we consider degree distributions with a tunable extent of degree heterogeneity, and we allow for degree-targeted seeding.

The first aspect, degree heterogeneity, was discussed in the SI of [48], where the authors present results for the LFR benchmark networks [44], and state that degree heterogeneity does not change the results qualitatively. Here, instead, we treat degree heterogeneity explicitly as a control parameter, quantified by $p_{\text{nest}} \in [0, 1]$. The parameter p_{nest} enters the analysis through the degree distribution:

$$p_k = p_{\text{nest}} p_k^{\text{pow}} + (1 - p_{\text{nest}}) p_k^{\text{poi}} \tag{4.24}$$

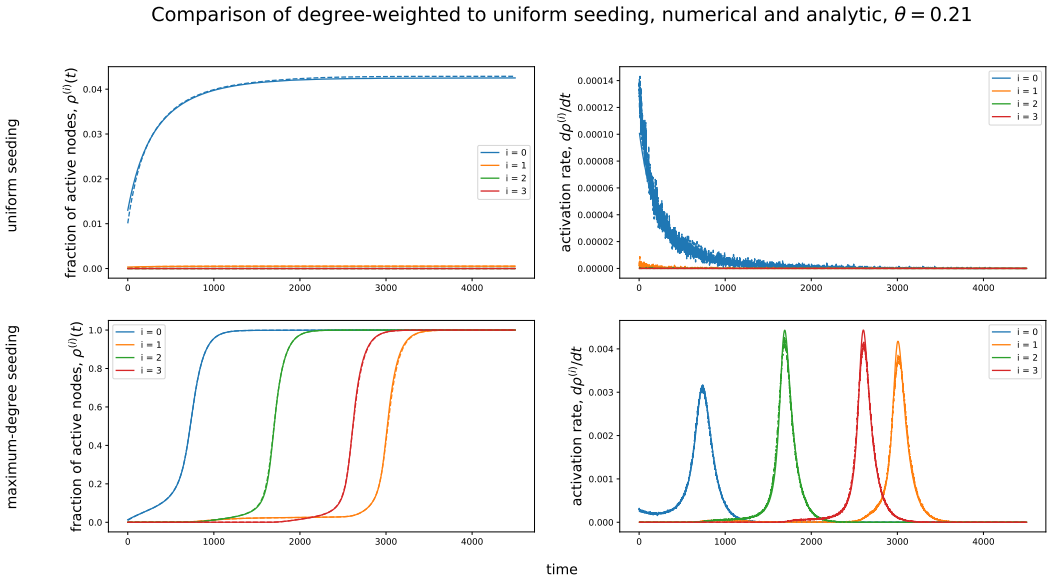

Figure 4.2. As in Figure 4.1, but with threshold $\theta = 0.21$. Note that with uniform seeding (top), the cascade does not reach the whole network, while for maximum-degree seeding, it does. Again we see excellent agreement between numerical results and analytic predictions.

where p^{pow} and p^{poi} are power law and Poisson degree distributions, respectively, each with mean z (which we take, for now, to be 20). For completeness, we have

$$p_k^{\text{pow}} = \frac{1}{\zeta(\gamma, \lambda)}(\lambda + k)^{-\gamma} \tag{4.25}$$

with λ chosen such that $\sum_k k p_k^{\text{pow}} = z$, and $\zeta(\gamma, \lambda) = \sum_k (\lambda + k)^{-\gamma}$ is the Hurwitz zeta function, and

$$p_k^{\text{poi}} = \frac{z^k \mathrm{e}^{-z}}{k!}. \tag{4.26}$$

The question of degree-targeted seeding was not addressed in [48], and we use the new analytic framework presented in subsection 4.3.3 to do so. Results are presented in Figure 4.3. The bottom row (uniform seeding) exactly recreates the results of [48]. We see that as nestedness (degree heterogeneity) increases, the region of incomplete cascade in the first module grows, indicating that the system becomes more resilient to cascading failures. However, we observe an opposite tendency in the top row (degree-targeted seeding). As degree heterogeneity grows, cases of partial cascade become less common, and the system is much more susceptible to a global cascade even with a small fraction of seed nodes.

