

# Causal Inference with Complex Dynamical Systems

George Stepaniants

February 1, 2019

## Abstract

We review in this paper the definitions of Granger Causality and Convergent Cross Mappings (CCM). We show that their application to linear and nonlinear physical oscillator systems fail to reproduce accurate network reconstructions. We find a regime that involves perturbing and heavily damping these systems under which both methods give back the correct network topologies. A perturbation based network inference approach is presented to solve this problem. It is evaluated on simulated data from systems of linear and nonlinear connected oscillators. Reasoning behind this method is discussed.

## 1 Introduction and Overview

## 2 Causal Inference Methods

### 2.1 Granger Causality

### 2.2 Convergent Cross Mapping

### 2.3 Perturbation Inference

## 3 Oscillator Systems

### 3.1 Harmonic Oscillators

The standard equation for a damped harmonic oscillator is

$$\frac{d^2x}{dt^2} =$$

The first system we investigate is that of  $n$  coupled harmonic oscillators where  $\mathbf{x}$  is the vector of displacements of all oscillators. This network of oscillators has fixed boundary conditions so oscillator 1 and oscillator  $n$  are attached to walls by springs. Also, pairs of oscillators in the network are connected to each other by springs with varying spring constants. Strength of coupling between oscillators and walls in this network is denoted by the  $n + 1 \times n + 1$  matrix of spring constants

$a$

Note that in a physical system of harmonic oscillators, the coupled interactions between pairs of oscillators are symmetric so the spring constant  $A_{ij}$  from oscillator  $i$  to oscillator  $j$  is equal to the spring constant  $A_{ji}$  from oscillator  $j$  to oscillator  $i$ . Therefore, in this system the matrix  $A$  is symmetric. The vector  $\mathbf{m}$  contains the masses of all harmonic oscillators in this system and  $c$  is the scalar damping coefficient.

A picture of a possible network configuration is shown in Figure [?]. The system of ODEs that governs the movement of these oscillators is

$$\mathbf{m} \cdot \frac{d^2\mathbf{x}}{dt^2} = A\mathbf{x} - c \frac{d\mathbf{x}}{dt}$$









<div> <div>Pred</div> <div>True</div> </div>				
	10	1	2	87
	0	0	1	99
	0	1	0	99
	1	0	0	99

Figure 1: Confusion matrix for GC on networks of two oscillators.

### 3.2 Directed Harmonic Oscillators

### 3.3 Kuramoto Oscillators

## 4 Numerical Experiments

Used MVGC toolbox for GC. Used R toolbox for CCM.

## 5 Computational Results

### 5.1 Inference for Two Oscillator System

Figure 1 shows the confusion matrix for GC inference over all possible directed networks of two harmonic or Kuramoto oscillators. As we can see, GC systematically predicts that all two oscillator networks are fully connected.

Figure 2 shows the confusion matrix for CCM inference over all possible directed networks of two harmonic or Kuramoto oscillators. As we can see, CCM systematically predicts that all two oscillator networks are fully connected.

Figure 3 shows the confusion matrix for perturbation inference over all possible directed networks of two harmonic or Kuramoto oscillators with one node perturbed.

Figure 4 shows the confusion matrix for perturbation inference over all possible directed networks of two harmonic or Kuramoto oscillators with both nodes perturbed.

### 5.2 Inference for N Oscillators

Give statistics like accuracy, TPR, FPR over different simulations of N oscillators. Don't damp or do anything special with the simulation parameters.









<div> <div>Pred</div> <div>True</div> </div>				
	100	0	0	0
	100	0	0	0
	0	0	100	0
	0	0	100	0

Figure 2: Confusion matrix for PC on networks of two oscillators with all nodes observed and only blue node perturbed.









<div> <div>Pred</div> <div>True</div> </div>				
	100	0	0	0
	0	100	0	0
	0	0	100	0
	0	0	0	100

Figure 3: Confusion matrix for PC on networks of two oscillators with all nodes observed and perturbed.

### **5.3 Inference with Damped Oscillator System**

Give statistics like accuracy, TPR, FPR over different simulations of  $N$  oscillators.

