Claremont Colleges
Scholarship @ Claremont

HMC Senior Theses

HMC Student Scholarship

2016

Topological Data Analysis for Systems of Coupled Oscillators

Alec Dunton
Harvey Mudd College

Recommended Citation

Dunton, Alec, "Topological Data Analysis for Systems of Coupled Oscillators" (2016). *HMC Senior Theses*. 79.
https://scholarship.claremont.edu/hmc_theses/79

This Open Access Senior Thesis is brought to you for free and open access by the HMC Student Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in HMC Senior Theses by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

Topological Data Analysis for Systems of Coupled Oscillators

Alec Dunton

Andrew Bernoff, Advisor

Chad Topaz, Reader

HARVEY
MUDD
COLLEGE

Department of Mathematics

May, 2016

Copyright © 2016 Alec Dunton.

The author grants Harvey Mudd College and the Claremont Colleges Library the nonexclusive right to make this work available for noncommercial, educational purposes, provided that this copyright statement appears on the reproduced materials and notice is given that the copying is by permission of the author. To disseminate otherwise or to republish requires written permission from the author.

Abstract

Coupled oscillators, such as groups of fireflies or clusters of neurons, are found throughout nature and are frequently modeled in the applied mathematics literature. Earlier work by Kuramoto, Strogatz, and others has led to a deep understanding of the emergent behavior of systems of such oscillators using traditional dynamical systems methods. In this project we outline the application of techniques from topological data analysis to understanding the dynamics of systems of coupled oscillators. This includes the examination of partitions, partial synchronization, and attractors. By looking for clustering in a data space consisting of the phase change of oscillators over a set of time delays we hope to reconstruct attractors and identify members of these clusters.

Contents

Abstract iii

Acknowledgments ix

1 Introduction 1
- 1.1 Coupled Oscillators 1
- 1.2 Topological Data Analysis 5

2 Synchronization in Kuramoto Oscillators 11
- 2.1 The Kuramoto Model 11
- 2.2 Introducing Noise to Kuramoto 15
- 2.3 Linearization about the Fixed Point 17
- 2.4 Synchronization and Ornstein-Uhlenbeck 20
- 2.5 A Hierarchy of Models 22

3 Clusters and Topological Signatures 27
- 3.1 Betti Zero 27
- 3.2 From One Cluster to Many 29
- 3.3 Cluster Identification 33
- 3.4 Efficacy of Our Cluster Identification Method 35
- 3.5 Cluster Detection under Nonlocal Coupling 38

4 Conclusions and Future Directions 41

Bibliography 43

List of Figures

1.1 A nighttime photograph of fireflies flashing in synchrony. (Taken from smithsonianmag.com) 2
1.2 A sample of different network structures 3
1.3 The construction of a simplicial complex in 2-space. 6
1.4 The color coded contour plot shown display the value of b_0 for a given pair (ϵ, t). 9

2.1 A plot of the phases of the oscillators governed by the Kuramoto model (note the invariance under addition of 2π). . 12
2.2 A plot of the phases of the oscillators governed by the Kuramoto model with $K = 2K_{crit}$ 13
2.3 A plot of the phases of the oscillators governed by the Kuramoto model with $K = 20K_{crit}$. 13
2.4 A plot of the phases of the oscillators governed by the Kuramoto model with $K = 200K_{crit}$. 14
2.5 A plot of the phases of the oscillators governed by the Kuramoto model with noise added as described. 15
2.6 The dependence of the spread of steady-state phases on the variance of white noise D, plotted on top of the function $11\sqrt{t}$. 16
2.7 A plot of the phases of the oscillators governed by equation 2.2. 17
2.8 A plot of the CDF of the phases of the oscillators governed by equation 2.2. 22
2.9 A plot of the phases of the oscillators governed by equation 2.3. 23
2.10 A plot of the phase distribution of the oscillators governed by equation 2.3. 24
2.11 A plot of the phases of the oscillators governed by equation 2.4. 25

2.12 A stationary plot of the phases of the oscillators governed by equation 2.4. 26

3.1 A plot of phases v. time for Kuramoto oscillators with one cluster built into the network structure (also known as all-to-all coupling). 29
3.2 A plot of phases v. time for Kuramoto oscillators with two clusters built into the network structure. 30
3.3 A plot of phases v. time for Kuramoto oscillators with four clusters built into the network structure. 33
3.4 A plot of B_0 versus the value of τ chosen. 34
3.5 Output Data Denoting Clusters Generated for a fixed value of Tau 34
3.6 Topological Barcode Plotted in terms of Tau 35
3.7 Dependence of Cluster Detection Accuracy on the Noise Amplitude 36
3.8 Dependence of Cluster Detection Accuracy on the Variance of ω and the Number of Oscillators 37
3.9 Kuramoto Oscillators with Weak Spatial Coupling 39
3.10 Kuramoto Oscillators with Strong Spatial Coupling 39

Acknowledgments

Thank you to Professor Bernoff for being an extremely helpful and patient advisor during this process. I would also like to thank Chad Topaz for his input and support as my second reader, as well as professors Thompson and Ziegelmeier for their guidance and support during the Harvey Mudd-Macalester REU last summer. Finally, I would like to thank all of my family and friends for their support during the past 4 years here at Harvey Mudd.

Chapter 1

Introduction

1.1 Coupled Oscillators

Oscillators, specifically coupled oscillators, are found throughout nature and an active area of research in the applied mathematics community. Examples of these oscillators include clustered groups of firing neurons and croaking frogs. When we say that a group of oscillators are *coupled*, we mean that each member's behavior influences that of the others in the system. In a biological context, we would say that fireflies exhibit this behavior. This is because a firefly's flashing causes other fireflies to light up in response, which leads to situation like that displayed in Figure 1.1. Numerous models have been developed to explain the behavior of systems of coupled oscillators. The most famous of these is that of Yoshiki Kuramoto, who developed the following system in 1975 (see (Kuramoto, 1975), (Strogatz, 2000), and (Acebrón et al., 2005)):

$$\dot{\theta}_i = \omega_i + \frac{1}{N}\sum_{j=1}^{N} K_{ij}\sin(\theta_j - \theta_i) \tag{1.1}$$

which is most commonly seen in the simplified form where $K_{ij} = K$:

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i) \tag{1.2}$$

The key components of this model include the coupling constant matrix with entries K_{ij}, the natural frequency of the i^{th} oscillator ω_i, and the coupling function $\sin(\psi)$. The coupling constant matrix with entries K_{ij} is an N-dimensional matrix which encodes the coupling strength K_{ij} of the j^{th}

Figure 1.1 A nighttime photograph of fireflies flashing in synchrony. (Taken from smithsonianmag.com)

oscillator on the i^{th} oscillator. In the symmetric case, the coupling constant K measures the strength of the influence of oscillators on one another: the larger K is, the stronger the coupling force between them. Here ω_i is an intrinsic property of each oscillator, like the tempo at which a metronome is set. The coupling function $\sin(\psi)$ represents the nature of the coupling between oscillators, and is the source of the synchronization observed in simulations of the model. Notice that when $\theta_j > \theta_i$, or $\psi > 0$, the coupling force exerted on oscillators i is positive. When $\psi < 0$, the coupling force is negative. This tends to make oscillators either 'catch up' to or 'wait up' for other oscillators, thereby leading to synchronization. (Acebrón et al., 2005)

For the purpose of extracting more exotic behavior from the Kuramoto model one can alter it. A common manner in which this is done involves changing the network structure of the oscillators. In the original model, Kuramoto assumed an all-to-all network structure, where every oscillator is coupled to every other oscillator. We can rewrite the model, accounting for the possibility of various network structures, such as those displayed in Figure 1.2 (see (Collins and Stewart, 1994) and (Easley and Kleinberg,

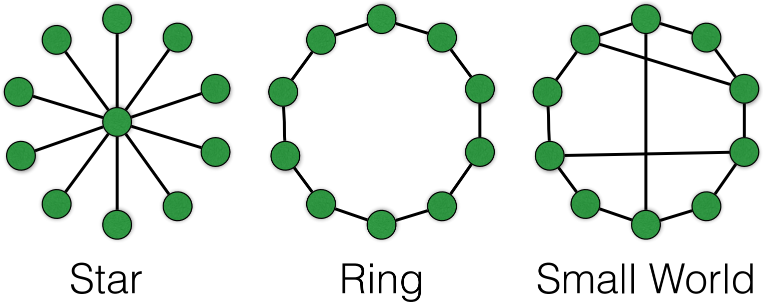


Figure 1.2 A sample of different network structures

2010)):

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \Delta_{ij} \sin(\theta_j - \theta_i)$$

where Δ_{ij} is an adjacency matrix which encodes the structure of the network. In the case of a ring network (see Figure 1.2), the matrix Δ may look like this:

$$\begin{bmatrix} 1 & 1 & 0 & \dots & 1 \\ 1 & 1 & 1 & \dots & 0 \\ 0 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 1 \end{bmatrix}$$

where in this case we have implemented the 'nearest neighbor' structure seen in Figure 1.2. That is, on the ring, two members are coupled only if they are immediately next to one another. In other networks, such as star or small-world, this adjacency matrix looks more complicated.

Perhaps the most relevant variation of the original Kuramoto model for our research is the partitioning of a network. A partition of a network is a set of sub-networks which are disjoint and whose union contains the original network. In terms of an adjacency matrix, this can be visualized as a block diagonal matrix whose blocks are all-ones matrices. These would look like

$$\begin{bmatrix} [1]_1 & 0 & 0 & \dots & 0 \\ 0 & [1]_2 & 0 & \dots & 0 \\ 0 & 0 & [1]_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & [1]_m \end{bmatrix}$$

where each [1] is a square matrix with all one entries. It is important to note that each [1] may be of any size less than or equal to the total number of oscillators in the system.

On top of varying the network structure in the Kuramoto model there are other variations including continuum, non-local coupling, and pulse-coupled models. The continuum model for Kuramoto oscillators prescribes a density in the place of a system of equations which model individual oscillators. The density of the oscillators in this variation of the original model obeys the Fokker-Planck equation. In a thorough review paper on the Kuramoto model, Stephen Strogatz outlines numerous characteristics of the model, including the steady-state of the model when white noise is added, as will be in Section 2.2. (Strogatz, 2000) A key result from the paper with respect to the stochastic Kuramoto model is that in limit as $N \to \infty$, the density function of the oscillators $\rho(\theta, t; \omega)$ satisfies the Fokker-Planck equation, which is defined as:

$$\frac{d\rho}{dt} = D\frac{D^2\rho}{d\theta^2} - \frac{d}{d\theta}(\rho v)$$

where

$$v(\theta, t; \omega) = \omega + Kr\sin(\psi - \theta)$$

and $re^{i\psi}$ is the order parameter, defined as $\frac{1}{N}\sum\limits_{j=1}^{N} e^{i\theta_j}$

This fact is extremely important when discussing the properties of clusters of oscillators and distributions of the phases of the oscillators within such clusters. (see (Strogatz and Mirollo, 1988), (Strogatz, 2000), (Balmforth and Sassi, 2000) and (Acebrón et al., 2005)). Furthermore, the relationship between the phase behavior of Kuramoto oscillators and Ornstein-Uhlenbeck processes is elucidated by this connection between Fokker-Planck and the phase density function ρ. This is discussed in detail in Section 2.4.

Non-local coupling models incorporate a spatial variable, usually labeled x, which is used to construct a coupling function, usually denoted $G(x_i, x_j)$, where x_i and x_j are the spatial positions of the i^{th} and j^{th} oscillators. Examples of coupling function include exponential decay models, such as $e^{-|x_i - x_j|}$, where the strength of coupling decreases as spatial distance between oscillators increases. (Kuramoto and Battogtokh, 2002) We will refer to these models in Section 3.5.