This difference aligns well with the intuition that degree-heterogeneous networks are "robust yet fragile", as famously demonstrated by the pair of papers by Cohen et. al. that treat percolation on networks with a power law degree distribution, where occupation probability is either uniform or targeted by highest degree [17, 18].

4.5 Conclusion

In this work we have considered a binary-state linear threshold model as a toy model for cascading extinctions in a mutualistic ecosystem. The linear threshold model allows us to account for degree heterogeneity and modularity in the underlying network, and we present a novel analytical approach that can illuminate the effect of degree-targeted seeding of cascades on such a network. The framework is very general and can be applied to analyze many different sorts of dynamics on networks, including degree-targeted percolation (as previously studied by other means [17, 18]).

We then applied this new analytical formulation to understand the effect of degree-targeted seeding in a modular, degree-heterogeneous network. Previous work on this system under the assumption of uniform seeding found that there is a finite range of "optimal modularity" within which a cascade initiated in one module will spread to the whole network, and that this conclusion is qualitatively unchanged by degree heterogeneity. The fact that too little modularity leads to a lack of global cascade can be understood by noticing that the more links exist between modules, the less likely it is for a node in the first module to be linked to a seed node.

In contrast, we find that seeding a cascade at the network's highest-degree nodes substantially weakens the damping effect of low modularity, leaving the system susceptible to a global cascade with a much number of seed nodes than are needed when seeds are selected independent of degree. Unsurprisingly, the size of this effect grows with degree heterogeneity.

