PERFORMANCE OF MEMORY FUNCTION BASED ON NATURALISTIC DATA

Mr. Chanin Kumpeerakij

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Physics
Department of Physics
FACULTY OF SCIENCE
Chulalongkorn University
Academic Year 2021

ประสิทธิภาพของฟังก์ชันความทรงจำที่ตั้งอยู่บนข้อมูลในธรรมชาติ

นายชนินทร์ คำภีรกิจ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาฟิสิกส์ ภาควิชาฟิสิกส์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2564

Thesis Title	PERFORMANCE OF MEMORY FUNCTION BASED ON NATURALISTIC DATA
By	Mr. Chanin Kumpeerakij
Field of Study	Physics
Thesis Advisor	Dr. THIPARAT CHOTIBUT
Thesis Co Advisor	Doctor Chaipat Chunharas Dr. Vudtiwat Ngampruetikorn

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

.... Dean of the FACULTY OF SCIENCE
(Professor Dr. POLKIT SANGVANICH)

THESIS COMMITTEE

.... Chairman
(Assistant Professor Dr. AUTTAKIT CHATRABHUTI)

.... Thesis Advisor
(Dr. THIPARAT CHOTIBUT)

.... Thesis Co-Advisor
(Doctor Chaipat Chunharas)

.... Thesis Co-Advisor
(Dr. Vudtiwat Ngampruetikorn)

.... Examiner
(Dr. ANNOP EKTARAWONG)

.... External Examiner
(Assistant Professor Dr. Chalermpol Kanchanawarin)

ชนินทร์ คำภีรกิจ : ประสิทธิภาพของฟังก์ชันความทรงจำที่ตั้งอยู่บนข้อมูลในธรรมชาติ. (PERFORMANCE OF MEMORY FUNCTION BASED ON NATURALISTIC DATA) อ.ที่ปรึกษาหลัก : อ. ดร.ธิปรัชต์ โชติบุตร, อ.ที่ปรึกษาร่วม : อ. ดร.นพ.ชัยภัทร ชุณหรัศมิ์,ดร.วุฒิวัฒน์ งามพฤฒิกร

ระบบพลศาสตร์ไม่เชิงเส้น อาทิ อัลกอริทึมโครงข่ายประสาทแบบวนซ้ำ เป็นรูปแบบอัลกอริทึมที่พิสูจน์ความสามารถในการทำนายข้อมูลที่เปลี่ยนแปลงตามช่วงเวลา อย่างไรก็ตาม การศึกษาวิธีการตั้งค่าคงที่ภายในระบบยังคงเป็นปัญหาที่ท้าทายเนื่องจากความซับซ้อนของระบบที่เกิดจากปฏิสัมพันธ์ของหน่วยย่อยภายในระบบ การศึกษานี้ได้เสนอรูปแบบการดำเนินการที่สามารถใช้ในการพิจารณาความเหมาะสมของระบบพลศาสตร์ไม่เชิงเส้นในการจดจำและทำนายข้อมูลที่เปลี่ยนแปลงตามช่วงเวลาผ่าน ผู้จัดทำได้ศึกษาระบบพลศาสตร์ไม่เชิงเส้นอย่างง่าย โดยใช้ปฏิสัมพันธ์ของออสซิลเลเตอร์สองหน่วย ซึ่งมีสัญญาณรบกวนและเสริมแรงผลักดันระบบจากสัญญาณภายนอกเพื่อให้ระบบจดจำข้อมูล ผู้จัดทำแสดงปริมาณความทรงจำและความสามารถในการทำนายข้อมูลจากปริมาณสารสนเทศร่วมระหว่างสถานะของออสซิลเลเตอร์และสัญญาณภายนอก ผลการทดลองได้แสดงว่า ความสัมพันธ์ระหว่างออสซิลเลเตอร์และความแตกต่างของลักษณะของออสซิลเลเตอร์ สามารถทำให้ระบบสามารถเพิ่มปริมาณข้อมูลที่ออสซิลเลเตอร์จดจำเกี่ยวกับสัญญาณภายนอกเนื่องจากระบบเพิ่มประสิทธิภาพของสัญญาณต่อสัญญาณรบกวน ซึ่งเป็นผลจากความเชื่อมโยงของออสซิลเลเตอร์ที่เพิ่มมากขึ้น วิทยานิพนธ์นี้นำเสนอได้นำเสนอวิธีการจัดการอย่างเป็นระบบในการปรับระบบพลศาสตร์ไม่เชิงเส้นให้สามารถมีความทรงจำและความสามารถในการทำนายได้สูงสุดสำหรับข้อมูลที่เปลี่ยน

สาขาวิชา	ฟิสิกส์	ลายมือชื่อนิสิต ....
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม ....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม ....

6270128123 : MAJOR PHYSICS
KEYWO
RD:

Chanin Kumpeerakij : PERFORMANCE OF MEMORY FUNCTION BASED ON NATURALISTIC DATA. Advisor: Dr. THIPARAT CHOTIBUT Co-advisor: Dr. Chaipat Chunharas,Dr. Vudtiwat Ngampruetikorn

Nonlinear dynamical systems, such as well-tuned recurrent neural networks, have proved a powerful tool for modeling temporal data. However, tuning such models to achieve the best performance remains an outstanding challenge, not least because of the complex behaviors that emerge from interacting microscopic constituents. Here, we consider a minimal model of two interacting phase oscillators coupled to a thermal bath and driven by a common signal. We quantify the memory and predictive capability of the system with the mutual information between the phases of oscillators and the signals at different times. We show that the interaction and heterogeneity between oscillators can increase the information between the system and the movement. We attribute this behavior to an increase in the effective signal-to-noise ratio, resulting from a stronger correlation between the oscillators. Our work offers the first step toward a systematic approach to optimizing interacting nonlinear dynamical systems for memorizing and predicting temporal patterns.

Field of Study:	Physics	Student's Signature ....
Academic Year:	2021	Advisor's Signature ....
		Co-advisor's Signature ....
		Co-advisor's Signature ....

ACKNOWLEDGEMENTS

I could not have undertaken this journey without Dr. Thiparat Chotibut, Dr. Vudtiwat Ngampruetikorn, and Dr. Chaipat Chunharas, my supervisors. They were always supporting knowledge, idea, patience, and feedback. Many thanks to my friends and members of Chula Complex and Intelligence System group (CHICS) for their programming and research experience suggestions. This study would not be possible without financial support from The Development and Promotion of Science and Technology Talents Project (DPST). I am also grateful to my family, especially my parents, sister, and cousins, for their moral support.

Chanin Kumpeerakij

TABLE OF CONTENTS

Page

List of Figure

Page

1. Introduction

1.1 Introduction

Dynamical systems are useful models for wide-ranging natural phenomena from population expansion (Strogatz, 2018) and bug outbreaks (Acebrón et al., 2005) to synchronized flashing of fireflies (Rodrigues et al., 2016). Indeed carefully designed dynamical systems, such as recurrent neural networks (RNN), have proved highly successful in modeling temporal data (Elsaraiti & Merabet, 2021; Muncharaz, 2020). However, tuning the parameters of large dynamical systems can be a challenge (Boedecker et al., 2012; Gupta et al., 2021; Nakajima & Fischer, 2021) ,especially because the interactions among microscopic elements can result in complex, high-dimensional dynamics.