## **6 Analysis**

### **6.1 Reasons why Granger Causality Fails for Harmonic Oscillator Systems**

Summary of GC Experiments for Harmonic Oscillators														
Experiment Name	n	A	K	m	$p_0$	$v_0$	$\gamma$	Forcing	N	$\Delta t$	T	s	Prep	
gcH1	2	All four two-node networks	0.1, 0.2, ..., 10	1	$[-0.5, 0.5]$	0	0	None	100	0.1	25	0.1	Detrend	
gcH5 (deviate with small mass)	5	ER $p = 0.5$	1	$a$	$b$	0	0	None	100	0.1	25	0.1	Detrend	
gcH6 (deviate with large damping, show diff K)	5	ER $p = 0.5$	0.1, 0.2, ..., 10	1	$b$	0	10	None	100	0.1	25	0.1	Detrend	

Summary of GC Experiments for Kuramoto Oscillators														
Experiment Name	n	A	K	$\theta_0$	$\omega$	$\gamma$	Forcing	N	$\Delta t$	T	s	Prep		
gcK1	2	All four two-node networks	0.1, 0.2, ..., 10	$[0, 2\pi]$	$[-1, 1]$	0	None	100	0.1	25	0.1	Detrend		
gcH6 (deviate with large damping, show diff K)	5	ER $p = 0.5$	0.1, 0.2, ..., 10	$[0, 2\pi]$	$[-1, 1]$	10	None	100	0.1	25	0.1	Detrend		

Summary of CCM Experiments for Harmonic Oscillators														
Experiment Name	n	A	K	m	$p_0$	$v_0$	$\gamma$	Forcing	N	$\Delta t$	T	s	Prep	
ccmH1	2	All four two-node networks	0.1, 0.2, ..., 10	1	$[-0.5, 0.5]$	0	0	None	100	0.1	25	0.1	Detrend	
ccmH5 (deviate with small mass)	5	ER $p = 0.5$	1	$a$	$b$	0	0	None	100	0.1	25	0.1	Detrend	
ccmH6 (deviate with large damping, show diff K)	5	ER $p = 0.5$	0.1, 0.2, ..., 10	1	$b$	0	10	None	100	0.1	25	0.1	Detrend	

Summary of CCM Experiments for Kuramoto Oscillators														
Experiment Name	n	A	K	$\theta_0$	$\omega$	$\gamma$	Forcing	N	$\Delta t$	T	s	Prep		
ccmK1	2	All four two-node networks	0.1, 0.2, ..., 10	$[0, 2\pi]$	$[-1, 1]$	0	None	100	0.1	25	0.1	Detrend		
ccmH6 (deviate with large damping, show diff K)	5	ER $p = 0.5$	0.1, 0.2, ..., 10	$[0, 2\pi]$	$[-1, 1]$	10	None	100	0.1	25	0.1	Detrend		

Summary of PC Experiments for Harmonic Oscillators														
Experiment Name	n	$n_{\text{pert}}$	$n_{\text{obs}}$	A	K	m	$p_0$	$v_0$	$\gamma$	Forcing	N	$\Delta t$	T	s
pcH1	2	1, 2	1, 2	All four two-node networks	0.1, 0.2, ..., 10	1	$[-0.5, 0.5]$	$[-1, 1]$	0.1	None	100	0.1	25	0.1
pcH1	2	1, 2	1, 2	All four two-node networks	0.1, 0.2, ..., 10	1	$[-0.5, 0.5]$	$[-1, 1]$	0.1	None	100	0.1	25	0.1
pcH10	10	1, 2, ..., 10	10	ER $p = 0.1, 0.2, \dots, 1$	1	1	$[-0.5, 0.5]$	$[-1, 1]$	0.3	Pulse <sup>b</sup>	10	0.1	<sup>c</sup>	0.1
pcH20	20	1, 2, ..., 20	1, 2, ..., 20	ER $p = 0.5$	0.1	1	$[-0.5, 0.5]$	$[-1, 1]$	0.3	Pulse <sup>b</sup>	10	0.1	<sup>c</sup>	0.1
pcH21	20	1, 2, ..., 20	1, 2, ..., 20	BA $m = 10$	0.1	1	$[-0.5, 0.5]$	$[-1, 1]$	0.3	Pulse <sup>b</sup>	10	0.1	<sup>c</sup>	0.1