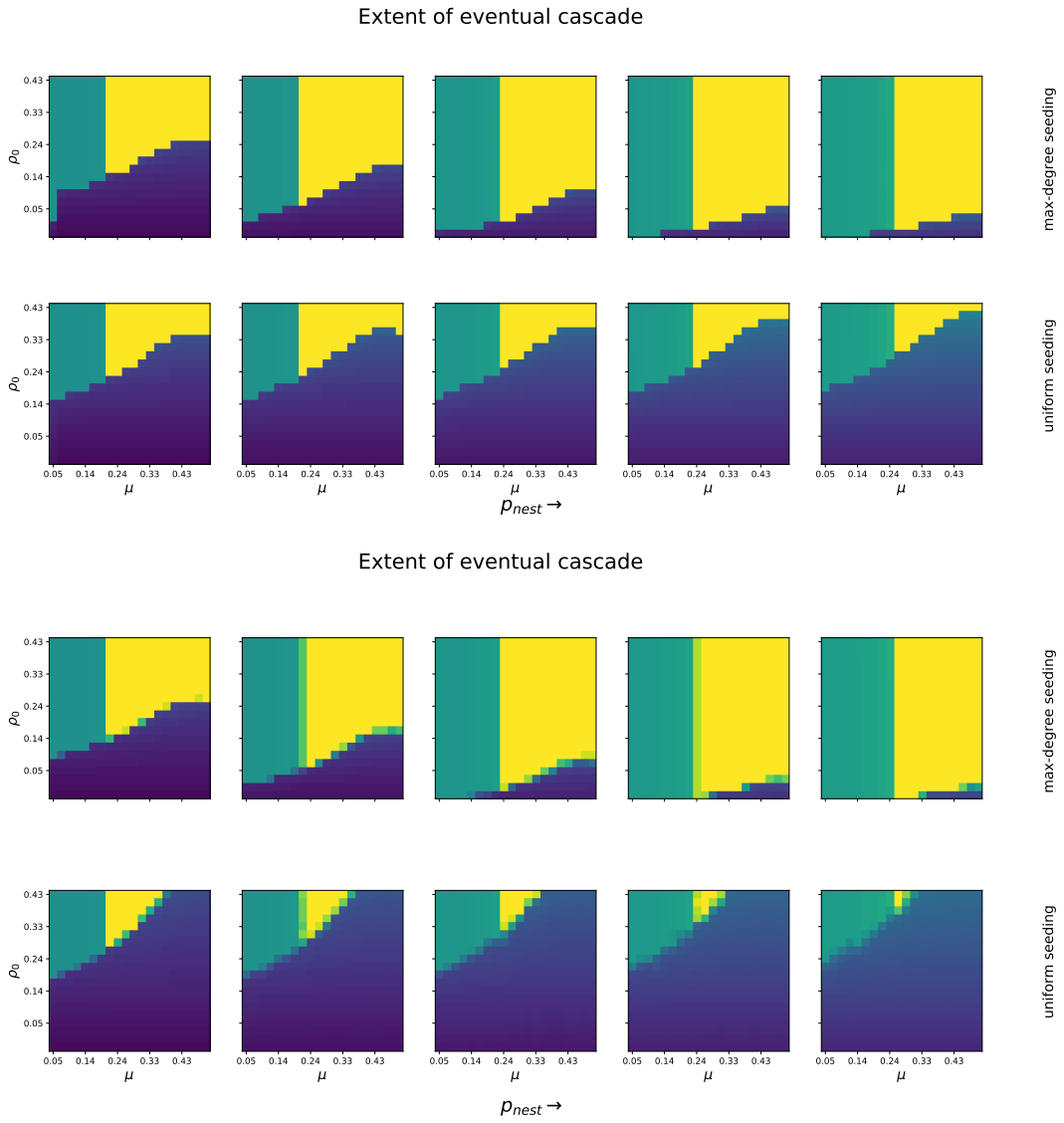

Figure 4.3. Summary of the joint effect of modularity and nestedness on the extent of cascade spreading, under both uniform and degree-targeted seeding. Color indicates extent of the eventual cascade, as a fraction of the whole network; yellow is 1, green is 0.5, and blue ranges between 0 and 0.3. The green region, present in every panel for $\mu \lesssim 0.2$, corresponds to the cascade completely covering the first module (where the seed nodes are located) and not spreading at all to the second. The blue region corresponds to the situation that the cascade spreads to only part of the first module. Parameters used here are: $\theta = 0.4$, $\overline{k} = 20$. In these figures, ρ_0 is the fraction of the first module that is infected, which is off from the notation in [48] by a factor of two. The top set of heatmaps is from iteration of the equations (4.19)–(4.17), and the bottom set is from averaging over direct simulation of the network dynamics, on networks of size $N = 2.5 \times 10^4$, averaged over ten realizations.

Chapter 5

Conclusion

5.1 Lessons Learned

What have we learned? The lessons I, personally, take from this dissertation operate at two levels. There are object-level lessons, insights about the systems under study. And then there are meta-lessons, insights that concern how and why we were able to obtain those object-level lessons.

5.1.1 Object-level lessons

From coupled entrainment, we learned that collective behavior can arise in more than one way. In particular, we considered top-down and bottom-up processes that establish frequency and phase order in a heterogeneous population of oscillators.

From coarse-graining, we observed that partially-synchronized networks of oscillators admit a dynamical description at a coarse-grained level. Moreover, that dynamical description has the same form as the one predicted in the $N \to \infty$ limit for the case of a modular coupling architecture.

From degree-targeted seeding, we learned that the initial conditions of a spreading process strongly influence the outcome of that spreading process. We also established quantitatively that the impact of targeting nodes by degree is greater for more degree-heterogeneous networks. Indeed, degree-targeting on a degree-regular network is identical to distributing nodes uniformly at random. More specifically, we found that degree targeting facilitates global cascades in modular networks.

5.1.2 Meta-lessons

First and foremost, this dissertation as a whole demonstrates the potential of toy models to help us understand the principles that underlie collective behavior. Any study of collective behavior that does not concern itself with concrete examples is ultimately impotent, but not all examples are created equal. Without sufficient simplification, general principles are liable to be lost in the details. In each chapter, we begin with a class of phenomena, and impose simplifying assumptions that allow us to derive precise relations between quantities of interest.

Next, there are commonalities in the mathematical reasons that each of the toy models in question is tractable. Broadly, tractability in the present work has its origin in linearity, symmetry, and low-dimensionality.

In coupled entrainment, the initial problem of locating a fixed point is in general nontrivial. But thanks to the simplifying assumptions we made, there exists a fixed point at which the influence of the coupling term exactly balances out to zero, irrespective of coupling strength. The fact that the fixed point in question stays at the same location for all values of coupling strength means that the linear stability problem is especially simple, since the state-dependence of the linearization is trivial. And of course, the very procedure of linearization takes advantage of the fact that it's straightforward to exactly solve the dynamics of a linear system.