Recurrent neural networks (RNNs) are a class of artificial neural networks (Schneidman et al., 2003). Such models are among the most widely used methods for predicting temporal data sequences and have been shown to excel in a diverse range of applications, including forecasting financial time series (Sako et al., 2022) ,sentimental analysis (Tangpanitanon et al., 2022) or text translation (Lipton, 2015). Deeper understanding of how RNNs achieve their performance is of both fundamental and practical interest. This requires a quantitative analysis of the interplay between nonlinearity and interactions between the building blocks of the dynamical systems on which RNNs are based.

However, the interaction between microscopic constituents of large nonlinear dynamical system is difficult to analyze. Here, I consider a simplified model of nonlinear dynamical systems that allows for an exploration of how system size, interactions and heterogeneity give rise to the ability to memorize and predict. More specifically, I consider the Kuramoto model, interacting phase oscillators system, because this model exhibits a number of interesting emergent behaviors (Acebrón et al., 2005; Rodrigues et al., 2016). The emergent properties of Kuramoto model used to explain in many physical system, including synchronization of metronomes placed on a freely moving base (Pantaleone, 2002), or neuron activities (Breakspear et al., 2010; Cumin & Unsworth, 2007; Schmidt et al., 2015). Consequently, I analyze the memory and predictive capabilities of the system through the principled lens of mutual information according to the brain functions.

Nevertheless, the Kuramoto model is too complicated to understand the interaction of microscopic elements due to the number of units. As a result, I consider two interacting phase oscillators based on Kuramoto model because it is the smallest system with complete system components. My model can be easily generalized to incorporate more units, different nonlinearity as well as nonreciprocal interactions. My results offer a first step towards more complete understanding of how computation emerges from interactions.

The thesis is divided into five chapters. Chapter two explains the background knowledge, including interacting phase oscillator systems and information theory. Chapter three presents the methodology, system overview, and how to quantify the performance of the system. Chapter four describes my findings and discusses the physical intuitions behind my numerical observations. Chapter five concludes the thesis and highlights directions for future work.

2. Background Knowledge

This chapter provides the necessary background information for understanding this thesis. The author divides this chapter into two sections. The theoretical examination of the interacting phase oscillators system is in the first section. This data describes the relationship between the system and its performance. Another section is information theory, which is the knowledge we use to quantify the information encoding capabilities of a system.

2.1 Interacting Phase Oscillators

The interacting phase oscillators system (see Figure 1) that we investigated is adapted from the "**Kuramoto model**" (Kuramoto & Nishikawa, 1987; Rodrigues et al., 2016). The system is constructed with N coupled phase oscillators ($\theta_i(t)$) that have their natural frequency (ω_i) drawn from a distribution $g(\omega)$. The following equation governs the dynamics of the system

$$\frac{d\theta_i}{dt} = \omega_i + \sum_{j=1}^{N} K_{ij} \sin(\theta_j - \theta_i) \tag{1}$$

where K_{ij} denotes the coupling strength between oscillator i and j. Intuitively, each oscillator runs with its frequency individually ($\theta_i(t + \Delta t) = \theta_i(0) + \omega_i \Delta t$), but (positive) couplings encourage their phases to align. For weak couplings, each oscillator behaves as if they are independent. On the other hand, strong (positive) couplings align the phases of oscillators and all oscillators share the same dynamics. This is the advantage that we tend to use to determine this thesis.

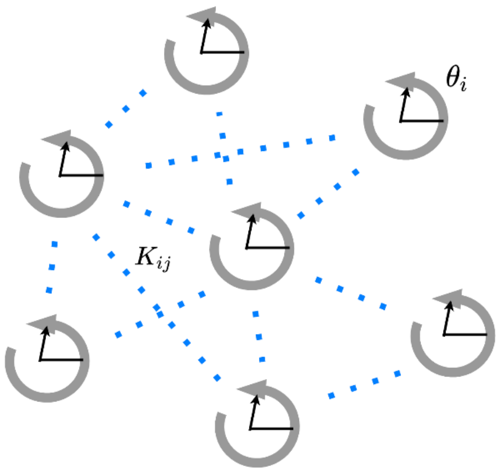


Figure 1 Illustration of interacting phase oscillators

Each circular arrow represents the time evolution of the phase of each oscillator. Blue dot line illustrates the interaction between oscillators.

2.1.1 Mean-field coupling model and Phase transition

Mean-field coupling model is the interacting phase oscillators system in the limit of infinite oscillators, $N \to \infty$, and identical coupling strength, $K_{ij} = K/N > 0$. The classical analysis of synchronization was in the form of time-dependent order parameter $r(t) \in [0,1]$, defined as

$$r(t)e^{i\psi(t)} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j(t)}, \qquad (2)$$

where $\psi(t)$ is the average phase of the system and the order parameter $r(t)$ denotes order parameter, $0 \leq r(t) \leq 1$. If $r(t) = 0$, it means oscillators does not align the average phase $\psi(t)$. So, $r(t) = 0$ implies that oscillators act incoherently. In contrast, $r(t) = 1$ indicates that oscillators run with the same phase as average phase. As a result, $r(t) = 1$ implies that oscillators are coherence. We can rearrange Equation 1 into

$$\dot{\theta}_i = \omega_i + Kr\sin(\psi - \theta_i)\ ;\ i = 1,2,3,\ldots,N \qquad (3)$$

and then the system is coupled to the average phase with coupling strength Kr. In thermodynamic limit, we can use the probability distribution of phase of oscillator$\rho(\theta,\omega,t)$ to find the order parameters, through the following equation averaged over phase and frequency.

$$re^{i\psi} = \int_{-\pi}^{\pi}\int_{-\infty}^{+\infty} e^{i\theta}\,\rho(\theta,\omega,t)g(\omega)d\theta d\omega \qquad (4)$$

where $\rho(\theta,\omega,t)$ can be obtained from the continuity equation with an angular or drift velocity $v = \omega + Kr\sin(\psi - \theta)$,

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial\theta}[\omega + Kr\sin(\psi - \theta)]\rho. \qquad (5)$$

Consequently, we can determine the system’s behavior using the order parameter and show it in the phase transition.

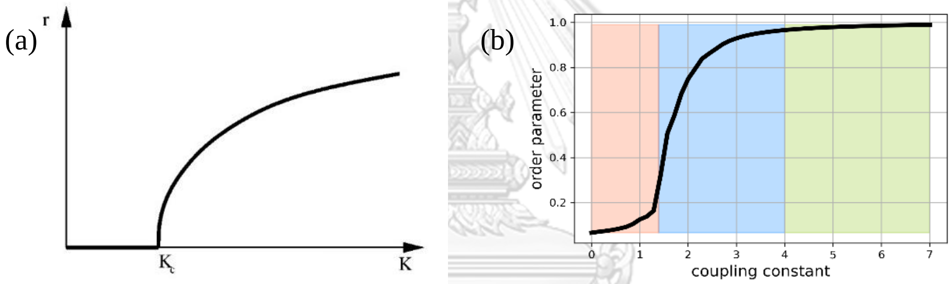


Figure 2 Phase transition of Mean-field model
(a) Theoretical results (b) simulation results of 200 oscillators and natural frequency drawn from Gaussian distribution.

As the coupling strength increases, the system will change from incoherence into partial coherence and then completely coherence. The transition happened at critical point point at $K_c = 2/[\pi g(0)]$ which depends on the distribution of natural frequency, $g(\omega)$ (Kuramoto & Nishikawa, 1987), (Acebrón et al., 2005). In a system with finite number of units (see Figure 2), the sharp phase transition manifests as a smooth crossover.