A final extension of the Kuramoto model includes the incorporation of time delays. Similar to spatial coupling functions, time delays are added to the model to better mimic the behavior of coupled oscillatory systems in nature - for example, oscillators whose coupling relies on sound stimulus, and is therefore dependent on the phase of surrounding oscillators at a previous time. (Yeung and Strogatz, 1999) This variation of the Kuramoto model may be of interest after we have thoroughly examined the original model.

1.2 Topological Data Analysis

At a high level, topological data analysis can be described as the study of the shape of data. In the context of coupled oscillators, it allows us to detect emergent behaviors such as clustering and synchronization. This is done by examining data from the phase-space of the system and approximating its structure with a topological object.

The foundational concepts which make topological data analysis possible lie within the field of algebraic topology. By constructing simplicial complexes and deriving their related homology, we can determine the topological invariants of the objects which we use to represent a set of data such as Betti numbers. Betti numbers are numbers which represent the number of connected components, topological circles, trapped volumes, etc. in a topological object. They are extremely useful in distinguishing different objects, as every topological object is uniquely identifiable by a sequence of Betti numbers.

We first consider a set of data points to be a sampling of a large space. (Topaz et al., 2015) We then connect points within this space if they are close, where closeness is defined in terms of a *proximity parameter*, usually denoted ϵ, giving us a *simplicial complex*. By varying ϵ, we can observe

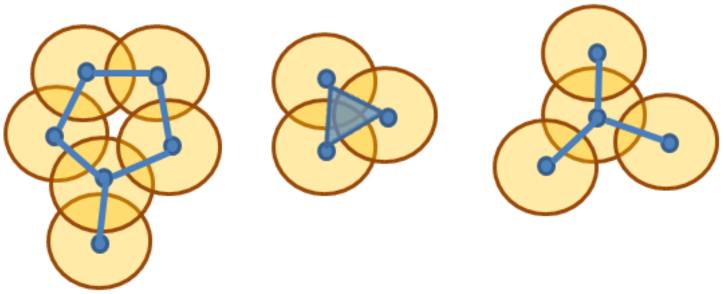


Figure 1.3 The construction of a simplicial complex in 2-space.

which characteristics of the space persist for the largest range of values. This gives us clues into the shape of the space from which our data is sampled. A visualization of this process is shown in Figure 1.3. Note that a simplicial complex is composed of points, lines, triangles, tetrahedra, and so forth. It is important to note that this definition of a simplicial complex is known as a *Vietoris-Rips* complex.

The formal definition of a simplicial complex is a set consisting of a finite collection of k-simplices, where a 0-simplex is a vertex, a 1-simplex is an edge, a 2-simplex is a triangle, a 3-simplex is a tetrahedron, etc. A simplex, which we denote S, satisfies two properties: for every set σ in S, every non-empty subset $\tau \subseteq \sigma$ is in S, and two k-simplices are either disjoint or have a non-empty intersection in a lower dimensional simplex. (Edelsbrunner et al., 2002) The importance of simplicial complexes is that they allow us to take point-cloud data, convert it into a topological object, and then use the concept of *Homology* to identifies characteristics of the data such as loops, components, and volumes. (Edelsbrunner et al., 2002) (Stolz, 2014)

In Figure 1.3, we have a visualization of a specific type of simplicial complex, the *Vietoris-Rips complex*. In this complex, for some $\epsilon > 0$, we have built a simplicial complex S_ϵ using the following rule: k +1 data points form a k-simplex if they all have pairwise distance less than ϵ. (Topaz et al., 2015). Note that in homology, it is necessary to ascribe an orientation on the vertices of each k-simplex. That is, a k-simplex $[v_1, v_2, ..., v_{k+1}]$ is equal to $-\sigma([v_1, v_2, ..., v_{k+1}])$ if σ is an odd permutation.

As the Vietoris-Rips method checks all pairwise distances between data points, it can be thought of as the maximal simplicial complex. (Topaz et al., 2015) One alternative construction of these complexes is called a witness complex. In a witness complex, one selects a subset of the points from the given point-cloud data called *landmark points* which are typically located in the densest regions of the Euclidean data. From there, edges are added to the complex based on their distance from landmark points, where we desire to draw edges between the two closest points to each landmark point. For a more thorough explanation, see De Silva and Carlsson (2004).

We said at the beginning of this section that the foundational ideas of topological data analysis lie in the discipline of algebraic topology. Thus far we have discussed the construction of topological spaces in the form of simplicial complexes, but have yet to give them algebraic structure. To do this, we construct an abstract vector space C_k with a basis comprised of the k-simplicies in S_ϵ. The elements in C_k are called *k-chains*. (Edelsbrunner et al., 2002)

We then proceed to define the boundary of a k-simplex. The *boundary* of a k-simplex is defined as the union of all of its (k-1)-subsimplices. For $k \geq 1$, the boundary map $\delta_k : C_k \to C_{k-1}$ is a linear transformation on a k-simplex $[v_0, v_1, ..., v_k]$ given by the formula

$$\delta_k([v_0, v_1, ..., v_k]) = \sum_{i=0}^{k} (-1)^i [v_0, ..., v_{i-1}, v_{i+1}, ..., v_k]$$

where we have obtained a (k-1)-simplex from $[v_0, v_1, ..., v_k]$ by removing v_i.

Boundary operators take the vector space C_k and turn it into a *chain complex*:

$$\ldots C_{k+1} \to C_k \to C_{k-1} \to \cdots \to C_1 \to C_0 \to 0$$

The following two subspaces of C_k are crucial to the material discussed in this paper:

k-cycles

$$Z_k := ker(\delta_k : C_k \to C_{k-1})$$

and k-boundaries

$$B_k := im(\delta_{k+1} : C_{k+1} \to C_k)$$

Thinking of these subspaces, we can now defined what it means for two cycles in Z_k to be *homologous*. Two cycles z_1, z_2 are said to be homologous

($z_1 \sim z_2$) if their difference is contained in B_k. The equivalence relation $\sim$ gives us a quotient space, which we call the kth homology of a simplicial complex, defined as

$$H_k := \{[z] | z \in Z_k\} \tag{1.3}$$

Where we can also write $H_k = Z_k / B_k$. We then define the kth Betti number, b_k, as the dimension of the kth homology, that is:

$$b_k = dim(H_k) = dim(Z_k) - dim(B_k) \tag{1.4}$$

or equivalently in linear algebraic terms:

$$b_k = [n_k - rank(\delta_k) - rank(\delta_{k+1})] \tag{1.5}$$

where b_k is equal to the number of independent holes of dimension k. For a quick example, if we were to calculate the values of the b_k's for a topological figure-eight, we would find that $b_0 = 1$, $b_1 = 2$, and $b_k = 0$ for all $k \geq 2$. This is because a figure eight is made up of one connected component, two circles, and cannot trap volumes of dimension higher than one.

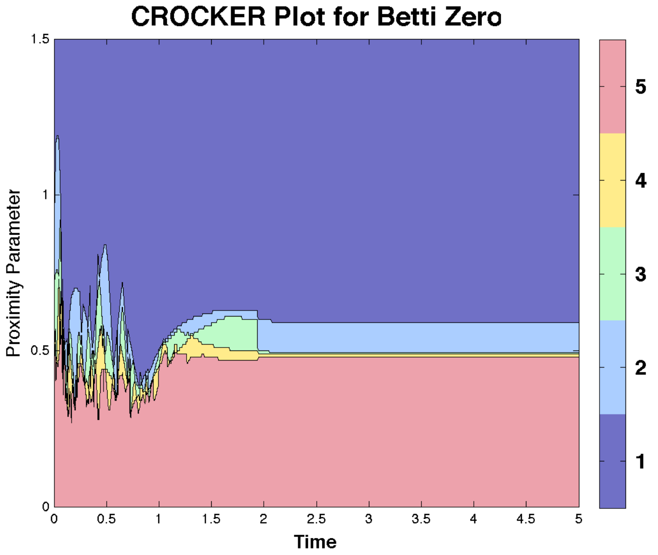
Now that we have the methodology for quantifying the topological characteristics of a space in b_k, we can discuss persistence. Up until now in our discussion we have thought of a k-simplex as a simplex constructed using a fixed proximity parameter. What we do now is take the parameter ϵ and vary it in increasing fashion. That is, for $\epsilon_1 \leq \epsilon_2 \leq \cdots \leq \epsilon_M$ we generate a sequential inclusion of complexes called a *filtration*. (Edelsbrunner et al., 2002)

$$S_{\epsilon_1} \subseteq S_{\epsilon_2} \subseteq \cdots \subseteq S_{\epsilon_M} \tag{1.6}$$

Tracking the features which persist over many values of ϵ then allows us to estimate the 'shape' of the data we are studying. A visually intuitive way in which this can be displayed is through the construction of a topological *barcode*.

A barcode encodes the topological features of a simplicial complex in relation to the proximity parameter which was used to construct it. The value of b_k for some k is represented by the number of horizontal bars intersecting a vertical line at some ϵ, which is placed on the x-axis.

A limitation of a barcode is that when working with time-dependent data, we are restricted to thinking of the data at a fixed time. In order to analyze emergent behavior of a dynamical system in terms of topological data



,

Figure 1.4 The color coded contour plot shown display the value of b_0 for a given pair (ϵ, t).

analysis, we must think of b_k as a function of time, not just ϵ. We then display this multivariable function using a contour plot, referred to by Topaz et al as a Contour Realization Of Computed k-dimensional hole Evolution in the Rips complex (CROCKER) plot. (Topaz et al., 2015) By examining a CROCKER plot for b_k, we can determine its value for some (ϵ, t). An example of a CROCKER plot is shown in Figure 1.4

Using topological data analysis, we can take data in Euclidean space and infer characteristics about its shape. We do this by defining a distance measure on the space and connecting points deemed 'close' enough to one another. After this is done, we can use ideas from homology to determine the Betti numbers of a topological object which approximates the data to transform the problem from a topological one to an algebraic one. Once we can calculate Betti numbers, we can calculate them for various values of ϵ, which defines 'closeness' in these problems. After doing this, we can graph the Betti numbers of a plot against either ϵ or both ϵ and time. (Topaz et al., 2015)

Applications of topological data analysis to dynamical systems can be found across applied mathematics publications. Of the papers addressing the subject, the most relevant is “Exploring the Topology of Dynamical Reconstructions" by Garland, Bradley, and Meiss. (Garland et al., 2015) In this paper, the authors outline a method for reconstructing the dynamics of the classic Lorenz system. They successfully reconstruct the attractor of the system without requiring diffeomorphism and take advantage of the parsimony of the witness construction method. In this thesis, we use time delays within the context of topological data analysis to examine the attractor of the system governed by the Kuramoto equations, as well as others based on pulse-coupling. Moreover, the paper by Garland et. al provides inspiration for using time delays to study the topology of coupled oscillators.