In coarse-graining, we take advantage of the fact that coupled oscillators tend to align with one another to conclude that the system can be effectively described with fewer dimensions than its native formulation. By making a fortuitous ansatz about the form of the dynamics of those reduced variables, the problem of inferring equations of motion becomes linear, and hence computationally tractable. Finally, we observe that the precision with which coarse-grained variables evolve autonomously depends on symmetry between the fine-grained variables corresponding to each cluster; within a cluster, all the oscillators should be both phase-locked and coupled in the same manner to the rest of the system. Said differently, the oscillators within each cluster should be symmetric with each other in the coupling network, as well as being phase locked.

In cascades, they key analytical reduction replaces each module in a network with a single variable, representing the probability that any given node in that module is active. On the face of it, it's remarkable that such a quantity should evolve autonomously, without needing to keep track of the whole state of the network. The reason this reduction is possible is that there is a statistical symmetry between oscillators within a given module, by construction of the networks we consider. That statistical symmetry, in turn, gives us a low-dimensional model that reproduces coarse observables to high accuracy.

5.2 Future directions

We finish by discussing some possible directions for further inquiry building off of the work presented here.

5.2.1 Coupled Entrainment

The work on coupled entrainment could be extended by bending or breaking some of the simplifying assumptions that were used. Perhaps most prominently, the entrainment signal considered is not possible to construct as an L^2 function. It would be interesting to consider approximations to that entrainment signal with finite input energy, and to see in what sense the behavior of the approximating systems approach the behavior of the limit system presented here. Conversely, it may be that a different choice of natural frequency distribution would admit a finite-energy entrainment signal that also leads to a splay state with comparable linear stability properties.

5.2.2 Coarse Graining

The coarse-graining work could be extended in several directions, as outlined in that chapter. I briefly recall them here:

5.2.2.1 Closure methods

It is possible to write exact evolution equations for the coarse-variables if you allow explicit time-dependence through the residuals $x_i(t) = z_\sigma(t) - \exp(\mathrm{i}\theta_i(t))$.

If one could obtain an accurate expression for x_i in terms of microscopic parameters ω_i, K_{ij}, and the coarse variables z_σ, then the explicit time-dependence could be removed,

leaving an autonomous ODE for the coarse variables $\{z_\sigma\}$. A promising approach could be to build on the work of Gottwald concerning collective coordinates [32].

Another possible way to remove the explicit time-dependence is essentially take a more data-driven approach to the above procedure. The time-dependence of the coefficients in the ODE governing the coarse variables can be computed directly from the fine-grained time series, so one could conceivably infer a functional relationship between those coefficients and the coarse-grained variables using a method analogous to the one we used to infer a functional relationship between $\dot{z}$ and z. If such a function fit is successful, it would render that ODE autonomous, albeit with potentially more intricate dependence on z_σ.

5.2.2.2 Data-driven function discovery

Stripping away yet another layer of theory, we can remain largely agnostic about the functional form of the equations that should govern $\dot{z}$ and infer them using a procedure such as SINDy [13]. Given the enormous complexity of the space of functions $f\colon \mathbb{C}^C \to \mathbb{C}^C$, it is prudent to impose certain restrictions on the function dictionary we choose.

First, the OA ansatz and its associated dimension reduction shows that the reduced model exhibits on-site and pairwise terms, and no higher-order (i.e. three-or-more-particle) terms. Moreover, each coupling term is of the same form. Therefore it is reasonable to suppose that our function basis should contain a set of onsite terms and a set of coupling terms, and that these terms should be replicated for each node and edge, respectively.

Next, it is well known that coarse-graining of dynamics can induce memory effects [80], so it is reasonable to expect that a coarse-grained model including memory or nonlocality (in the form of polyadic coupling terms) would be effective. Data-driven discovery of coarse-grained dynamical models including memory is discussed in [56]. Polyadic coupling terms make no conceptual difference, but combinatorial explosion may make their inclusion computationally costly.