2.1.2 Mean-field coupling with white noise and phase transition

This section focuses on an alternative model of mean-field coupling oscillators by adding white noise into the system and specifying the natural frequency distribution. The natural frequency is drawn from the discrete bimodal distribution, which limits the value to $\pm\omega_0$ where $\omega_0 \in \mathbb{R}$. The dynamics of this system is governed by the following equation

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \sin\left(\theta_j - \theta_i\right) + \eta_i(t),\ i = 1,2,3,\ldots N. \tag{6}$$

where $\eta_i(t)$ is the independent white-noise from stochastic process with expectation value; D is Diffusion coefficient

$$\langle\eta_i(t)\rangle = 0, \langle\eta_i(t)\eta_j(t')\rangle = 2D\delta(t - t')\delta_{ij}, . \tag{7}$$

In addition to order parameter analysis, we can reformulate eq. 6 into the following equation

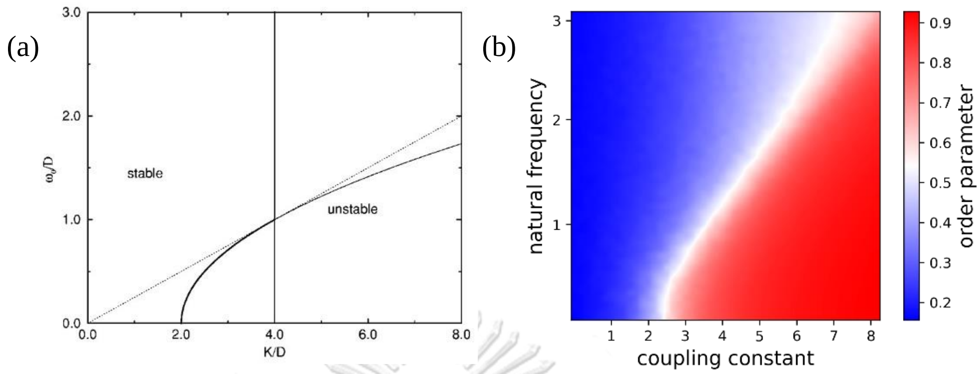
$$\dot{\theta}_i = \omega_i + Kr\sin(\psi - \theta_i) + \eta_i(t),\ i = 1,2,3,\ldots N, . \tag{8}$$

The presence of the noise term means that we have to analyze the Fokker-Planck equation instead of the continuity equation to compute the probability distribution of phase oscillators. However, it is not easy to analyze the Fokker-Planck equation of a large system. Fortunately, we can study the probability distribution of one oscillator as the following equation to understand the dynamic of system.

$$\frac{\partial\rho}{\partial t} = D\frac{\partial^2\rho}{\partial\theta^2} - \frac{\partial}{\partial\theta}(v\rho) \tag{9}$$

$$v(\theta, \omega, t) = \omega + Krsin(\psi - \theta) \tag{10}$$

Consequently, the behavior analysis shifts to focus on the state of probability distribution instead because once we find the phase distribution, we can express the behavior of the system.



(a) Stability analysis of probability distribution ((Bonilla et al., 1998)) (b) simulation

Figure 3 Phase transition of Mean-field model with white noise
results of 30 oscillators.

The stability analysis of phase probability distribution for incoherent initialization ($\rho_0 = 1/2\pi$) is illustrated in fig.3(a). (Bonilla et al., 1998) separate the *stable* and *unstable* with the line $K/D = 2[1 + (\omega_0/D)^2]$. The stable indicates that probability distribution of the phase is uniformly distributed over time. The stable probability distribution leads to an incoherent system because there is no change in population density. The unstable probability distribution evolves into high density at a specific phase, which means they are coherent. A simulation of 30 oscillators demonstrated the similar properties of the system that variate the behavior from incoherence to coherence while changing the parameters, see Figure 3(b)

In this thesis, I investigated the mean-field coupling model with white noise because it reduces the number of control parameters and it is consistent with the natural frequency distribution.

2.2 Information Theory

Information theory is a study of quantifying information in delivered messages for communication. It can be used to quantify how much information is in an event or in a random variable. Information theory was proposed and developed by Claude Shannon (MacKay, 2002). Quantifying information is the concept of measuring how much surprise the occurrence is. If it hardly happens, it has high information because we are surprised when the event occurred, and vice versa. Consequently, we can determine the information through the probability of an event.

$$I(x) = -\log_2\big(p(x)\big). \tag{11}$$

Equation (11) demonstrates the quantifying of information of an event x through the probability that the event x will occur. The minus sign restricts the value in the domain $[0, \infty)$. Zero information means that there is no surprise in the event because it certainly happens. The base-2 logarithm is used to represent the information in *“bit”* unit.

Example: Consider the information of tossing coin experiment

In the previous topic, the tossing coin experiment is one of a stochastic process. Tossing coin provides the random variable with “head" and “tail" values. A fair coin says that it has a probability of facing “head" and “tail" at 0.5. So, we can determine the information in a fair coin.

$$I(fair\ coin) = -\log_2(0.5) = -\log_2\left(\frac{1}{2}\right) = 1.00\ bits$$

Information in a fair coin is 1 bit. It shows that a fair coin has two representations that could face up.

2.2.1 Shannon Entropy

Shannon entropy or information entropy ($H(x)$) quantifies information uncertainty in a random variable (MacKay, 2002). Intuitively, information entropy is the average information of available states in the random variable. The previous topic shows the information value expressing how rare the event is. So, the information entropy explains the average number of information required to represent the random variable. Information entropy of a random variable X can be calculated with N discrete states from the following equation.

$$H(X) = -\sum_{i=1}^{N} p\,(x_i)\log_2\big(p(x_i)\big)\ ;\ x_i \in X. \tag{12}$$

2.2.2 Joint Entropy and Conditional Entropy

From Equation 12, we can substitute with joint probability or conditional probability. So, we obtain the information entropy that express the uncertainty about the correlation between two random variables instead. The joint entropy and conditional entropy can be calculated from the following equations.

$$H(X,Y) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p\left(x_i, y_j\right) \log_2\left(p\left(x_i, y_j\right)\right) \; ; \; x_i \in X, y_j \in Y \tag{13}$$

$$H(X|Y) = -\sum_{j=1}^{M} p\left(y_j\right) \sum_{i=1}^{N} p\left(x_i|y_j\right) \log_2\left(p\left(x_i|y_j\right)\right) \tag{14}$$

$$= -\sum_{i=1}^{N}\sum_{j=1}^{M} p\left(x_i, y_j\right) \log_2\left(p\left(x_i|y_j\right)\right) \; ; \; x_i \in X, y_j \in Y \tag{15}$$

Equation 13 is the calculation of joint entropy, which use the joint probability between discrete random variables X and Y. Another equation, Equation 14, is the general form of conditional entropy. Equation 14 can be written in an alternative form as Equation 15.

2.2.3 Mutual Information

In the previous topic, conditional entropy can measure the uncertainty about the correlation between two random variables. We can say that conditional entropy is information that Y does not know about X, referred to as Equation 14 because we already knew information about Y in the condition. Still, there is uncertainty about knowing X.