In “Topological Data Analysis of Biological Aggregation Models", Topaz et. al use the methods outlined in this paper to study the shapes which emerge from biological models such as the d’Orsogna model for swarms. In their paper the authors are able to use the topology of position-velocity data from swarms to distinguish between single and double mills - patterns typical found in schools of fish. (Topaz et al., 2015)

Chapter 2

Synchronization in Kuramoto Oscillators

2.1 The Kuramoto Model

Of all of the emergent behaviors observed in the Kuramoto Model, the most well-known is that of synchronization. (see (Balmforth and Sassi, 2000), (Strogatz, 2000), (Acebrón et al., 2005), (Chopra and Spong, 2005), and (Cumin and Unsworth, 2007)) In effect, synchronization describes the steady-state of the system in which all oscillators have the same phase velocity. This long-term velocity is equal to the average of the natural frequencies:

$$\lim_{t \to \infty} \dot{\theta}(t) = \frac{1}{N} \sum_{i=0}^{N} \omega_i = \overline{\omega}$$

In the original model, Kuramoto assumed that $\overline{\omega} = 0$. This would tell us that, in the limit, oscillators tend to stop moving. It is important to note that this holds for this model when all-to-all coupling governs the system. In other cases, such as when coupling is local or sections of the network topology are disjoint, this changes. Of the analytic results of the Kuramoto model present in the literature, one of the most relevant to this paper is the notion of a critical coupling constant, usually denoted K_{crit}. From "On the critical coupling for Kuramoto oscillators" we have that this value is sufficiently large when it is greater $\frac{N(\omega_{max} - \omega_{min})}{2(N-1)}$ where N is the number of oscillators in the system. (Dörfler and Bullo, 2011) If the system is partitioned into separate components, then this critical value of K also depends on the number of such components. In this paper, we use the threshold value of

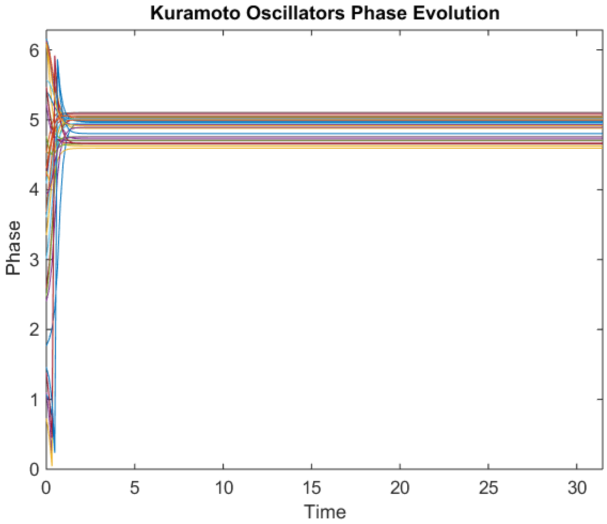


Figure 2.1 A plot of the phases of the oscillators governed by the Kuramoto model (note the invariance under addition of 2π).

$\frac{2M}{\omega_{max}-\omega_{min}}$, where M is the number of components in the given partition of the network. This value of K is sufficiently large to lead to synchronization, and is the value we refer to as K_{crit} in this thesis.

If the exact value of the critical coupling constant is sought, it can be solved implicitly using the following set of equations (Dörfler and Bullo, 2011):

$$K_{crit} = nu^* / \sum_{i=1}^{N} \frac{1}{\sqrt{1-(\frac{\Omega_i}{u^*})^2}}$$

where $\Omega_i = \omega_i - \frac{1}{N}\sum_{j=1}^{N}\omega_j$ and $u^* \in [\|\Omega\|_\infty, 2\|\Omega\|_\infty]$ is the unique solution to

$$2\sum_{i=1}^{N}\sqrt{1-(\frac{\Omega_i}{u^*})^2} = \sum_{i=1}^{N}\frac{1}{\sqrt{1-(\frac{\Omega_i}{u^*})^2}}$$

It is important to reflect on the definition of synchronization and note that it does not imply that the phases of the oscillators are all the same as $t \to \infty$. This situation, referred to as global synchronization (Acebrón et al., 2005),

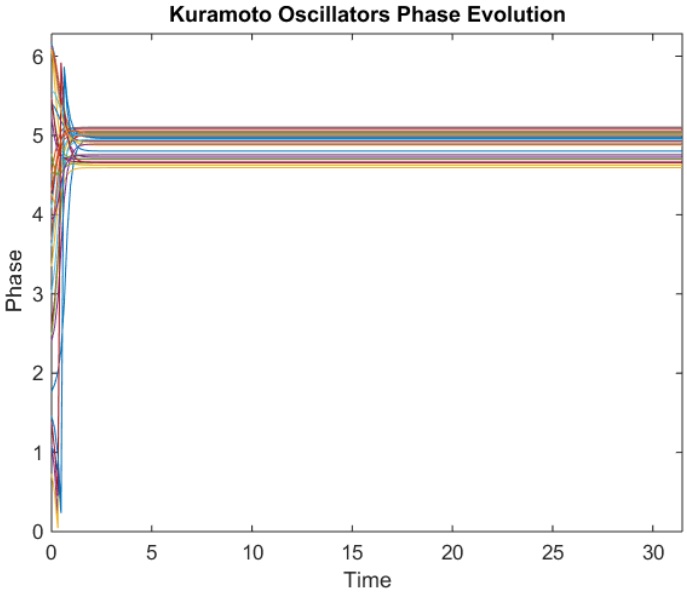


Figure 2.2 A plot of the phases of the oscillators governed by the Kuramoto model with $K = 2K_{crit}$

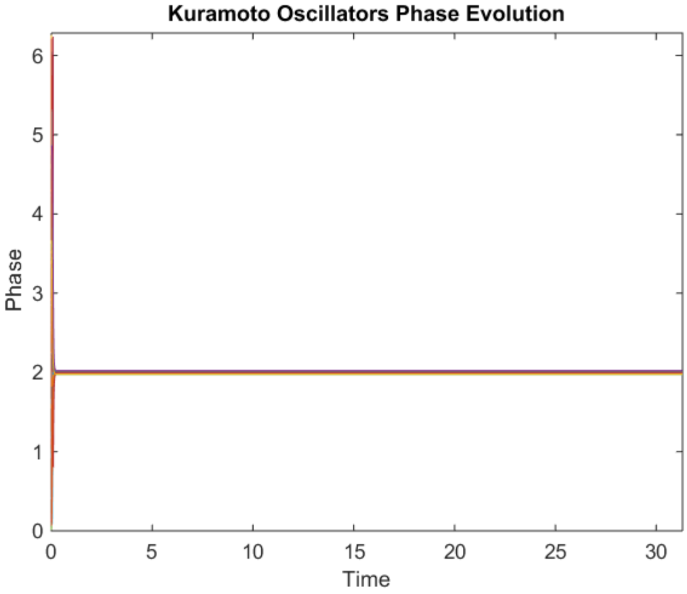


Figure 2.3 A plot of the phases of the oscillators governed by the Kuramoto model with $K = 20K_{crit}$.

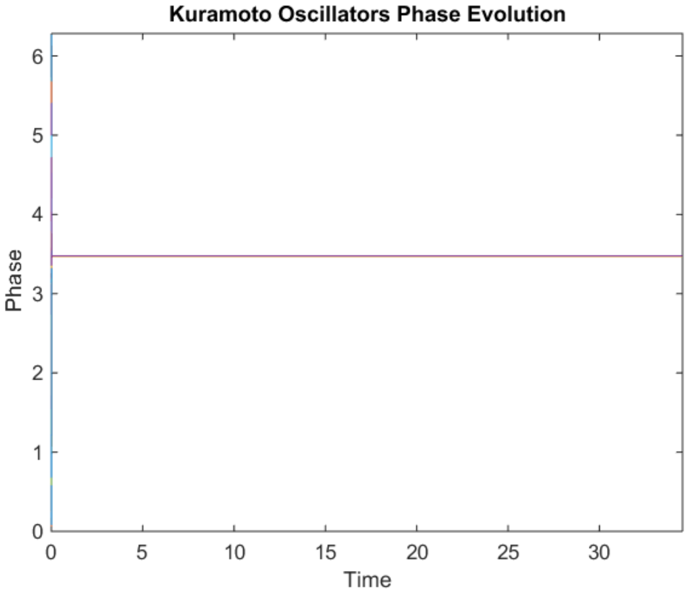


Figure 2.4 A plot of the phases of the oscillators governed by the Kuramoto model with $K = 200K_{crit}$.

occurs only when $K \to \infty$ or when all of the oscillators have the same natural frequency. When we refer to synchronization in this paper, we are referring to the case in which the phase velocities of oscillators converge.

When oscillators synchronize under this definition, their phase velocities and phases lie within respective distributions. That is to say, the angular velocities of each oscillator does not actually reach zero. Moreover, the phases of these oscillators move in time with a banded structure, demonstrated in Figure 2.1. We now discuss the distribution of these phases after sufficient time has elapsed. Examining the plot in Figure 2.1, we see that they seem to obey some sort of distribution, specifically the group of oscillators centered approximately about 4.5. We let the spread of the phases of the oscillators be denoted σ_θ.

Because phase locking occurs when $K \to \infty$, we should expect that the spread of the distribution is related to K in an inverse fashion. That is, $\sigma_\theta \propto \frac{1}{K^\alpha}$. We can see this visually when we let $K = 2K_{crit}$, $K = 20K_{crit}$, and $K = 200K_{crit}$ successively in Figures 2.2, 2.3, and 2.4.

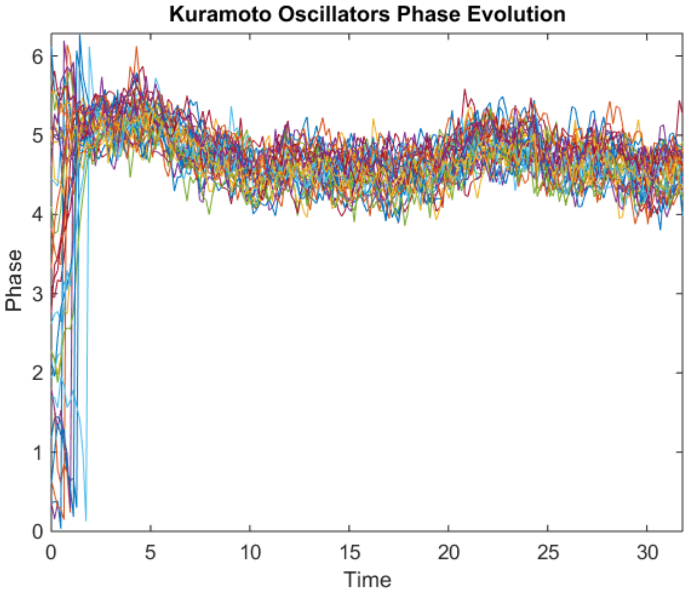


Figure 2.5 A plot of the phases of the oscillators governed by the Kuramoto model with noise added as described.

Evidently, increasing the value of K leads to a decrease in the spread of the distribution of phases after a few characteristic times have elapsed, where we define a *characteristic time* to be equal to $\frac{2\pi}{\overline{\omega}+1}$.

2.2 Introducing Noise to Kuramoto

The Kuramoto model is extremely elegant and therefore ideal for traditional analysis. However, there have been over 1000 papers written on it in which stochasticity is incorporated. For these variations of the model, typical dynamical systems methods break down. A standard way to introduce noise to this model is to simply rewrite it as

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=0}^{N}\sin(\theta_j - \theta_i) + \xi_i(t)$$

where $\xi(t)$ is an independent, gaussian white noise process with the properties $< \xi >= 0$ and $< \xi^2 >= 1$. In “Stochastic Runge-Kutta algorithms I. White Noise", Honeycutt presents a second order numerical scheme which

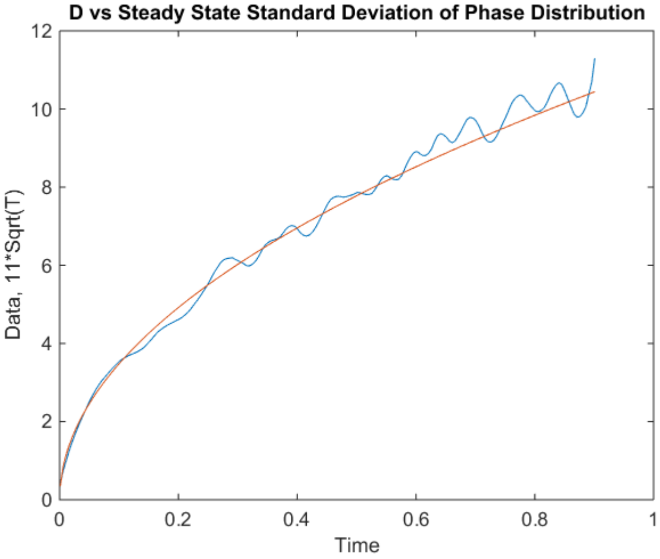


Figure 2.6 The dependence of the spread of steady-state phases on the variance of white noise D, plotted on top of the function $11\sqrt{t}$.

can be used to simulate the Kuramoto model with noise added. (Honeycutt, 1992) Implementing the stochastic Kuramoto model, we use a variation of the Runge Kutta order 2 method:

$$\begin{aligned}
x(T+\Delta t) &= x(T) + \frac{1}{2}\Delta t(F_1 + F_2) + \psi\sqrt{2D\Delta t} \\
F_1 &= f(x_0) \\
F_2 &= f(x_0 + \Delta t F_1 + \psi\sqrt{2D\Delta t})
\end{aligned}$$

where ψ is a random variable with variance 1 and mean 0. The error for this method is $O(\Delta t^2)$ locally, and therefore first-order globally.