Finally we remark on some physical principles that inform our methods for inferring coarse-grained dynamical models. In the coarse-grained equations arising from the OA

ansatz, the coefficient of the linear term has negative real part, and the coupling matrix $N_\sigma K^{\sigma\sigma'}/2N$ is real. These conditions together imply that the C-fold product of the unit disk is invariant; in other words, if all complex phases z_σ initially have magnitude not greater than 1, then they will continue to have magnitude not greater than 1 for all time. This condition breaks down if we allow the coupling matrix to have an imaginary part or the coefficient of the linear term to have positive real part, and so we impose those constraints during our inference procedure.

Those constraints can be thought of in a more general context. We know that the coarse-grained variables z_σ should always remain bounded in the unit disk, because they are averages of quantities on the unit circle. Knowing this, we can perform *constrained* inference of the right-hand side of an ODE governing those variables, with the constraint being that a function is feasible only if the corresponding ODE leaves the appropriate set in state space invariant.

Such a constraint may appear unruly. Notice, however, that we only need to impose that the vector field never points outward at the boundary. If we alternately write the system in terms of amplitudes and phases, that condition is simply $\dot{r}_\sigma \leq 0$ whenever $r_\sigma = 1$ and $r_{\sigma'} \leq 1$ for all σ'. Moreover, this constraint is convex! That means that if we construct a basis of functions $f_1, \ldots, f_L$ that all obey the disk-invariance condition, then any convex combination of those basis functions will obey it as well. Given such a basis, we can fit $\dot{z} = a_1 f_1(z) + \cdots + a_L f_L(z)$ with $a_i \geq 0$, and rest assured that the resulting right-hand side will obey the constraint that the coarse variables z remain bounded in the appropriate way.

5.2.3 Degree-targeting Cascades

Finally, the framework of degree-targeted seeding of discrete processes on modular networks has several possible extensions and applications. We presently only apply our framework to the Watts model on a two-module network, while it is applicable much more broadly. Certainly varying network topologies would be very interesting, but perhaps more expansive would be consideration of other dynamics. For example, we consider only the case of a fixed threshold for the entire network. We could relax this assumption

to allow for a distribution of thresholds, or a number threshold as opposed to a fractional threshold. Degree-targeting is also much more general than the case of maximum-degree seeding that we consider here. The framework we present could also be used, for example, to analyze targeting of low-degree nodes. Lastly, we focus primarily here on the asymptotic state of the network, while we could also look in more detail at transient dynamics to uncover information about the routes over which cascades tend to travel.

References

[1] Juan A. Acebrón, L. L. Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137–185, 2005.

[2] Robert Adler. A Study of Locking Phenomena in Oscillators. *Proceedings of the IRE*, 34(6):351–357, jun 1946.

[3] Réka Albert and Hawoong Jeong. Diameter of the World-Wide Web Growth dynamics of the World-Wide Web. *Nature communications*, 401(September):398–399, 1999.

[4] TM Antonsen Jr, RT Faghih, M Girvan, E Ott, and J Platig. External periodic driving of large systems of globally coupled phase oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(3):037112, 2008.

[5] Alexandre Arenas and A. Díaz-Guilera. Synchronization and modularity in complex networks. *European Physical Journal: Special Topics*, 143(1):19–25, 2007.

[6] Alexandre Arenas, Albert Díaz-Guilera, and Conrad J. Pérez-Vicente. Synchronization processes in complex networks. *Physica D: Nonlinear Phenomena*, 224(1-2):27–34, 2006.

[7] Alexandre Arenas, Albert Díaz-Guilera, and Conrad J. Pérez-Vicente. Synchronization reveals topological scales in complex networks. *Physical Review Letters*, 96(11):1–4, 2006.

[8] W Atmar and B D Patterson. The measure of order and disorder in the distribution of species in fragmental habitat. *Oecologia*, 96:373–382, 1993.

[9] A. C. Barnes, R. C. Roberts, N. C. Tien, C. A. Zorman, and Philip X L Feng. Silicon carbide (SiC) membrane nanomechanical resonators with multiple vibrational modes. *2011 16th International Solid-State Sensors, Actuators and Microsystems Conference, TRANSDUCERS'11*, (d):2614–2617, 2011.