In the context of Equation 14, if we know the information entropy of X and the conditional entropy of X given Y, then we can determine the information that Y knows about X. We define the term information when two random variables know information about each other as “**mutual information**" ($I(X;Y)$), where the information entropy of X is called **output entropy**, and the conditional entropy is called **noise entropy**.

$$I(X;Y) = H(X) - H(X|Y). \tag{16}$$

Figure 4 illustrates the relation between information entropy and mutual information. There are two random variables, X, and Y, sharing the mutual information that both random variables know about each other.

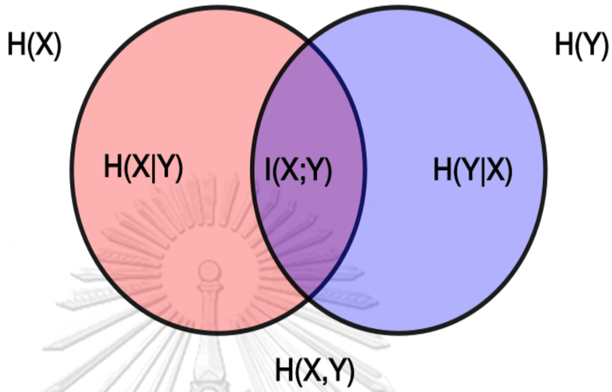


Figure 4 Entropy-Mutual Information Relation in Venn Diagram

3. Methodology

In this thesis, I used an information-theoretic framework to investigate the impact of interacting phase oscillator characteristics on memory and predictive capability (Palmer et al., 2015a). According to the interacting phase oscillator model, we have three parameters to experiment: natural frequency, coupling constant, and input coupling constant, which is the strength of external signals. Consequently, we must vary the system parameters and investigate the memory and predictive capability. This section will provide the approach and hypothesize the results.

3.1 System Overview

According to the interacting phase oscillators, a system was built from the natural frequency of each oscillator, coupling terms, and white noise. The system will be able to encode terms based on the external input. Consequently, the external input-driven terms, $K_{in}\sin(\theta_i - x_{in})$, were then added to the system, Equation 17.

$$\frac{d\theta_i}{dt} = \omega_i + \frac{1}{N}\sum_{j=1}^{N} K_{ij}\sin\left(\theta_j - \theta_i\right) + K_{in}\sin(\theta_i - x_{in}) + \eta_i(t) \quad (17)$$

where θ_i and ω_i denote the phase and natural frequency of each oscillator, N number of oscillators, K_{ij} coupling constant between oscillator i^{th} and j^{th}, K_{in} input-coupling constant, x_{in} external input signal, and $\eta_i(t)$ Gaussian white noise. The input-coupling term will assign the external signal, which will be coupled with oscillators i^{th}, whose strength will be controlled by K_{in}. In this work, we focused on two interacting phase oscillators and binary signal input drawn from the Poisson process. Consequently, the Equation 17 will be reduced to Equation 18.

$$\begin{aligned} \frac{d\theta_1}{dt} &= \omega_1 + \frac{1}{2}K_{12}\sin(\theta_2 - \theta_1) + K_{in}\sin(\theta_1 - x_{in}) + \eta_1(t) \\ \frac{d\theta_2}{dt} &= \omega_2 + \frac{1}{2}K_{21}\sin(\theta_1 - \theta_2) + K_{in}\sin(\theta_2 - x_{in}) + \eta_2(t) \end{aligned} \quad (18)$$

The system is illustrated in the following figure, Figure 5.

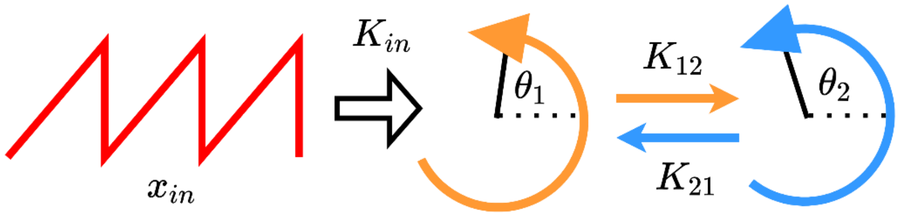


Figure 5 Illustration of two interacting phase oscillators

Two oscillators are interacting with each other and coupling to the external signal. Variable in the system relied on Equation 18

Next, we computed the time-evolution of the phases of oscillators through stochastic integration method (Rößler, 2010).

3.2 Estimation Mutual Information between phases of oscillators and inputs

To estimate mutual information between two variables required a collection of the event for both variables. Consequently, we collected the time evolution of oscillator's phases from each setup's parameters with *5000 realizations of input signal*. Then we can estimate the mutual information between the phases of oscillators and inputs to quantify memory and predictive capability.

3.2.1 Oscillator phase discretization

The phases of oscillators are continuous variables. It is not easy to compute mutual information with a binary state, a discrete variable. So, we propose discretizing the phases of oscillators into a finite number of states. The procedure for discretizing is done through the following steps.

1. Re-scale phase in the domain $[0,2\pi]$ since the phase obtained from the dynamical equation is a continuous variable in real space, so the phase still goes around in circle.

2. Discretize re-scaled phase by using bins over a circle. For example, if we need to separate phases into four states, then we binned phase on the circle into four parts, $\left(\left[0,\frac{\pi}{2}\right],\left[\frac{\pi}{2},\pi\right],\left[\pi,\frac{3\pi}{2}\right],\left[\frac{3\pi}{2},2\pi\right]\right)$. Next, we labeled the phases of oscillators according to the label that each phase fell in. (*Methodology Figure*)

In this thesis, we focused on four discretization states because they can capture the activity that occurred and support computational efficiency.

3.2.2 Estimating Mutual Information

We estimated the collection of the phases of oscillators and input signal at different times to quantify the memory and predictive capability of the system. The estimation is done through the following equation.

$$I\big(\theta_1(t), \theta_2(t); x_{in}(t')\big) = \sum_{\theta_1,\theta_2} \sum_{x_{in}} p\,(\theta_1, \theta_2|x_{in}) p(x_{in}) \log_2 \left[\frac{p(\theta_1, \theta_2|x_{in})}{p(\theta_1, \theta_2)}\right] \quad (19)$$

According to Equation 19, this is the mutual information between the system state (the phases of two oscillators) at time t and the input signal at time t'. Consequently, we quantified **memory** by estimating mutual information of the collection of the input signal in the past $(t' < t)$. The **predictive capability**, on the other hand, is estimated with the input signal collection in the future $(t' > t)$, (Palmer et al., 2015b) .

3.3 Effect of three parameters on the system characteristic

The setup of parameters varies the system characteristics. So, we aim to show how each parameter affects memory and predictive capability based on behavior analysis. Here, we adjust three parameters to one at a time investigated their effects.

3.3.1 Input-Coupling Constant

The first parameter that we investigate is the input-coupling constant (K_{in}) because its strength could show how strongly the external signal affects the system. Thus, we hypothesized that the amount of memory and predictive capability would increase as K_{in} was increased.

However, if K_{in} is too strong, the system will track the input rather than encoding any information about it. So, we would see the mutual information between the phases of oscillators and input signal as the same information between input and input signal.

3.3.2 Coupling Constant

The coupling constant (K_{ij}) is the next parameter we looked at. This thesis only determined the fully connected scheme ($K_{12} = K_{21} = K$) because we could readily link back to the behavior of the broader system.

The effect of the coupling constant is to adjust the dynamics of oscillators to be synchronized. Intuitively, this means each oscillator interacted with each other and shared information. So, we hypothesized that interaction between oscillators helps the system increase memory and predictive capability.