Implementing this scheme, we obtain a plot of phases against time for 32 oscillators, $D = 0.1$, and $K = 4K_{crit}$, shown in Figure 2.5.

We see that even adding noise where $D = 0.1$ makes the phases of the oscillators far less coherent. This becomes much more significant when there are many partitions of the network in question. In fact, the spread of the phases over time is also proportional to the square root of the spread of the noise, which is shown in Figure 2.6.

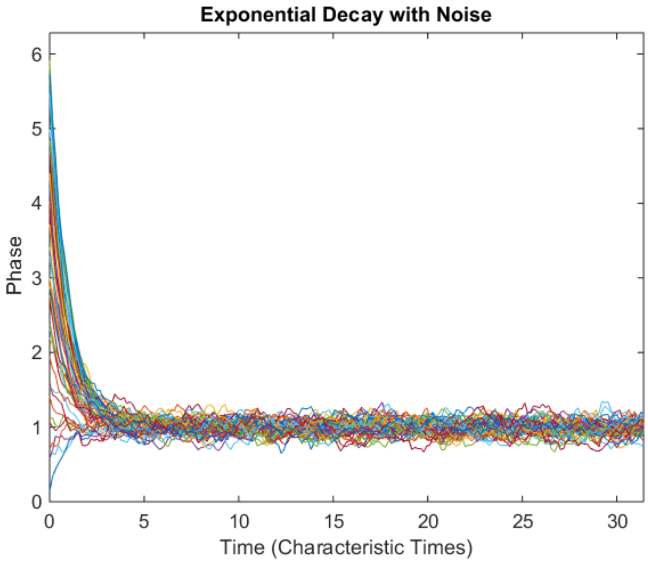


Figure 2.7 A plot of the phases of the oscillators governed by equation 2.2.

Taking a more traditional analytic approach to the stochastic Kuramoto model, Bag et. al describe the influence of noise on the order parameter of synchronization, as well as how it impacts the critical coupling threshold for the system. (Bag et al., 2007) They find that the characteristics of the noise added to the model have a substantial impact on the coupling strength require to drive the Kuramoto system to synchronization, and determine the relationship between D and both the critical value of the coupling constant and the infinite time limit value of the order parameter.

2.3 Linearization about the Fixed Point

The most famous behavior exhibited by coupled oscillators governed by the Kuramoto model is synchronization. When oscillators with identical natural frequencies synchronize, we have that their phases are all approximately the same. We also note that the angular velocity of these synchronized oscillators is equal to $\bar{\omega}$, the natural frequency of all of the oscillators.

We first observe the Kuramoto model in its original form:

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i)$$

We then make the simplifying assumption that every oscillator has the same natural frequency Ω. We then may simply use a change of variables $\theta_i = \theta_i - \Omega t$ to eliminate the frequency Ω. We can now rewrite the Kuramoto model as

$$\dot{\theta}_i = \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i)$$

When we are looking at a synchronized system, we note that for the phase of each oscillator that $\theta_i \approx \theta_j$. Thus we may once again simplify this model using a taylor approximation at the fixed point to be:

$$\dot{\theta}_i = \frac{K}{N}\sum_{j=1}^{N}(\theta_j - \theta_i) \tag{2.1}$$

We then rearrange the terms in the sum and divide through by K to obtain:

$$\frac{\dot{\theta}_i}{K} = -\theta_i + \frac{1}{N}\sum_{j=1}^{N}\theta_j$$

Rewriting this in matrix form, we obtain

$$\frac{N\dot{\vec{\theta}}_i}{K} = \begin{bmatrix} 1-N & 1 & 1 & \dots & 1 \\ 1 & 1-N & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \dots & \dots & 1-N \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_N \end{bmatrix}$$

where every diagonal entry in the operator matrix is $1 - N$ and every other entry is 1. We can now rewrite this as a vector equation

$$\frac{N\dot{\vec{\theta}}}{K} = (P - NI)\vec{\theta}$$

where P is the circulant matrix whose entries are all 1, I is just the identity matrix, and $\vec{\theta}$ encodes the phase of every oscillator. We now assume that the solution is of the form

$$\vec{\theta} = \vec{V}e^{\lambda t}$$

which yields the equation

$$\frac{N}{K}\vec{V}\lambda e^{\lambda t} = (P - NI)\vec{V}e^{\lambda t}$$

which we simplify to obtain

$$\frac{\lambda N\vec{V}}{K} = P\vec{V} - N\vec{V}$$

which we then rearrange to obtain the eigenvalue equation

$$(\frac{\lambda N}{K} + N)\vec{V} = P\vec{V}$$

We then let $\Lambda = \frac{\lambda N}{K} + N$, then

$$\Lambda\vec{V} = P\vec{V}$$

We see that P, being a matrix whose entries are all equal to one, is rank one. Moreover, P has one non-zero eigenvalue $\Lambda_1 = N$, which corresponds to the eigenvector

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

whose entries are all equivalent. The other $N - 1$ eigenvalues of the matrix P, Λ_k where $k = 2...N$ are all zero, and correspond to the vectors $\vec{v}_i$ where

$$\vec{v}_i = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_N \end{bmatrix}$$

where $\sum\limits_{i=1}^{N} \theta_i = 0$. Note that this set of vectors indeed forms a basis for a vector space of $N - 1$ dimensions, as we are given $N - 1$ degrees of freedom in choosing θ_i. We now solve for the eigenvalues λ_k and see that $\lambda_1 = 0$ and $\lambda_k = -K$ when $k > 1$. λ_1 corresponds to a null vector related to the symmetry mode of the Kuramoto model, while the other λ_k correspond to the exponential decay observed in $N - 1$ dimensions. Thus in $N - 1$

dimensions of the eigenspace of the linearization of the Kuramoto model about its fixed point exponential decay behavior, indicating the stability of the fixed point in these dimensions. When we add white noise to the model as follows

$$\dot{\theta}_i = \omega + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \xi_i(t)$$

we expect that the distribution of phases after sufficient time has elapsed for synchronization to occur is that predicted by Ornstein-Uhlenbeck in the $N-1$ dimensional eigenspace corresponding to the non-zero eigenvalues of the linearization about the fixed, synchronized point of this system (discussed in the next section). What happens when the ω's of each oscillator are allowed to take on a distribution as opposed to being identical is discussed in Section 2.5.

2.4 Synchronization and Ornstein-Uhlenbeck

An Ornstein-Uhlenbeck process $\{x_t\}$ is defined by the stochastic differential equation

$$dx_t = \theta(\mu - x_t)dt + \sigma dW_t$$

where x_t is a random variable, W_t is the Wiener process, and $\theta \geq 0$ and $\sigma \geq 0$. The PDF of an Orstein-Uhlenbeck process $f(x,t)$ follows the Fokker-Planck equation, which is defined as

$$\frac{df}{dt} = \theta\frac{d}{dx}[(x-\mu)f] + \frac{\sigma^2}{2}\frac{d^2f}{dx^2}$$

Where in the Kuramoto model $f = \rho$, $\frac{\sigma^2}{2} = D$, $\mu = \bar{\rho}$, $x = \theta$ and $\theta = K$. A key feature of this type of process is its steady-state distribution in the limit as time goes to infinity:

$$f_s(\theta) = \sqrt{\frac{K}{2\pi D}}e^{\frac{-K(\theta-\bar{\rho})^2}{2D}}$$

The relevance of Ornstein-Uhlenbeck to the Kuramoto model with noise incorporated lies in the steady-state phase distribution of Kuramoto oscillators. Given zero initial conditions and identical oscillators, the cumulative distribution function of the phases of Kuramoto oscillators after synchronization with standard normal white noise is an error function, which corresponds to the probability density function given by $f_s(\theta)$.

We can work backwards from the Fokker-Planck equation to obtain an equation which describes the Ornstein-Uhlenbeck process. We first start with the original equation

$$\frac{d\rho}{dt} + \frac{d}{d\theta}(\rho v) = D\frac{d^2\rho}{d\theta^2}$$

We first note that the function $v(\theta, t; \omega)$ is equal to $\frac{d\theta}{dt}$, which, in the continuous setting, is given by:

$$v(\theta, t) = \frac{d\theta}{dt} = \bar{\omega} + K\int\limits_{\Omega} \rho(\theta')\sin(\theta' - \theta)d\theta'$$

where ρ is the phase density and Ω is the domain of θ'. When we assume that $\theta' \approx \theta$, we can linearize just as we did in Section 2.3 and obtain:

$$v(\theta, t) = \bar{\omega} + K\int\limits_{\Omega} \rho(\theta')(\theta' - \theta)d\theta'$$

which simplifies to

$$v(\theta, t) = \bar{\omega} + K(\bar{\rho} - \theta)$$

substituting this back into the Fokker-Planck equation we obtain

$$\frac{d\rho}{dt} + \frac{d}{d\theta}(\rho(\bar{\omega} + K(\bar{\rho} - \theta))) = D\frac{d^2\rho}{d\theta^2}$$

we then let $\bar{\omega} = 0$ without loss of generality:

$$\frac{d\rho}{dt} + K\frac{d}{d\theta}(\rho(\bar{\rho} - \theta)) = D\frac{d^2\rho}{d\theta^2}$$

which, when we consider ρ to be the probability density function of a random variable ψ, corresponds to the stochastic differential equation for an Ornstein-Uhlenbeck process:

$$d\psi_t = K[(\bar{\rho} - \psi_t)]dt + DdW_t$$

Thus when we linearize about the fixed point of the Kuramoto model, we see in fact that the phase distribution of the oscillators behaves as predicted by Ornstein-Uhlenbeck, and more specifically that the width of the distribution can be predicted by considering the parameters D and K.

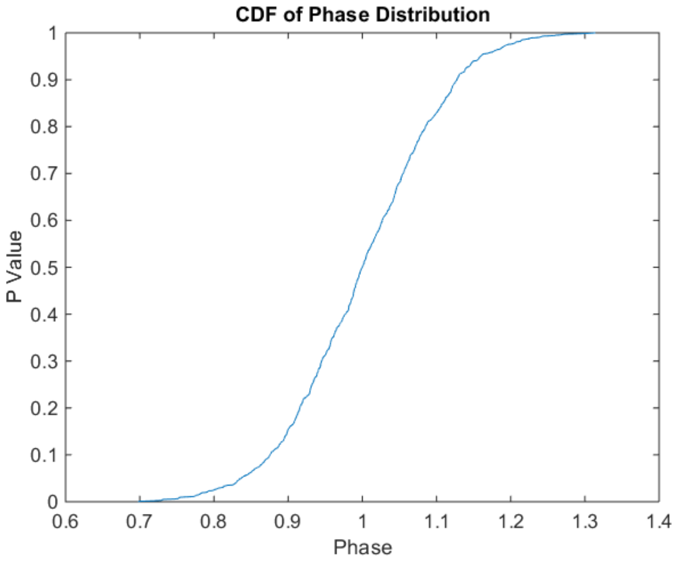


Figure 2.8 A plot of the CDF of the phases of the oscillators governed by equation 2.2.