Summary of PC Experiments for Kuramoto Oscillators														
Experiment Name	n	$n_{\text{pert}}$	$n_{\text{obs}}$	A	K	$\theta_0$	$\omega$	$\gamma$	Forcing	N	$\Delta t$	T	s	Method
pcK1	2	1, 2	1, 2	All four two-node networks	0.1, 0.2, ..., 10	$[0, 2\pi]$	$[-1, 1]$	0	None	100	0.1	25	0.1	meanvar
pcK1	2	1, 2	1, 2	All four two-node networks	0.1, 0.2, ..., 10	$[0, 2\pi]$	$[-1, 1]$	0	None	100	0.1	25	0.1	chngrpt
pcK20	20	1, 2, ..., 20	1, 2, ..., 20	ER $p = 0.5$	0.1	$[0, 2\pi]$	$[-1, 1]$	0	Pulse <sup>b</sup>	10	0.1	<sup>c</sup>	0.1	meanvar
pcK21	20	1, 2, ..., 20	1, 2, ..., 20	BA $m = 10$	0.1	$[0, 2\pi]$	$[-1, 1]$	0	Pulse <sup>b</sup>	10	0.1	<sup>c</sup>	0.1	meanvar

<sup>a</sup> Make one mass very small compared to the rest

<sup>b</sup> Displace only one mass from equilibrium

<sup>b</sup> Short pulses of force for each perturbation

<sup>c</sup>  $T$  changes in size to fit all perturbations in simulation

## 7 Appendix A

### 7.1 Calculating Wait Time for Damped Harmonic Oscillators

Our perturbation approach requires that the forcing perturbation in the oscillator system decays to zero before the next forcing perturbation can be applied. This requires that we calculate a wait time between perturbations for every simulation. Below we discuss how this wait time is computed for the case of coupled harmonic oscillators. The general equation for a harmonic oscillator network is

$$\frac{d^2\mathbf{x}}{dt^2} + c\frac{d\mathbf{x}}{dt} - M\mathbf{x} = \mathbf{F}(t)$$

where the displacements  $\mathbf{x}$  and forcing function  $\mathbf{F}$  are vectors of length  $n$ . In our simulations, we make our forcing function be a step function of the form

$$\mathbf{F}(t) = \begin{cases} \mathbf{0} & t < t_0 \\ \mathbf{f} & t_0 \leq t \leq t_0 + \Delta t \\ \mathbf{0} & t > t_0 + \Delta t \end{cases}$$

where  $t_0$  is the time of the perturbation and  $\Delta t$  is the length of time under which the system is forced.

To solve this equation, let's first solve the homogeneous equation

$$\frac{d^2\mathbf{x}}{dt^2} + c\frac{d\mathbf{x}}{dt} - M\mathbf{x} = 0$$

This equation has  $n$  linearly independent solutions of the form  $\mathbf{x}_i(t) = \mathbf{v}_i e^{\omega_i t}$  for  $1 \leq i \leq n$ . Substituting  $\mathbf{x}_i$  into the homogeneous equation we have

$$(\omega_i^2 + c\omega_i - M)\mathbf{v}_i = 0$$

This implies that  $M\mathbf{v}_i = (\omega_i^2 + c\omega_i)\mathbf{v}_i$  so  $\mathbf{v}_i$  is an eigenvector of  $M$  with eigenvalue  $\lambda_i = \omega_i^2 + c\omega_i$ . Solving the quadratic equation for  $\omega_i$ , we have that

$$\omega_i = \frac{-c \pm \sqrt{c^2 + 4\lambda_i}}{2}$$

Now note that if  $w_i$  has a positive real part, then our solution to this differential equation will diverge. This agrees with experimental simulations. !!!Not sure how to deal with  $\pm$  in quadratic equation solution so just taking plus here!!!