Nevertheless, coupling constant could help increasing information. Coupling constant also drives the system into the synchronized state where all the oscillators behave the same. Consequently, this scheme would decrease memory and predictive capability when K was too strong because both oscillators acted like one.

3.3.3 Natural Frequency

The natural frequency (ω_i) is the last parameter we considered. Natural frequency is the characteristic of an oscillator. In this experiment, ω_i is drawn from set $\{1, -1\}$. So, there are only three combinations: $\{1,1\}, \{1, -1\}, \{-1, -1\}$.

We hypothesized that the setup of $\{\omega_1, \omega_2\}$ as $\{1, -1\}$ could help the system increasing memory and predictive capability because this structure sets each oscillator to have different character. That means both oscillators would behave differently and could encode other parts of the input.

3.4 Methodology Conclusion

We start the process by collecting the response from the dynamical equation with the different realization of inputs. After that, we discretize the state of oscillators and estimate the mutual information between oscillators and inputs at different times. Consequently, we variate the setup parameter of the system, including input-coupling constant (K_{in}), coupling constant (K), and natural frequency configuration, to demonstrate the relationship between the system's characteristic and encoding efficiency.

4. Results

In this chapter, we investigated the memory and predictive capabilities for noninteracting and interacting systems with homogeneous and heterogeneous natural frequencies.

4.1 Effect of Input-Coupling Constant

4.1.1 Estimation Mutual Information at Different Time

First, we estimated mutual information between the phases of both oscillators and the external signals. The result is shown in Figure 6.

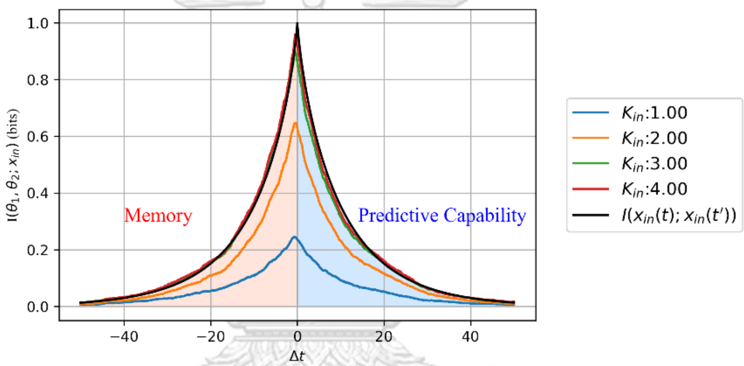


Figure 6 Mutual information compared at different times

This results compared different setup of input coupling constant. The solid black ($I(x_{in}(t); x_{in}(t'))$) label is mutual information estimated from probability distribution of input compared at different times. The input's mutual information maximize at 1 bits because of binary signal. It demonstrates that strength of external signal help increasing both memory and predictive capability, referred to the area under curves.

The results reveal that memory and predictive capability improve as the input-coupling constant increases. Similarly, as the input-coupling constant increases, so do the memory and predictive capabilities. Consequently, we used equal time mutual information ($I(\theta_1, \theta_2; x_{in})[\Delta t = 0]$) to show memory and predictive capability as encoded information in the system in the next sections.

4.1.2 Equal Time Mutual Information Without Interaction Between Oscillators

Figure 6 shows the equal time mutual information changed in the input-coupling constant with respect to .

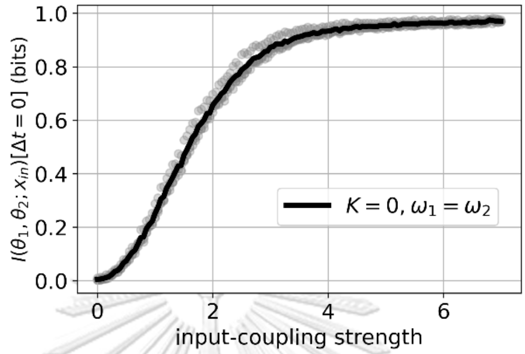
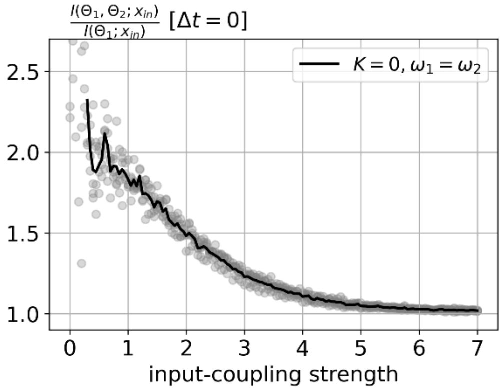


Figure 7 Equal time mutual information without interaction

When K_{in} is closed to zero, memory and predictive capability are approximately zero because the system encodes noise instead of the external signal. Next, the encoded information increases along with the strength of the external signal until it reaches the maximum value at one bit because K_{in} increases the signal to noise ratio.

4.1.3 State Space Efficiency

Next, we compared the information about external input in one and two oscillators to show how effectively the system uses state-space to encode information. Intuitively, two oscillators would encode more information than one oscillator because we increased statistical independence. However, K_{in} caused both oscillators to act similarly, and therefore two oscillators behave like one instead.

(a) 

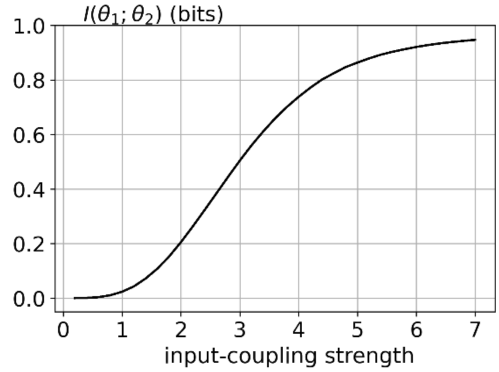
(b) 

Figure 8 How system without interaction using state space to encode

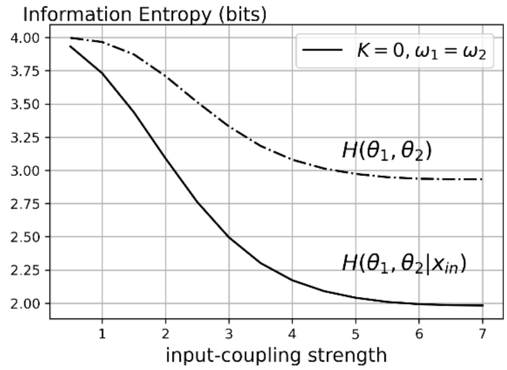
(a) Ratio between mutual information estimated from two oscillators and one oscillator. (b) Mutual information between oscillators' phases versus K_{in}

Figure 8(a) shows two main points at the extreme limit. First, the two oscillators behaved differently in a noise-dominated domain with a low input-coupling constant because Gaussian white noise disrupted the system ability to encode information. As a result, two oscillators can use their state-space to encode information more efficiently than one oscillator because noise increases the statistical independence between oscillators. We can inspect the mutual information between oscillators as shown in Figure 8(b). It demonstrates that both oscillators know almost nothing about each other, which means they do not correlate with each other.

On the other hand, the oscillators were likely to behave similarly as the input-coupling constant was increased until they behaved identically in the input-dominated regime. However, as seen in Figure 8(b) the mutual information between oscillator phases approached one, which means the system replicates input. As a result, the oscillators were inefficient in using state-space in this regime.