2.5 A Hierarchy of Models

2.5.1 The Equations

To understand how the stochastic Kuramoto model defined by

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \xi_i(t)$$

has a phase distribution corresponding to that of an Ornstein-Uhlenbeck process, we first establish a hierarchy of models which can allow us to understand the parameter space of the Kuramoto model and its relationship with this steady state distribution. The three models, placed in order of increasing complexity, are

$$\dot{\theta}_i = -\theta_i + \xi_i(t) \tag{2.2}$$

which is a standard model whose steady state corresponds directly to Ornstein-Uhlenbeck,

$$\dot{\theta}_i = \bar{\omega} + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \xi_i(t) \tag{2.3}$$

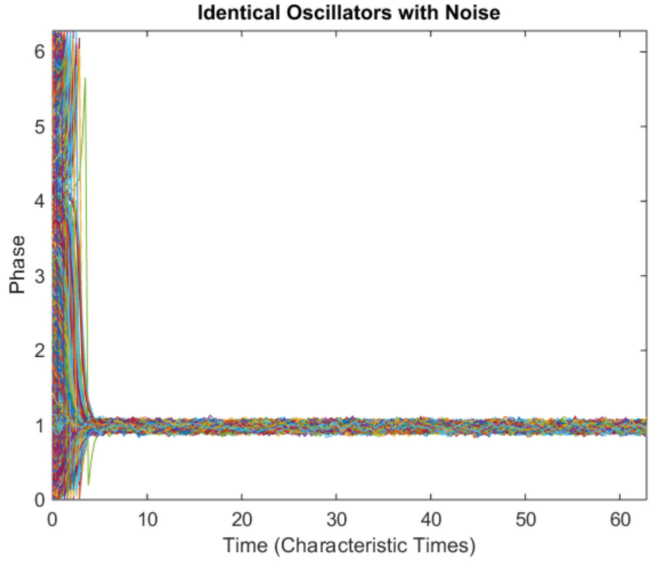


Figure 2.9 A plot of the phases of the oscillators governed by equation 2.3.

where every oscillator has an identical natural frequency, and

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \xi_i(t) \tag{2.4}$$

We note the similarity between the exponential decay model (2.2) and the linearization of Kuramoto about its synchronized fixed point, as near that point the Kuramoto model is approximately a set of N equations which correspond to exponential decay, as verified by the eigenvalues derived in Section 2.3. Thus model 2.3 should have a steady-state distribution whose CDF is a error function, which would match the result predicted by the theory of Ornstein-Uhlenbeck. In model 2.4, we observe unique behavior which is discussed at the end of this section.

2.5.2 Numerics

In the first model (equation 2.2), we observe a combination of exponential decay in θ_i for each oscillator, as well as the presence of white noise. A plot of phase v. time is shown in Figure 2.7. And the steady state distribution, an error function, is shown in Figure 2.8. The behavior of the phases in time is that of exponential decay to a steady state distribution, whose CDF

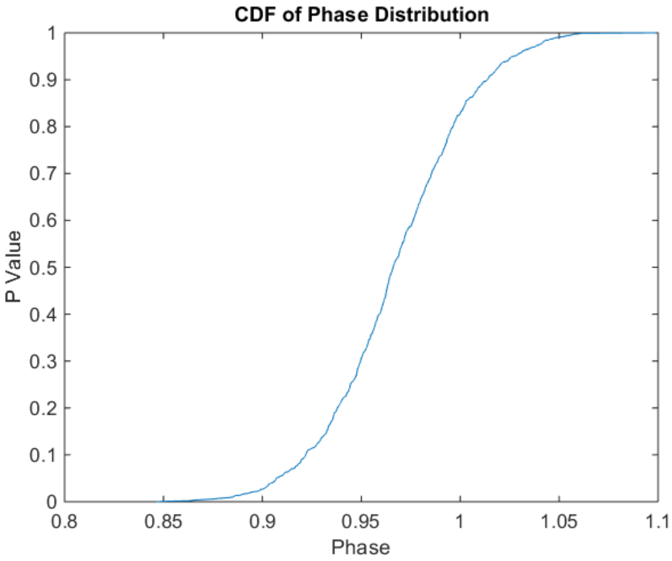


Figure 2.10 A plot of the phase distribution of the oscillators governed by equation 2.3.

is precisely that predicted by Ornstein-Uhlenbeck (the integral of a Gaussian).

We now examine the second model, in which each oscillators has an identical natural frequency (equation 2.3). As we showed in Section 2.3, in $N-1$ dimensions, this model ought to behave as predicted by Ornstein-Uhlenbeck. That is, its steady state phase distribution should have a CDF which is an error function. We first examine a plot of the phases of oscillators against time for this model, shown in Figure 2.9. We again observe synchronization, where the phases of the oscillators approach a distribution, whose CDF is shown in Figure 2.10. We again see that the steady state behavior matches that predicted by Ornstein-Uhlenbeck. Instead of having a system of coupled oscillators which are identical, we instead have one in which each oscillator has a unique natural frequency. In this case, we observe that although the oscillators do in fact synchronize, each oscillator behaves differently as an individual in this case.

Because the natural frequencies are not identical, the phases reach a state in which the distance between them remains constant. It is important to note that this implies that the phases do not all have to be the same. Because

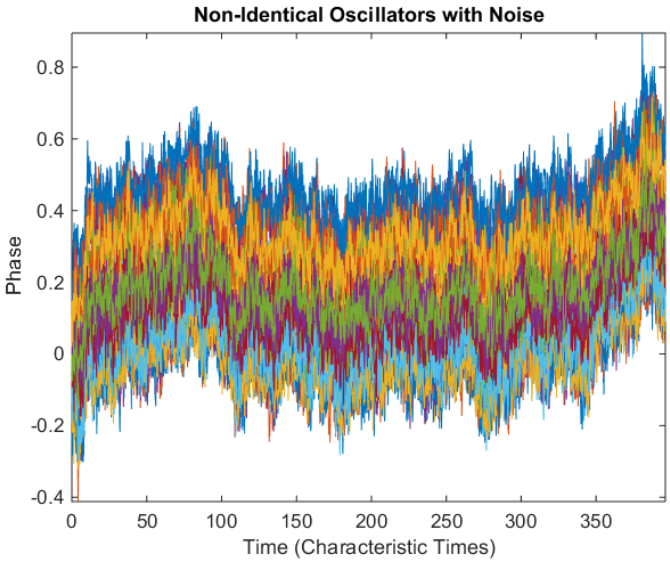


Figure 2.11 A plot of the phases of the oscillators governed by equation 2.4.

of this, we observe different qualitative behavior in plots of the phases of theses oscillators vs. time (Figure 2.11)

In Figure 2.11, we see colored stripes in the plot of phase vs. time for the oscillators in the simulation. This corresponds directly to the behavior just described; each oscillator moves randomly about a fixed value of θ, which leads to the formation of these bands.

It is also important to note that the width of the phase distribution does not change in time. We can see this visually by looking at a phase vs. time plot with the mean phase subtracted out at each time step (see Figure 2.12).

We have that the spread of the distribution is independent of time which reflects the stationarity of the distribution described in Section 2.4. Note that when we call the phase distribution of oscillators *stationary*, we mean that its mean and variance do not change over time. We also see that the variance of each of these distributions is constant in time.

In this section, we thoroughly examined the Kuramoto model and its tendency to exhibit synchronous behavior. We then introduced white noise to the model, taking note of the impact of this noise on the emergent be-

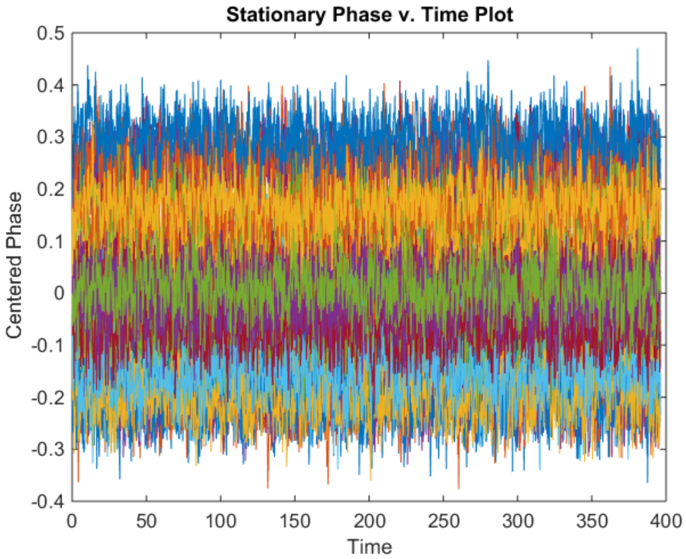


Figure 2.12 A stationary plot of the phases of the oscillators governed by equation 2.4.

havior of the system. We then linearized the model about its fixed point and showed how the stochastic Kuramoto model's steady-state phase distribution is that of an Ornstein-Uhlenbeck process. Finally, we addressed the most complex version of the stochastic Kuramoto model in this paper: the case in which the natural frequencies of the oscillators are taken from a distribution. In that case, we observed a distinct striped pattern in the phase v. time plot of the oscillators. In the next chapter we discuss how computational topological methods can shed light on the behavior of the Kuramoto model.

Chapter 3

Clusters and Topological Signatures

3.1 Betti Zero

Under the usual all-to-all coupling structure Kuramoto oscillators tend to synchronize into one cluster of oscillators. When we alter this and impose a subsystem structure on the model, the oscillators break into multiple clusters which behave as a standard Kuramoto system individually. That is, each cluster exhibits phase-locked behavior and moves with some frequency $\bar{\omega}_j$, the average natural frequency of every oscillator in the j^{th}cluster.

In order to conduct topological data analysis on phase data from the Kuramoto model, we must first convert the phase data into point-cloud data. We first take the input data, which is a time series of phase values for each oscillators, and then turn it into points in Euclidean space (which may also be a time series).

For example, If we are to simulate the behavior of 40 oscillators over ten characteristic times τ, where our time steps are a twentieth of a characteristic time, we would have a phase vector whose dimensions are 40×2001. At each time-step, we have a column vector of length 40, which we label $\vec{\theta}_n(t)$ where n is the time-step.

Once we have the set of vectors $\vec{\theta}_n(t)$, we must first determine how to convert this information into point cloud data. We will utilize the following coordinate system, which lies on the n-torus:

$$\begin{bmatrix} \cos(\theta(t-\tau)) \\ \sin(\theta(t-\tau)) \\ \vdots \\ \cos(\theta(t-n\tau)) \\ \sin(\theta(t-n\tau)) \end{bmatrix}$$

At first glance, the choice of sine and cosine as $observation\ functions$ seems odd. However, what using these two functions allows us to do is to map the data onto a familiarly space, namely the n-torus, which is the direct product of S^1 with itself n times.

We have incorporated time delays as a proxy for angular velocity, partly inspired by *Takens' Embedding Theorem*, which states that for a dynamical system with strange attractor A of box-counting dimension d_A, the attractor A can be embedded in Euclidean space with dimension k where $k > 2d_A$. (**?**) In most cases, we will use the sine and cosine of phase differences, though we will incorporate more in cases in which noise or small variances in natural frequency complicate the process of splitting data points into clusters.