Now let's assume that  $\text{Re}(w_i) < 0$  for all  $1 \leq i \leq n$ . Using our forcing function  $\mathbf{F}(t)$  defined above, let's solve our differential equation in the time  $t \in [t_0, t_0 + \Delta t]$ . In this time interval,  $\mathbf{F}(t) = \mathbf{f}$  for some constant  $\mathbf{f} \in \mathbb{R}^n$  so

$$\frac{d^2\mathbf{x}}{dt^2} + c\frac{d\mathbf{x}}{dt} - M\mathbf{x} = \mathbf{f}$$

If  $\mathbf{x}_{hom}$  is the homogeneous solution to the equation above, then as long as the matrix  $M$  is invertible, the general solution to our ODE is

$$\mathbf{x} = \mathbf{x}_{hom} - M^{-1}\mathbf{f}$$

This indeed agrees with our simulations as on this time interval when forcing is applied, each  $i$ th node's trajectory oscillates about  $(-M^{-1}\mathbf{f})_i$ . Now, after time  $t_0 + \Delta t$ , the forcing stops and  $\mathbf{F}(t) = \mathbf{0}$ . At time  $t = t_0 + \Delta t$ , on average our oscillators have positions  $\mathbf{x}(t_0 + \Delta t) = -M^{-1}\mathbf{f}$  and now their trajectories are once again dictated by the solution of the homogenous equation. So now we want to solve the ODE

$$\frac{d^2\mathbf{x}}{dt^2} + c\frac{d\mathbf{x}}{dt} - M\mathbf{x} = 0$$

on the interval  $t \in [t_0 + \Delta t, \infty)$  where  $\mathbf{x}(t_0 + \Delta t) = -M^{-1}\mathbf{f}$ . For simplicity, we can imagine that we are solving this problem on the time interval  $t \in [0, \infty)$  with initial conditions  $\mathbf{x}(0) = -M^{-1}\mathbf{f}$ .

We would like to compute an upper bound for the amount of time it would take the magnitude of our solution to decay to  $\epsilon > 0$ . Writing out the general form of the solution we have

$$\mathbf{x}(t) = \sum_{i=1}^n \mathbf{x}_i(t) = \sum_{i=1}^n \mathbf{v}_i e^{\omega_i t}, \quad \mathbf{x}(0) = -M^{-1}\mathbf{f}$$

Now we want to find a time  $t_{wait}$  so that for all  $t > t_{wait}$  we have  $\|\mathbf{x}(t)\| < \epsilon$ . Taking the magnitude of our solution

$$\|\mathbf{x}(t)\| = \left\| \sum_{i=1}^n \mathbf{v}_i e^{\omega_i t} \right\|$$

But note that all the eigenvectors  $\mathbf{v}_i$  are mutually orthogonal so

$$\|\mathbf{x}(t)\| = \sum_{i=1}^n \|\mathbf{v}_i e^{\omega_i t}\| \leq \sum_{i=1}^n \|\mathbf{v}_i\| e^{\text{Re}(\omega_i)t}$$

We also know from our initial conditions that the magnitudes of the eigenvectors  $v_i$  are chosen such that  $\mathbf{x}(0) = \sum_{i=1}^n \mathbf{v}_i = -M^{-1}\mathbf{f}$  which implies that  $\sum_{i=1}^n \|\mathbf{v}_i\| = \|M^{-1}\mathbf{f}\|$ . If  $\omega_{max}$  is the frequency with the largest (nonpositive) real part, we can write out

$$\sum_{i=1}^n \|\mathbf{v}_i\| e^{\text{Re}(\omega_i)t} \leq \|M^{-1}\mathbf{f}\| e^{\text{Re}(\omega_{max})t} < \epsilon$$

So for all  $t > \frac{\log(\epsilon/\|M^{-1}\mathbf{f}\|)}{\text{Re}(\omega_{max})}$ , we have that  $\|\mathbf{x}(t)\| < \epsilon$ . Therefore, the wait time between perturbations for a network of damped harmonic oscillators is

$$t_{wait} = \frac{\log(\epsilon/\|M^{-1}\mathbf{f}\|)}{\text{Re}(\omega_{max})}$$

## 8 Appendix B

### References

- [1]
- [2]