4.1.4 Output and Noise Entropy

This section inspects the result in Figure 7 by using the extraction of mutual information, output entropy, and noise entropy. Because we used the combination of four discretization states of both oscillators, one and two oscillators have output entropy maximum at 2 and 4 bits.

(a)


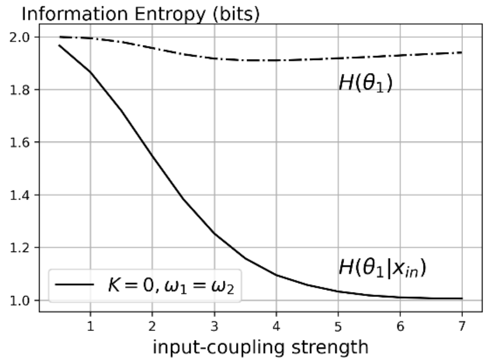
(b)


Figure 9 Output and Noise entropy of system without interaction

(a)Two oscillators' response compare with input (b) One oscillators' response compare with input

Figure 9(a) and 9(b) showed that noise entropy, or the input information that oscillators were unaware of, was very high when the system was set up in a noise-dominated regime. Then noise entropy decreased while the input-coupling constant increased until saturated at the input-dominated regime. Output entropy also performed the same trends, but the gap between output and noise entropy indicateds that the system gains information depending on the external input strength.

4.2 Effect of Coupling Constant

We still investigated the system performance in the same scheme, but we now added the coupling constant and interaction between oscillators.

4.2.1 Equal Time Mutual Information With Interaction Between Oscillators

Two oscillators can share their present state simultaneously because of their interaction. This activity demonstrated how two oscillators share information. As a result, the ability to encode data is improved, see Figure 10.

As expected, the interaction aids the system in encoding more data. On the other hand, the encoding performance depends on the strength of the external signal, referred to as the system with interaction, which is not significantly different from the system without interaction.

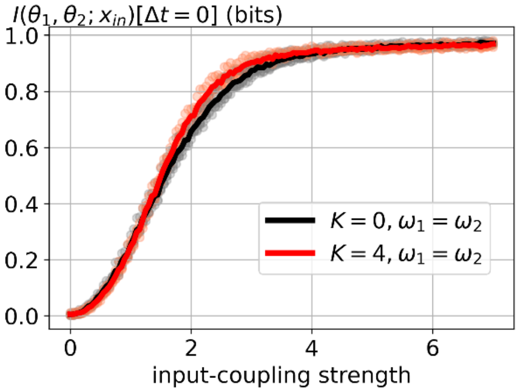


Figure 10 Equal time mutual information with interaction

Red line and *Black line* are the results from the systems with and without interaction. The results are estimated from the phase of two oscillators and inputs.

4.2.2 State Space Efficiency

Now we can see how effectively a system with interaction use state-space to encode external signal.

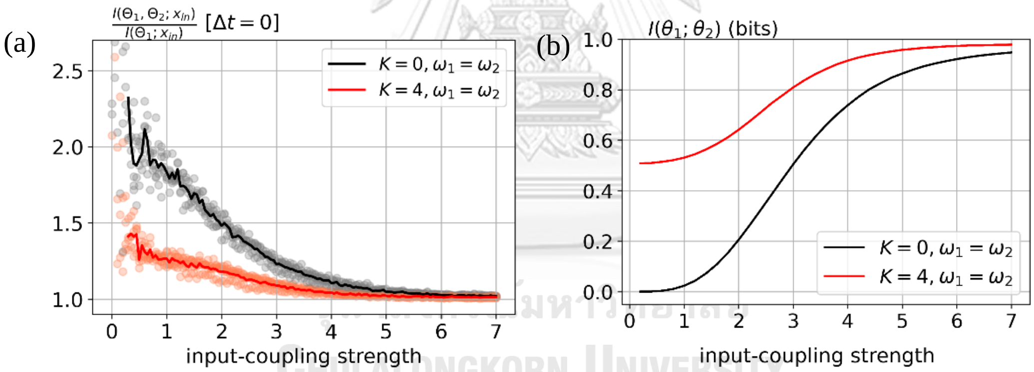


Figure 11 Comparing Encoding efficiency between systems with and without interaction

(a) Efficiency ratio between estimation of two and one oscillator. (b) Mutual information between oscillators' phases while changing K_{in}

Trends for a system with interaction were similar to the system without interaction, but the ratio was lower, as shown in Figure 11(a). It signified that the interaction caused both oscillators to behave more similarly. As same as results in Figure 11(b), mutual information between both oscillators also showed that both oscillators act equally with each other. Then the similarity of the phases of oscillators caused the system uses less state space to encode information. However, both oscillators performed the same as the system without interaction in an input-dominated regime.

4.2.3 Output and Noise Entropy

According to the results of the system without interaction, we will investigate output and noise entropy, comparing previous results with a system with interaction through the following figure.

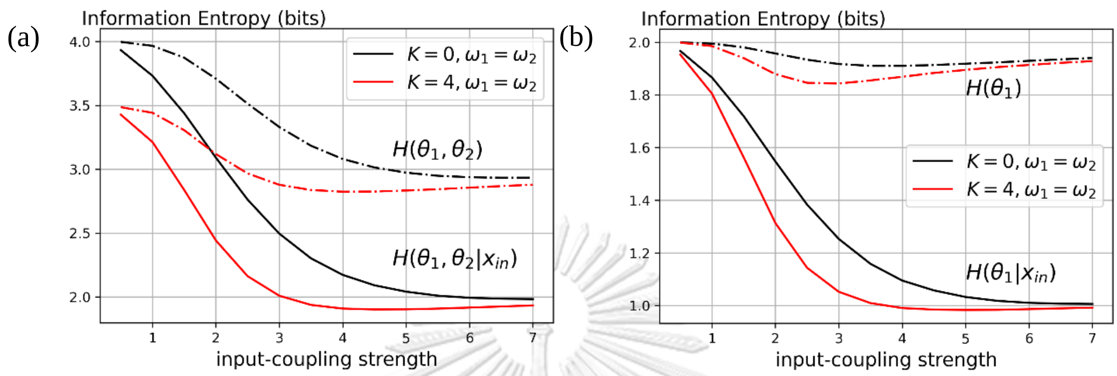


Figure 12 Output and noise entropy comparing the system with and without interaction

(a) Output and noise entropy for two oscillators (b) Output and noise entropy estimated from one oscillators' response

We can see that the output entropy of the system with interaction decreases because the system uses less state space than the system without interaction. However, the noise entropy lowered when the system added the interaction. This result revealed that the interaction helps oscillators share information and decrease noise.

4.3 Effect of Setup of Natural Frequency

This section will look at the impact of different natural frequency setups on encoding efficiency and compare our findings to prior ones.

4.3.1 Equal Time Mutual Information

First, we considered equal time mutual information set up with different natural frequency compared with previous results through the following figure, Figure 13.