We also note that an *attractor* of a dynamical system is a set of values to which the systems tends, and that the *box-counting dimension* is just a way of determining the fractal dimension of a set in Euclidean space. Takens's theorem therefore gives us a sufficient condition for embedding the attractor of the Kuramoto model, which is necessarily a fixed point or a limit cycle. (Giacomin et al., 2012) Note that the case in which the oscillators reach a fixed point corresponds to the case in which the mean natural frequency is zero, whereas the limit cycle case emerges when $\bar{\omega} = 0$.

Once we have time series point-cloud data selected from the options provided previously in this section, we are left with the decision to examine the data at individual time-steps, which lends itself to realizations such as topological barcodes, or to look at every timestep and use a CROCKER plot to realize the Betti numbers for the data over many time-steps. (Topaz et al., 2015) This is done using implementations from the Javaplex library. (Adams and Tausz, 2014) For the intents and purposes of this paper, we will rely on topological barcodes.

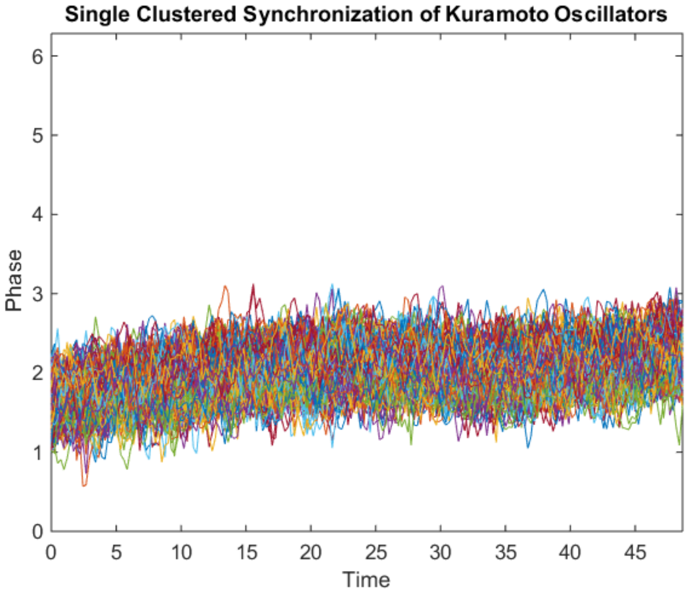


Figure 3.1 A plot of phases v. time for Kuramoto oscillators with one cluster built into the network structure (also known as all-to-all coupling).

3.2 From One Cluster to Many

This section will first be a brief review of the previous chapter, in the context of clustering. The simplest case is that in which Kuramoto oscillators form one cluster. An example of system in which there is only one cluster is given in Figure 3.1. When we discuss clustering in the context of 1 cluster, we are simply describing the synchronization of oscillators governed by the standard Kuramoto model (Kuramoto, 1975). In order to ensure that the application of persistent homology indeed yields a B_0 value of 1, we consider a simplified representation of Kuramoto oscillators subject noise and split into two clusters, evolving over time, where $f_1(t)$ describes the evolution of cluster 1, and $f_2(t)$ describes that of cluster 2. Why we chose to consider the two cluster case will soon become evident. First, the phase evolution of the simulation which we will estimate using linear functions of the form $at + b$ is shown in Figure 3.2.

$$f_1(t) = \alpha t + \alpha_0 \tag{3.1}$$

$$f_2(t) = \beta t + \beta_0 \tag{3.2}$$

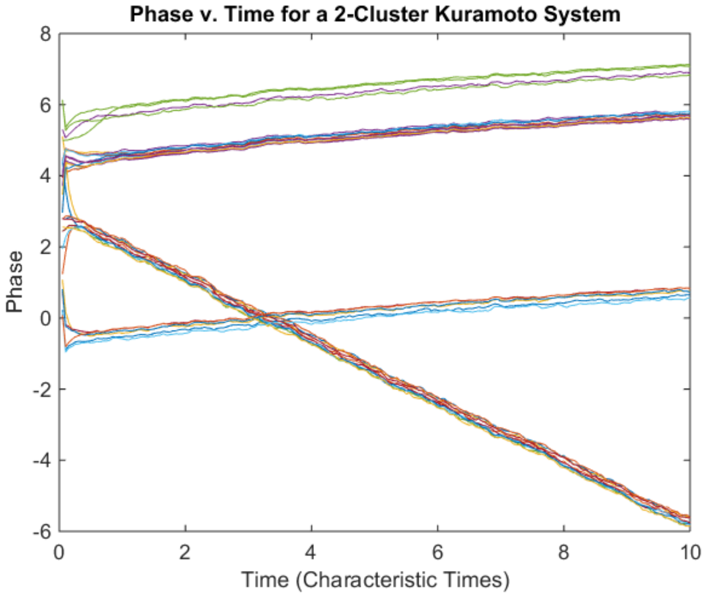


Figure 3.2 A plot of phases v. time for Kuramoto oscillators with two clusters built into the network structure.

We now assume that we have generated a numerical simulation described by these two equations. Taking the simulation data, we convert it into our point-cloud format, which requires us to sample f_1 and f_2 at two times, t_1 and t_2. Letting t_2 - t_1 = τ, we calculate $f_1(t-\tau)$ and $f_2(t-\tau)$, and then vectorize these values as follows:

$$\begin{bmatrix}\cos(f_1(t-\tau))\\ \sin(f_1(t-\tau))\end{bmatrix} \tag{3.3}$$

$$\begin{bmatrix}\cos(f_2(t-\tau))\\ \sin(f_2(t-\tau))\end{bmatrix} \tag{3.4}$$

We are now interested in determining the Euclidean distance between these two points, which is a rough estimation of the distance between two points, each members of a different cluster. Using the standard euclidean 2-norm, we find the distance between these two points to be;

$$\sqrt{(\cos(f_1(t-\tau)) - \cos(f_2(t-\tau)))^2 + (\sin(f_1(t-\tau)) - \sin(f_2(t-\tau)))^2} \tag{3.5}$$

which simplifies to

$$g(\tau) = \sqrt{(\cos(\alpha\tau) - \cos(\beta\tau))^2 + (\sin(\alpha\tau) - \sin(\beta\tau))^2} \tag{3.6}$$

where we have used the fact that f_1 and f_2 have constant derivatives. We simplify the function $g(\tau)$ one more time:

$$g(\tau) = \sqrt{2 - 2\cos[(\alpha - \beta)\tau]} \tag{3.7}$$

We now seek to maximize the value of this function, as the maximum of $g(\tau)$ occurs at the τ value which maximizes the distance between two clusters in our euclidean data space. This in turn gives us the best chance of detect the correct number of clusters. We find that the maxima of this function occur at all τ where

$$\tau = \frac{(2n+1)\pi}{\alpha - \beta} \qquad n \in \mathbb{Z} \tag{3.8}$$

In fact, it can be shown that this τ value is a maximizer of the distance between two clusters when we use any number of time lags to construct data of the form:

$$\begin{bmatrix} \cos(\theta(t-\tau)) \\ \sin(\theta(t-\tau)) \\ \vdots \\ \cos(\theta(t-n\tau)) \\ \sin(\theta(t-n\tau)) \end{bmatrix}$$

We also note that at $\alpha = \beta$, this quantity is undefined. This case corresponds to the one cluster case, in which all oscillators have synchronized and move together at some angular velocity $\bar{\omega}$ as described by the Kuramoto model. This is to say, if all oscillators are truly in one cluster, and for sufficiently large ϵ, that no value of τ will allow us to detect two clusters of data points. For the two cluster case, we simply choose τ according to this formula.

It is usually the case that we wish to choose $\tau > 0$, as well as the smallest τ, that is when $n = 0$. This can also be explained qualitatively: for τ

values too small, the noise in the model makes detecting clusters substantially more difficult. On the other hand, when τ is too large, we observe phase-winding, which is not necessarily a problem given the periodicity of our cloud-point data, but is nevertheless one more thing to worry about in our analysis.

Typically we will use three time delays to embed the time-series data for each oscillator in 6-space, that is, we will be using data of the form:

$$\begin{bmatrix} \cos(\theta(t-\tau)) \\ \sin(\theta(t-\tau)) \\ \cos(\theta(t-2\tau)) \\ \sin(\theta(t-2\tau)) \\ \cos(\theta(t-3\tau)) \\ \sin(\theta(t-3\tau)) \end{bmatrix}$$

It is important to note a few properties of data of this form. The most relevant is the manifold on which the data lives. If we evaluate the norm of a point $\overrightarrow{x}$ in a set of data of this form, we find that

$$||\overrightarrow{x}|| = \sqrt{3}$$

We also know that any two points which lie in this data set can be no more than $2\sqrt{3}$ apart, as that maximum is achieved, for example, when the two points are given by

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Thus we have no need to consider a filtration value (proximity parameter value) any larger than $2\sqrt{3}$ when using this specific form of data.

Things get more complicated when we introduce network structures which force the oscillators into more than two clusters. (see Figure 3.3) It becomes much more difficult to determine an optimal τ value analytically, and we therefore rely on computational techniques, which are outlined in the following section.

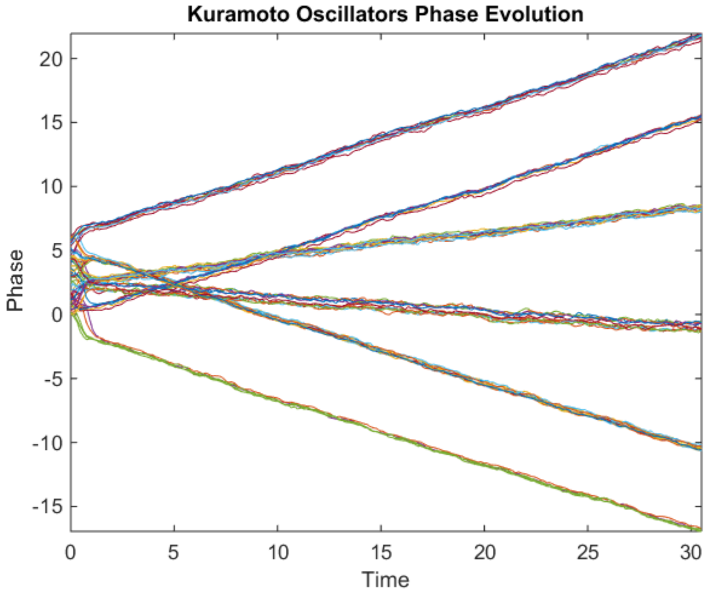


Figure 3.3 A plot of phases v. time for Kuramoto oscillators with four clusters built into the network structure.