[10] Ernest Barreto, Brian Hunt, Edward Ott, and Paul So. Synchronization in networks of networks: The onset of coherent collective behavior in systems of interacting populations of heterogeneous oscillators. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 77(3):1–7, 2008.

[11] Jordi Bascompte, P. Jordano, C. J. Melian, and Jens M. Olesen. The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences*, 100(16):9383–9387, 2003.

[12] Eric Brown, Philip Holmes, and Jeff Moehlis. Globally coupled oscillator networks. In *Perspectives and Problems in Nolinear Science*, pages 183–215. Springer, 2003.

[13] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, apr 2016.

[14] C. Campbell, S. Yang, R. Albert, and K. Shea. A network model for plant-pollinator community assembly. *Proceedings of the National Academy of Sciences*, 108(1):197–202, jan 2011.

[15] Lauren M. Childs and Steven H. Strogatz. Stability diagram for the forced Kuramoto model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(4):43128, 2008.

[16] ShiNung Ching, Emery N Brown, and Mark A Kramer. Distributed control in a mean-field cortical network model: implications for seizure suppression. *Physical Review E*, 86(2):021920, 2012.

[17] Reuven Cohen, Keren Erez, Daniel Ben-Avraham, and Shlomo Havlin. Resilience of the Internet to Random Breakdowns. *Physical Review Letters*, 85(21), dec 2000.

[18] Reuven Cohen, Keren Erez, Daniel Ben-Avraham, and Shlomo Havlin. Breakdown of the Internet under Intentional Attack. *Physical Review Letters*, 86(16):3682–3685, apr 2001.

[19] Gheorghe Craciun and Martin Feinberg. Multiple Equilibria in Complex Chemical Reaction Networks: II. The Species-Reaction Graph. *SIAM Journal on Applied Mathematics*, 66(4):1321–1338, jan 2006.

[20] Hiroaki Daido. Lower critical dimension for populations of oscillators with randomly distributed frequencies: A renormalization-group analysis. *Physical Review Letters*, 61(2):231–234, 1988.

[21] Vasilis Dakos and Jordi Bascompte. Critical slowing down as early warning for the onset of collapse in mutualistic communities. *Proceedings of the National Academy of Sciences*, 111(49):17546–17551, 2014.

[22] Albert Diaz-Guilera. Dynamics towards synchronization in hierarchical networks. *Journal Of Physics A-Mathematical And Theoretical*, 41(22):224007, 2008.

[23] Eusebius J. Doedel, Thomas F. Fairgrieve, Björn Sandstede, Alan R. Champneys, Yuri A. Kuznetsov, and Xianjun Wang. Auto-07p: Continuation and bifurcation software for ordinary differential equations. Technical report, 2007.

[24] Florian Dörfler and Francesco Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014.

[25] Leonardo Dueñas-Osorio and Srivishnu Mohan Vemuru. Cascading failures in complex infrastructure systems. *Structural Safety*, 31(2):157–167, mar 2009.

[26] D. Efimov, P. Sacré, and R. Sepulchre. Controlling the Phase of an Oscillator: A Phase Response Curve Approach. In *Joint 48th Conference on Decision and Control*, pages 7692–7697, December 2009.

[27] Bard Ermentrout. Type I membranes, phase resetting curves, and synchrony. *Neural computation*, 8(5):979–1001, 1996.

[28] Diego Garlaschelli, F. den Hollander, J. M. Meylahn, and Benthen Zeegers. Synchronization of Phase Oscillators on the Hierarchical Lattice. *Journal of Statistical Physics*, 174(1):188–218, jan 2019.

[29] James P. Gleeson. Cascades on correlated and modular random networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 77(4):1–10, 2008.

[30] Jesús Gómez-Gardeñes, Yamir Moreno, and Alexandre Arenas. Paths to synchronization on complex networks. *Physical review letters*, 98(3):34101, 2007.

[31] Levi B Good, Shivkumar Sabesan, Steven T Marsh, Kostas Tsakalis, David Treiman, and Leon Iasemidis. Control of synchronization of brain dynamics leads to control of epileptic seizures in rodents. *International journal of neural systems*, 19(03):173–196, 2009.