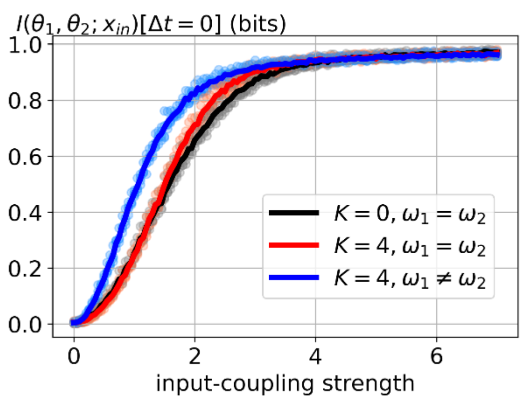


Figure 13 Equal time mutual information with interaction and differentiate natural frequency

The results revealed similar tendencies to previous findings, although it gets more information than in the system with interaction in the noise-dominated regime. Because of increasing heterogeneity of the system, oscillators with different characteristic helps to suppress noise even more. However, the trend converged to the maximum value in an input-dominated regime.

4.3.2 Input Response Efficiency

The interesting point is around noise-dominated regimes as in Figure 14(a) and 14(b). In both natural frequency with interaction configurations, the trend of the ratio of mutual information calculated from two and one oscillators is almost identical.

(a)

(b)

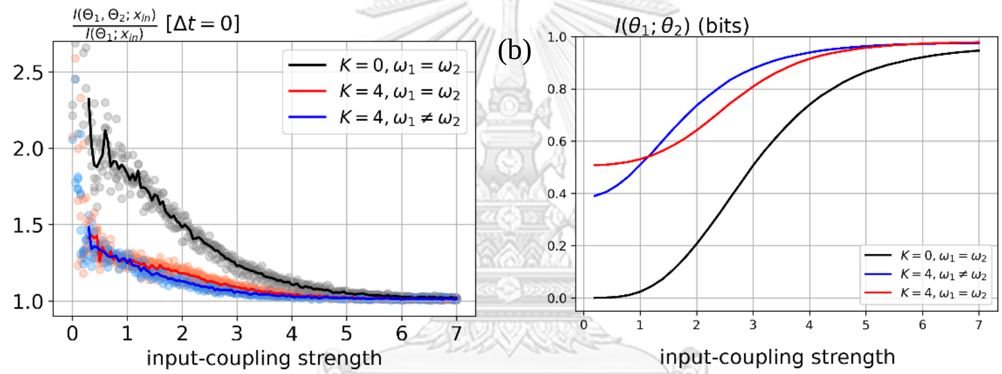


Figure 14 Input responded efficiency of the system with different natural frequency configuration (a) Ratio between mutual information estimated from two oscillators and one oscillator (b) Mutual information between oscillators' phases while changing

Figure 14(b) indicates that both oscillators in a system with different natural frequency configurations act differently in a noise-dominated regime. Surprisingly, the system with the same natural frequency in the center of the trend of mutual information between oscillators behaves more like both oscillators with distinct natural frequency configurations. The cause of this occurrence is that the noise was not only disturb the system, but it also perturbed the system into other state. Hence, the interacting phase oscillators reached the undercover state to encode the information. However, each oscillators shared information together and then oscillator acted similar to one oscillator because they shared state together.

4.3.3 Output and Noise Entropy

We compared output and noise entropy between a system with different natural frequency configurations and prior results in this topic. We illustrate the results through Figure 15.

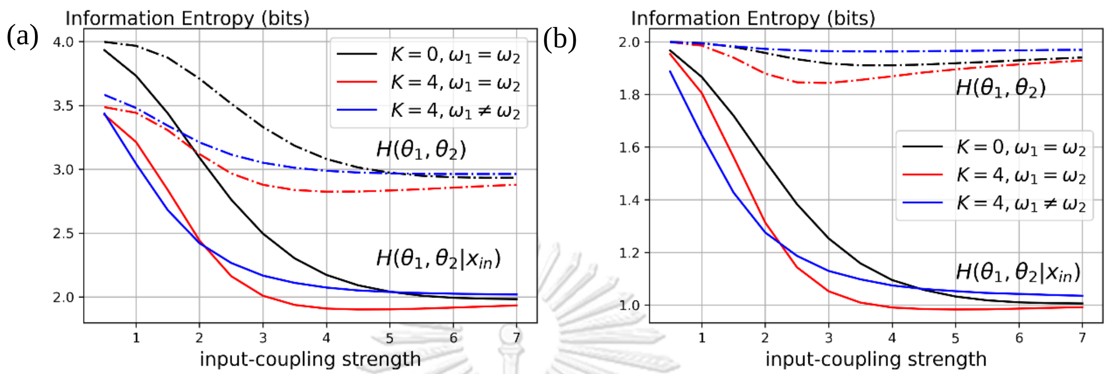


Figure 15 Output and noise entropy of system with different configurations
(a) Output and noise entropy for two oscillators (b) Output and noise entropy for one oscillator

Natural frequency configurations enhanced output entropy, as expected because the varied features of oscillators promote statistical independence. However, as statistical independence increased, noise entropy increased as well. Noise entropy was low and significant in noise-dominated regimes because of the external input response of the system. Furthermore, we found that the different natural frequency configurations increase output entropy in one oscillator more than the system without interaction, followed by the increase of noise entropy in an input-dominated regime. The results indicate that systems with interaction and different natural frequency configurations can simultaneously reduce noise effects and respond to the external signal.

5. Summary

5.1 Summary

The purpose of this graduation thesis, titled *Performance of Memory Function Based on the Naturalistic Data*, is to examine responses to the external input of a simplified system of interacting phase oscillators at various system parameters. Dynamical systems, and recurrent neural network in particular, have proven to be powerful for memorizing and predicting temporal data. However, the performance of these systems relies heavily on carefully tuned parameters. Finding the right operating point is a challenging problem especially for large dynamical systems, required for complex, high-dimensional data. To understand how system parameters affect performance, I consider a minimal model of two interacting oscillators and investigate its memory and predictive capabilities, quantified with mutual information.

The first exploration is an effect of the strength of coupling to external input signals. It indicated that the input-coupling constant affects encoding efficiency because the amount of encoded information depends on this value. The second exploration is the effect of the coupling strength between oscillators. It demonstrates that sharing information between oscillators help increases encoded information, but reduce of statistical independence. The last exploration is an effect of natural frequency configuration. The natural frequency is a characterization of each oscillator. Difference in natural frequency configuration helps increase the possible encoding state of the system and actively respond to the input. If we combine the different configuration of the natural frequency with interaction, we can increase the encoding capability of the system.

These results illustrate the computational performance of dynamic systems, which emerges from the collective behavior. Even though the current study focuses on a simple system, it is easy to generalize to larger, more complex systems.

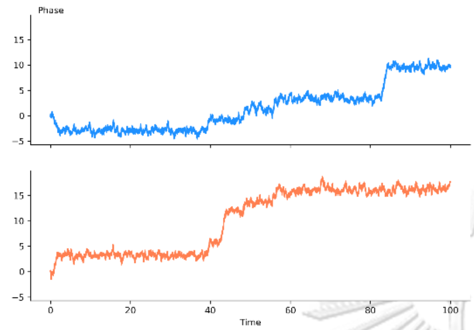
5.2 Future Work

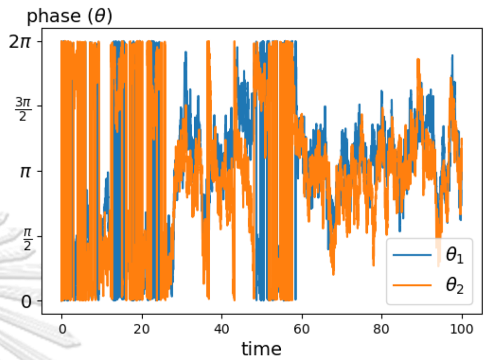
There are available points to study this simple system. First, this thesis examines only the fully connected interacting phase oscillator system. We can adjust the interaction strength and arrangement to check the system’s efficiency. Next, we aim to study this system with other external input to express the encoding efficiency. After that, we then experiment with a larger system to determine the memory and predictive capabilities.