3.3 Cluster Identification

Because we are interested in being able to detect any number of clusters, we are forced to abandon analytic techniques like that outlined in the previous section and optimize our time delay measurement τ using computation. In order to do this, we outline an algorithm to help optimize time delay values, from which we proceed to use a union-find algorithm to identify the members of each cluster (Agarwal et al., 2006):

Step 1: Estimate the noise amplitude in simulation data, denoted D
Step 2: Choose a proximity parameter value ϵ equal to $4D$ to ensure that oscillators which are synchronized are decidedly connected in our euclidean data space
Step 3: Fix ϵ and vary τ
Step 4: Determine which values of τ return the correct B_0 value
Step 5: Use union-find to identify cluster members for some subset of the ideal values of τ
Step 6: Determine which oscillators remain in the same cluster across multiple values of τ

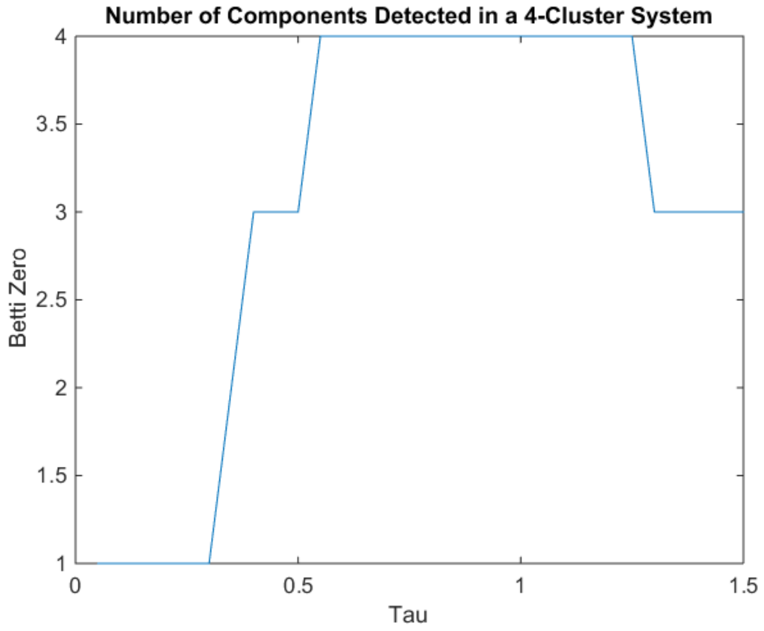
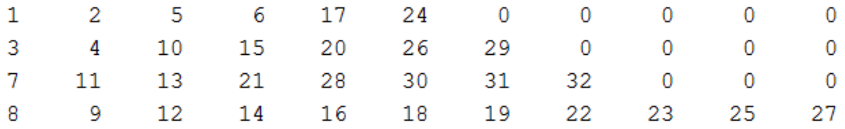


Figure 3.4 A plot of B_0 versus the value of τ chosen.



```
1     2     5     6    17    24     0     0     0     0     0
3     4    10    15    20    26    29     0     0     0     0
7    11    13    21    28    30    31    32     0     0     0
8     9    12    14    16    18    19    22    23    25    27
```

Figure 3.5 Output Data Denoting Clusters Generated for a fixed value of Tau

Step 7: Use this information to determine the number of clusters and their members

We now describe these individual steps.

1. For the purpose of proof of concept, we use the fact that D is an input parameter to carry out this step
2. This step follows quickly from the previous one.
3. For this step, we loop over all values of τ between 0 and $1.5T$, or 1.5 characteristic times. Doing so generates data which we can plot as we do in Figure 3.4.

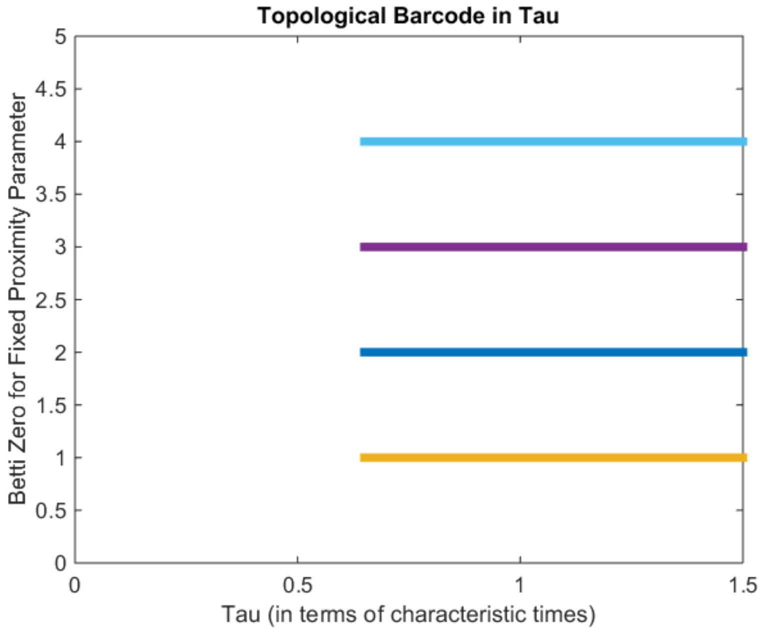


Figure 3.6 Topological Barcode Plotted in terms of Tau

4. We now choose a selection of τ values τ_0, τ_1, τ_2, ..., τ_N which return the proper number of components. In the purely empirical case in which we are unaware of the number of clusters already, we would defer to choosing the maximum value attained by a function such as that displayed in Figure 3.4.

5. Running a union-find algorithm on the simplicial complexes generated by using τ_0, τ_1, τ_2, ..., τ_N, we are given a list of clusters and their respective members as output. An example of this output is shown in Figure 3.5.

6. We now see which oscillators are members of the same cluster across τ_0, τ_1, τ_2, ..., τ_N, which gives us the number of clusters and their members. A barcode depicting this data is shown in Figure 3.6.

3.4 Efficacy of Our Cluster Identification Method

In order to determine how well this approach works, we test the effectiveness of the algorithm's dependence on the noise amplitude. Running 100

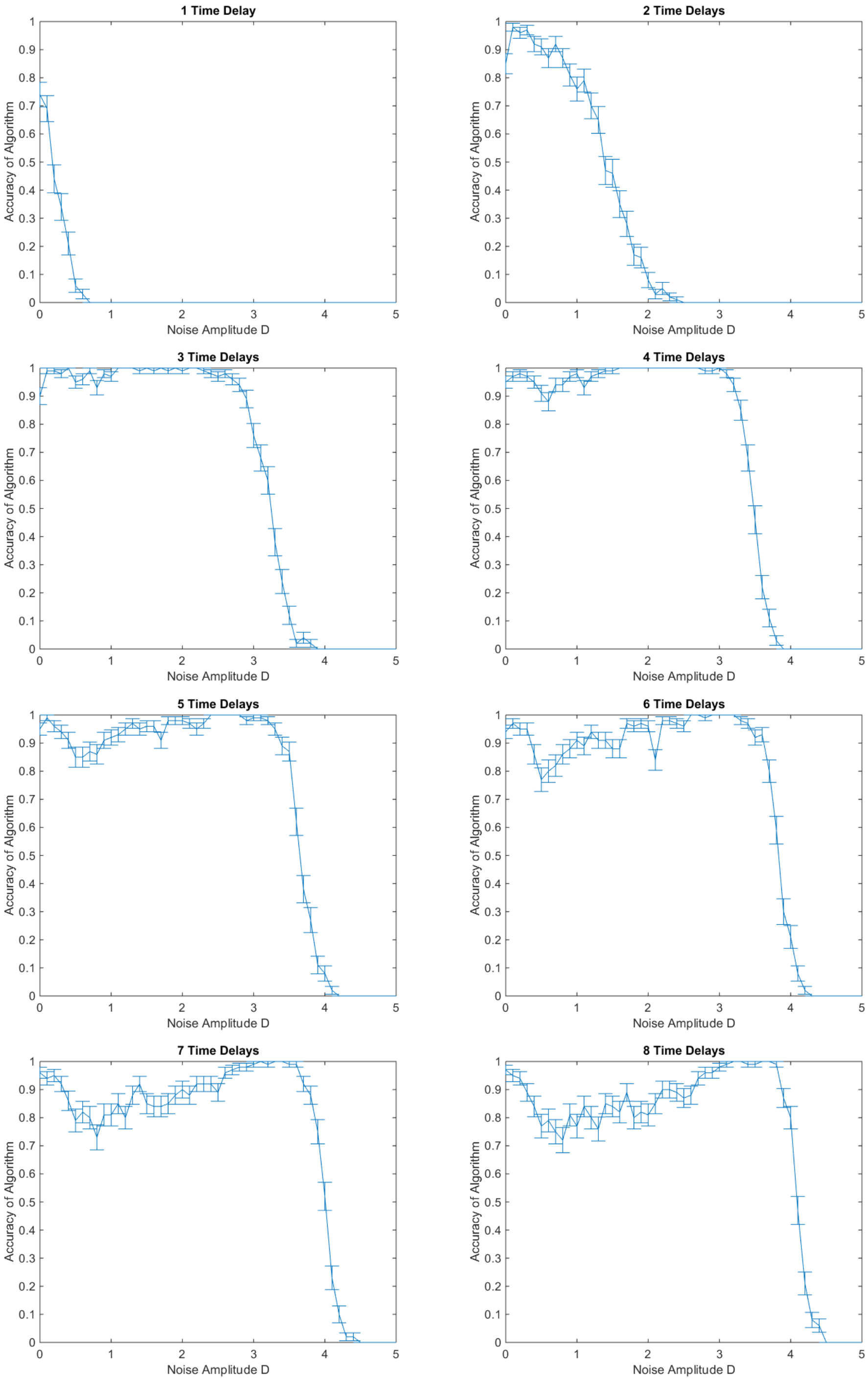


Figure 3.7 Dependence of Cluster Detection Accuracy on the Noise Amplitude

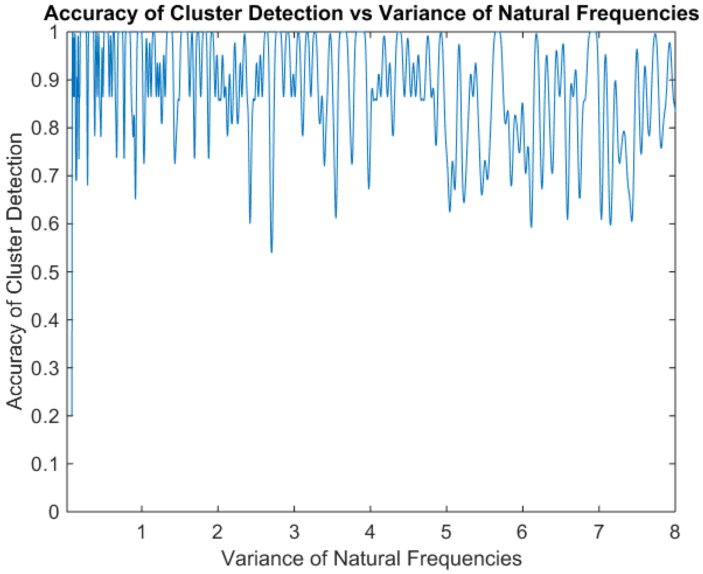


Figure 3.8 Dependence of Cluster Detection Accuracy on the Variance of ω and the Number of Oscillators

simulations for 100 values of D, the noise amplitude, between $D = 0.01$ and $D = 10$, we obtain the results shown in Figure 3.7. We observe that using a higher embedding dimension is generally better for higher levels of noise. However, at low noise levels, a higher embedding dimension is not necessarily better. This may be a result of transience in the solution, but is left to be explained.

We observe that there is little effect on the efficacy of our cluster-detection method when the variance of the distribution of the natural frequencies of the oscillators is changed. This is shown graphically in Figure 3.8. We also note that this algorithm's efficacy is not dependent on the size of the system, which is also demonstrated in Figure 3.8.

3.5 Cluster Detection under Nonlocal Coupling

We can alter the Kuramoto model in ways other than by changing its network structure. Recall the stochastic Kuramoto model

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \xi_i(t)$$

in which the coupling strength K is constant. By introducing a spatial variable x, which we assign to each oscillator, we can rewrite the system as

$$\dot{\theta}_i = \omega_i + \frac{1}{N}\sum_{j=1}^{N}K(x_i, x_j)\sin(\theta_j - \theta_i) + \xi_i(t)$$

We now must choose a function $K(x_i, x_j)$ which makes physical sense. To impose the principles of nonlocal coupling, we want a function which assigns high coupling values to oscillators which are close together in physical space and low values to those which are far apart. We also want the function to be even and positive everywhere. Taking inspiration from Kuramoto and Battogtokh, as well as others who have written on nonlocally coupled oscillators, we choose the function

$$K(x_i, x_j) = \frac{\kappa}{2}e^{-|x_i - x_j|}$$

(Kuramoto and Battogtokh, 2002) and end up with the system

$$\dot{\theta}_i = \omega_i + \frac{\kappa}{2N}\sum_{j=1}^{N}e^{-|x_i - x_j|}\sin(\theta_j - \theta_i) + \xi_i(t)$$

for which we have example simulations shown in Figures 3.10 and 3.9. Using the algorithm outlined in Section 3.3, we can detect the number of clusters into which the oscillators break up. This case is distinct from those explored in 3.4, as we retain the all-to-all coupling structure, so the clustering in this case comes from spatial dependence, not connectivity of the system itself. In Figure 3.9 we observe a strongly decoupled system of oscillators, with the topological barcode corresponding to the simulation data from it is shown in Figure 3.9.