[32] Georg A. Gottwald. Model reduction for networks of coupled oscillators. *Chaos*, 25(5), 2015.

[33] Bruce Hannon. The structure of ecosystems. *Journal of Theoretical Biology*, 41(3):535–546, oct 1973.

[34] Frank C Hoppensteadt and Eugene M Izhikevich. *Weakly connected neural networks*, volume 126. Springer Science & Business Media, 2012.

[35] J. Hunter and J. Milton. Amplitude and Frequency Dependence of Spike Timing: Implications for Dynamic Regulation. *Journal of Neurophysiology*, 90:387–394, 2003.

[36] 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.

[37] Samuel Johnson, Virginia Domínguez-García, and Miguel A. Muñoz. Factors Determining Nestedness in Complex Networks. *PLoS ONE*, 8(9):e74025, sep 2013.

[38] Walid Klibi, Alain Martel, and Adel Guitouni. The design of robust value-creating supply chain networks: A critical review. *European Journal of Operational Research*, 203(2):283–293, 2010.

[39] Oleg Kogan, Jeffrey L. Rogers, M. C. Cross, and G. Refael. Renormalization group approach to oscillator synchronization. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 80(3):1–12, 2009.

[40] P Kundur, NJ Balu, and MG Lauby. *Power system stability and control*. 1994.

[41] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, volume 39, pages 420–422, Berlin/Heidelberg, 1975. Springer, Springer-Verlag.

[42] Yoshiki Kuramoto. Cooperative Dynamics of Oscillator Community. *Progress of Theoretical Physics Supplement*, 79(0):223–240, 1984.

[43] Yoshiki Kuramoto. *Chemical oscillations, waves, and turbulence*. Springer Science & Business Media, 2012.

[44] Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi. Benchmark graphs for testing community detection algorithms. *Physical Review E*, 78(4):046110, oct 2008.

[45] Erik D. Lumer and Bernardo A. Huberman. Hierarchical dynamics in large assemblies of interacting oscillators. *Physics Letters A*, 160(3):227–232, 1991.

[46] Renato E. Mirollo and Steven H. Strogatz. Jump Bifurcation and Hysteresis in an Infinite-Dimensional Dynamical System of Coupled Spins. *SIAM Journal on Applied Mathematics*, 50(1):108–124, feb 1990.

[47] Hiroya Nakao. Phase reduction approach to synchronisation of nonlinear oscillators. *Contemporary Physics*, 57(2):188–214, 2016.

[48] Azadeh Nematzadeh, Emilio Ferrara, Alessandro Flammini, and Yong Yeol Ahn. Optimal network modularity for information diffusion. *Physical Review Letters*, 113(8):1–5, 2014.

[49] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[50] E. Oh, K. Rho, H. Hong, and B. Kahng. Modular synchronization in complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 72(4):1–4, 2005.

[51] Jens M. Olesen, Jordi Bascompte, Y. L. Dupont, and P. Jordano. The modularity of pollination networks. *Proceedings of the National Academy of Sciences*, 104(50):19891–19896, 2007.

[52] Edward Ott and Thomas M. Antonsen. Low dimensional behavior of large systems of globally coupled oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(3):37113, 2008.

[53] Edward Ott and Thomas M. Antonsen. Long time evolution of phase oscillator systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(2):023117, jun 2009.

[54] Edward Ott, Brian R. Hunt, and Thomas M. Antonsen. Comment on "Long time evolution of phase oscillator systems" [Chaos 19 , 023117 (2009)]. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(2):025112, jun 2011.

[55] Bertrand Ottino-Löffler and Steven H. Strogatz. Kuramoto model with uniformly spaced frequencies: Finite-N asymptotics of the locking threshold. *Physical Review E*, 93(6):062220, jun 2016.

[56] Shaowu Pan and Karthik Duraisamy. Data-driven Discovery of Closure Models. 17(4), 2018.

[57] Diego Pazó. Thermodynamic limit of the first-order phase transition in the Kuramoto model. *Physical Review E*, 72(4):046211, oct 2005.