Methodology Figure

Discretization phases of oscillators

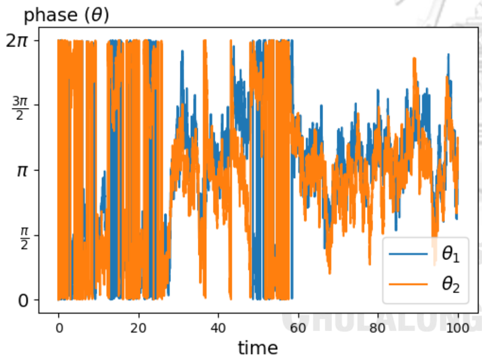
Re-scaling phases of oscillators

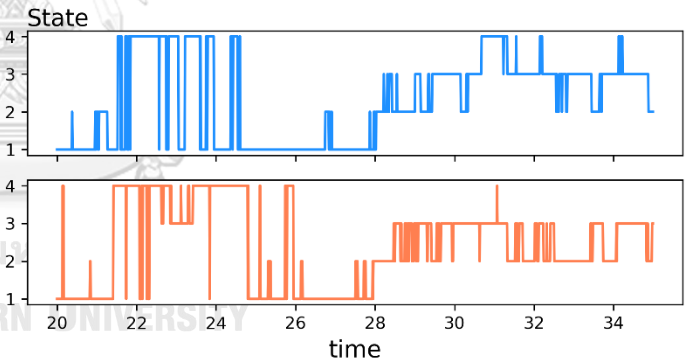
(a)


(b)


(a) Raw phase collected from each oscillator (b) Re-scaled phases

Phase discretization

(a)


(b)


(a) Re-scaled phases of oscillators (b) Discretizing phases into 4 states

REFERENCES

Acebrón, J. A., Bonilla, L. L., Vicente, C. J. P., Ritort, F., & Spigler, R. (2005). The Kuramoto model: A simple paradigm for synchronization phenomena [https://doi.org/10.1103/revmodphys.77.137]. *Reviews of Modern Physics*, *77*(1), 137–185-137–185.

Boedecker, J., Obst, O., Lizier, J. T., Mayer, N. M., & Asada, M. (2012). Information processing in echo state networks at the edge of chaos [https://doi.org/10.1007/s12064-011-0146-8]. *Theory in Biosciences*, *131*, 205-213.

Bonilla, L. L., Vicente, C. J. P., & Spigler, R. (1998). Time-periodic phases in populations of nonlinearly coupled oscillators with bimodal frequency distributions [https://doi.org/10.1016/s0167-2789(97)00187-5]. *Physica D: Nonlinear Phenomena*, *113*(1), 79–97-79–97.

Breakspear, M., Heitmann, S., & Daffertshofer, A. (2010). Generative models of cortical oscillations: From kuramoto to the nonlinear Fokker–Planck equation. *Front. Hum. Neurosci*, *4*.

Cumin, D., & Unsworth, C. P. (2007). Generalising the Kuramoto model for the study of neuronal synchronisation in the brain. *Physica D: Nonlinear Phenomena*, *226*(2), 181–196-181–196.

Elsaraiti, M., & Merabet, A. (2021). A Comparative Analysis of the ARIMA and LSTM Predictive Models and Their Effectiveness for Predicting Wind Speed [https://doi.org/10.3390/en14206782]. *Energies*, *14*(20), 6782-6782.

Gupta, M., Rajnish, K., & Bhattacharjee, V. (2021). Impact of Parameter Tuning for Optimizing Deep Neural Network Models for Predicting Software Faults [https://doi.org/10.1155/2021/6662932]. *Scientific Programming*, *2021*, 1–17-11–17.

Kuramoto, Y., & Nishikawa, I. (1987). Statistical macrodynamics of large dynamical systems. Case of a phase transition in oscillator communities [https://doi.org/10.1007/bf01009349]. *Journal of Statistical Physics*, *49*(3-4), 569–605-569–605.

Lipton, Z. C. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning. *CoRR*, *abs/1506.00019*. http://arxiv.org/abs/1506.00019

MacKay, D. J. C. (2002). *Information Theory, Inference & Learning Algorithms*. Cambridge University Press.

Muncharaz, J. O. (2020). Comparing classic time series models and the {LSTM} recurrent neural network: An application to S{\&}amp$\mathsemicolon$P 500 stocks [https://doi.org/10.46503/zvbs2781]. *Finance, Markets and Valuation*, *6*(2), 137–148-137–148.

Nakajima, K., & Fischer, I. (2021). *Reservoir Computing* [https://doi.org/10.1007/978-981-13-1687-6]. Springer Singapore. https://link.springer.com/10.1007/978-981-13-1687-6

Palmer, S. E., Marre, O., Berry, M. J., & Bialek, W. (2015a). Predictive information in a sensory population. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(22), 6908-6913. https://doi.org/10.1073/pnas.1506855112

Palmer, S. E., Marre, O., Berry, M. J., & Bialek, W. (2015b). Predictive information in a sensory population [https://doi.org/10.1073/pnas.1506855112]. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 6908-6913.

Pantaleone, J. (2002). Synchronization of metronomes. *American Journal of Physics*, *70*(10), 992–1000-1992–1000.

Rodrigues, F. A., Peron, T. K. D. M., Ji, P., & Kurths, J. (2016). The Kuramoto model in complex networks [https://doi.org/10.1016/j.physrep.2015.10.008]. *Physics Reports*, *610*, 1-98. https://www.sciencedirect.com/science/article/pii/S0370157315004408

Rößler, A. (2010). Runge–Kutta Methods for the Strong Approximation of Solutions of Stochastic Differential Equations [https://doi.org/10.1137/09076636X]. *SIAM Journal on Numerical Analysis*, *48*(3), 922-952. https://doi.org/10.1137/09076636X

Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022). Neural Networks for Financial Time Series Forecasting [https://doi.org/10.3390/e24050657]. *Entropy*, *24*(5). https://www.mdpi.com/1099-4300/24/5/657

Schmidt, R., LaFleur, K. J. R., de Reus, M. A., van den Berg, L. H., & van den Heuvel, M. P. (2015). Kuramoto model simulation of neural hubs and dynamic synchrony in the human cerebral connectome. *BMC neuroscience*, *16*(1), 1–13-11–13.

Schneidman, E., Bialek, W., & Ii, M. J. B. (2003). *Behavioral/Systems/Cognitive Synergy, Redundancy, and Independence in Population Codes*.

Strogatz, S. H. (2018). *Nonlinear Dynamics and Chaos* [https://doi.org/10.1201/9780429492563]. CRC Press.

Tangpanitanon, J., Mangkang, C., Bhadola, P., Minato, Y., Angelakis, D. G., & Chotibut, T. (2022). Explainable natural language processing with matrix product states [https://doi.org/10.1088/1367-2630/ac6232]. *New Journal of Physics*, *24*(5), 053032-053032.

VITA

NAME	Chanin Kumpeerakij
DATE OF BIRTH	22 October 1996
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	B.Sc. Kasetsart University
HOME ADDRESS	39/147 Bang-Mae-Nang, Bang-Yai, Nonthaburi