We see that in the uncoupled case the algorithm detects the proper number of clusters: almost every oscillator is moving at a unique frequency,

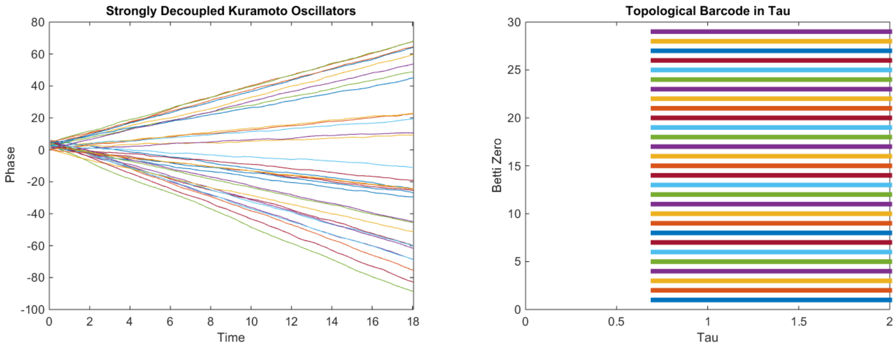


Figure 3.9 Kuramoto Oscillators with Weak Spatial Coupling

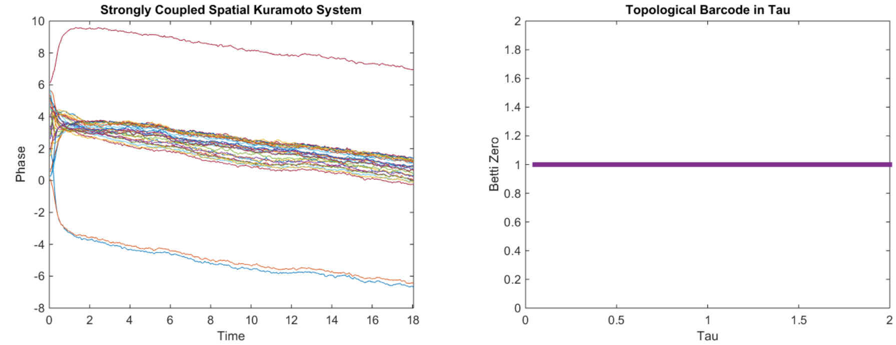


Figure 3.10 Kuramoto Oscillators with Strong Spatial Coupling

which, by the nature of the point cloud in which we embed our phase data, suggests that we ought to expend at B_0 value which is approximately the number of oscillators, in this case 32.

When we increase the value of κ enough, we observe synchronization under the nonlocal coupling scheme, demonstrated in Figure 3.10. In this case, the algorithm correctly identifies that the system is completely synchronized, for a wide range of values of τ (shown in Figure 3.10.) An interesting and expected finding during this test of our algorithm was that it was better to choose lower embedding dimensions (roughly four or six), which was quite efficient computationally, as opposed to having to use higher dimensional data. This is most likely a consequence of the fact that we chose to use low noise amplitude values in these simulations (0.1). At much higher values of D, one should embed in higher dimensions.

Chapter 4

Conclusions and Future Directions

In this thesis we began with an introduction to the Kuramoto model for coupled oscillators. We then moved on to the concepts behind topological data analysis, and how ideas like persistent and simplicial homology can be applied to simulation data generated from the Kuramoto mode. We then introduced noise to the Kuramoto model, and examined how this changed the long-term behavior of the system. By linearizing about the system's stable fixed point, we were able to derive a result which matches the theory of Ornstein and Uhlenbeck: the oscillators tend to become distributed normally about an average phase value as $t \to \infty$. We also identified the $N-1$ dimensional stable eigenspace which corresponds to synchrony and the 1 dimensional eigenspace which corresponds to the symmetry mode of the Kuramoto model.

After developing the proper theory to understand the connection between the coupling constant K, the noise amplitude D, and the width of the steady-state distribution of the stochastic Kuramoto model, we utilized the tools outlined in section 1.2 to study simulation data generated from the model. We embedded the time series data for each individual oscillator on the n-torus by using indicator functions inspired by the work of Floris Takens. From this, we fixed a value for the proximity parameter ϵ and generated topological barcodes in τ, the time delay. By identifying optimal values of both ϵ and τ using both analytical and empirical methods, we were able to develop a clustering-detection algorithm which is over 97 percent accurate for different ranges of noise, depending on the number of time lags used to construct the point cloud (see Figure 3.7).

Possible extensions of this research include the examination of other models for coupled oscillators, namely Van der Pol oscillators and pulse-coupled oscillators. With these other models come different questions, such as, what is synchronization in the context of pulse-coupling? Or, how can we use the ideas of persistent homology, reconstruction of attractors through time delays per Takens's theorem, and embedding of phase data in high-dimensional euclidean data to detect synchrony, cluster formation, and other behavior typically seen in systems of coupled oscillators? There is also the question of the consequences of imposing different network structures on systems of coupled oscillators. There has been some work done in this area, but little to none involving methods from topological data analysis.

Bibliography

Acebrón, Juan A, Luis L Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. 2005. The Kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics* 77(1):1 – 117.

Adams, Henry, and Andrew Tausz. 2014. Javaplex tutorial. *https://codegooglecom/p/javaplex/wiki/Tutorial* URL https://code.google.com/p/javaplex/wiki/Tutorial.

Agarwal, Pankaj K., Lars Arge, and Ke Yi. 2006. I/o-efficient batched union-find and its applications to terrain analysis. In *Proceedings of the Twenty-second Annual Symposium on Computational Geometry*, 167–176. SCG '06, New York, NY, USA: ACM. doi:10.1145/1137856.1137884. URL http://doi.acm.org/10.1145/1137856.1137884.

Arenas, Alex, Albert Díaz-Guilera, and Conrad J Pérez-Vicente. 2006. Synchronization reveals topological scales in complex networks. *Phys Rev Lett* 96(11):114,102.

Bag, Bidhan Chandra, KG Petrosyan, and Chin-Kun Hu. 2007. Influence of noise on the synchronization of the stochastic Kuramoto model. *Physical Review E* 76(5):056,210.

Balmforth, Neil J, and Roberto Sassi. 2000. A shocking display of synchrony. *Physica D: Nonlinear Phenomena* 143(1):21–55.

Boissonnat, Jean-Daniel, and Monique Teillaud. 2007. *Effective computational geometry for curves and surfaces*. Springer.

Breakspear, Michael, Stewart Heitmann, and Andreas Daffertshofer. 2010. Generative models of cortical oscillations: neurobiological implications of the Kuramoto model. *Frontiers in human neuroscience* 4:1–14.

Carlsson, Gunnar, and Vin De Silva. 2010. Zigzag persistence. *Foundations of computational mathematics* 10(4):367–405.

Chopra, Nikhil, and Mark W Spong. 2005. On synchronization of Kuramoto oscillators. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, 3916–3922. IEEE.

Collins, James J, and Ian Stewart. 1994. A group-theoretic approach to rings of coupled biological oscillators. *Biological cybernetics* 71(2):95–103.

Cumin, D, and CP Unsworth. 2007. Generalising the Kuramoto model for the study of neuronal synchronisation in the brain. *Phys D: Nonlin Phenom* 226(2):181–196.

De Silva, Vin, and Gunnar Carlsson. 2004. Topological estimation using witness complexes. *Proc Sympos Point-Based Graphics* 157–166.

Dörfler, Florian, and Francesco Bullo. 2011. On the critical coupling for Kuramoto oscillators. *SIAM Journal on Applied Dynamical Systems* 10(3):1070–1099.

Easley, David, and Jon Kleinberg. 2010. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press.

Edelsbrunner, Herbert, David Letscher, and Afra Zomorodian. 2002. Topological persistence and simplification. *Discrete and Computational Geometry* 28(4):511–533.

Filatrella, G, Niels Falsig Pedersen, and K Wiesenfeld. 2007. Generalized coupling in the Kuramoto model. *Phys Rev E* 75(1):017,201.

Garland, Joshua, Elizabeth Bradley, and James D Meiss. 2015. Exploring the topology of dynamical reconstructions. *arXiv preprint* .

Giacomin, Giambattista, Khashayar Pakdaman, and Xavier Pellegrin. 2012. Global attractor and asymptotic dynamics in the kuramoto model for coupled noisy phase oscillators. *Nonlinearity* 25(5):1247.

Honeycutt, Rebecca L. 1992. Stochastic runge-kutta algorithms i white noise. *Physical Review A* 45(2):600.

Hsiang, Solomon M. 2015. Matlab code to generate and analyze random networks. *http://wwwsolomonhsiangcom/computing/matlab-code* URL http://www.solomonhsiang.com/computing/matlab-code.

Kuramoto, Yoshiki. 1975. Self-entrainment of a population of coupled non-linear oscillators. *International symposium on mathematical problems in theoretical physics* 420–422.

Kuramoto, Yoshiki, and Dorjsuren Battogtokh. 2002. Coexistence of coherence and incoherence in nonlocally coupled phase oscillators. *Nonlinear Phemon Complex Sys 5* 380 – 385.

Mirollo, Renato E, and Steven H Strogatz. 1990. Synchronization of pulse-coupled biological oscillators. *SIAM J Appl Math* 50(6):1645–1662.

Niebur, Ernst, Heinz G Schuster, Daniel M Kammen, and Christof Koch. 1991. Oscillator-phase coupling for different two-dimensional network connectivities. *Physical Review A* 44(10):6895–6940.

O'Keeffe, Kevin P, Pavel L Krapivsky, and Steven H Strogatz. 2015. Synchronization as aggregation: Cluster kinetics of pulse-coupled oscillators. *arXiv preprint arXiv:150104115* .

Peskin, CS. 1975. Self-synchronization of the cardiac pacemaker. *Mathematical Aspects of Heart Physiology, New York University: New York* 268–278.

Schmidt, Lennart, Konrad Schönleber, Katharina Krischer, and Vladimir García-Morales. 2014. Coexistence of synchrony and incoherence in oscillatory media under nonlinear global coupling. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24(1):1–7.

Stolz, Bernadette. 2014. Computational topology in neuroscience (master's thesis). Oxford University .

Strogatz, Steven H. 2000. From Kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Phys D: Nonlin Phenom* 143(1):1–20.

Strogatz, Steven H, and Renato E Mirollo. 1988. Collective synchronisation in lattices of nonlinear oscillators with randomness. *Journal of Physics A: Mathematical and General* 21(13):L699–L705.

Takens, Floris. 1981. *Dynamical Systems and Turbulence, Warwick 1980: Proceedings of a Symposium Held at the University of Warwick 1979/80*, chap. Detecting strange attractors in turbulence, 366–381. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/BFb0091924. URL http://dx.doi.org/10.1007/BFb0091924.

Topaz, Chad M, Lori Ziegelmeier, and Tom Halverson. 2015. Topological data analysis of biological aggregation models. *PLoS ONE* 10(5):e0126,383.

Yeung, MK Stephen, and Steven H Strogatz. 1999. Time delay in the Kuramoto model of coupled oscillators. *Physical Review Letters* 82(3):648–651.