[58] Noa Pinter-Wollman, Elizabeth A. Hobson, Jennifer E. Smith, Andrew J. Edelman, Daizaburo Shizuka, S. de Silva, James S. Waters, Steven D. Prager, Takao Sasaki, George Wittemyer, Jennifer Fewell, and David B. McDonald. The dynamics of animal social networks: analytical, conceptual, and theoretical advances. *Behavioral Ecology*, 25(2):242–255, mar 2014.

[59] Keith G Provan and H. Brinton Milward. Do Networks Really Work? A Framework for Evaluating Public-Sector Organizational Networks. *Public Administration Review*, 61(4):414–423, jul 2001.

[60] B. Razavi. A study of injection pulling and locking in oscillators. *Proceedings of the IEEE 2003 Custom Integrated Circuits Conference, 2003.*, 39(9):1415–1424, 2003.

[61] Serguei Saavedra, Felix Reed-Tsochas, and Brian Uzzi. A simple model of bipartite cooperation for ecological and organizational networks. *Nature*, 457(7228):463–466, 2009.

[62] Serguei Saavedra, Rudolf P. Rohr, Vasilis Dakos, and Jordi Bascompte. Estimating the tolerance of species to the effects of global environmental change. *Nature Communications*, 4:1–6, 2013.

[63] H. Sakaguchi. Cooperative Phenomena in Coupled Oscillator Systems under External Fields. *Progress of Theoretical Physics*, 79(1):39–46, jan 1988.

[64] Michael A Schwemmer and Timothy J Lewis. The theory of weakly coupled oscillators. In *Phase response curves in neuroscience*, pages 3–31. Springer, 2012.

[65] Per Sebastian Skardal and Juan G. Restrepo. Hierarchical synchrony of phase oscillators in modular networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85(1):1–8, 2012.

[66] 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, sep 2000.

[67] Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Westview Press, 2001.

[68] Hisa-Aki Tanaka. Entrainment limit of weakly forced nonlinear oscillators. In *Mathematical Approaches to Biological Systems*, pages 77–93. Springer, 2015.

[69] David G Tarboton, Rafael L Bras, and Ignacio Rodriguez-Iturbe. The fractal nature of river networks. *Water Resources Research*, 24(8):1317–1322, aug 1988.

[70] Xianjun Wang, Eusebius J. Doedel, Thomas F. Fairgrieve, Björn Sandstede, Alan R. Champneys, and Yuri A. Kuznetsov. AUTO-07p: Continuation and bifurcation software for ordinary differential equations.

[71] Duncan J. Watts. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci.*, 99(9):5766–5771, 2002.

[72] G. B. West. A General Model for the Origin of Allometric Scaling Laws in Biology. *Science*, 276(5309):122–126, apr 1997.

[73] Kenneth G. Wilson. Renormalization group and critical phenomena. II. Phase-space cell analysis of critical behavior. *Physical Review B*, 4(9):3184–3205, 1971.

[74] Arthur T Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of theoretical biology*, 16(1):15–42, 1967.

[75] Arthur T Winfree. *The geometry of biological time*, volume 12. Springer Science & Business Media, 2001.

[76] Anatoly Zlotnik, Yifei Chen, István Z. Kiss, Hisa-Aki Tanaka, and Jr-Shin Li. Optimal Waveform for Fast Entrainment of Weakly Forced Nonlinear Oscillators. *Physical Review Letters*, 111(2):024102, jul 2013.

[77] Anatoly Zlotnik and Jr-Shin Li. Optimal entrainment of neural oscillator ensembles. *Journal of Neural Engineering*, 9(4):046015, 2012.

[78] Anatoly Zlotnik and Jr-Shin Li. Optimal Subharmonic Entrainment of Weakly Forced Nonlinear Oscillators. *SIAM Journal on Applied Dynamical Systems*, 13(4):1654–1693, jan 2014.

[79] Anatoly Zlotnik, Raphael Nagao, István Z. Kiss, and Jr-Shin Li. Phase-selective entrainment of nonlinear oscillator ensembles. *Nature communications*, 7, 2016.

[80] Robert Zwanzig. Memory effects in irreversible thermodynamics. *Physical Review*, 124(4):983–992, 1961.