An investigation into the occurrence of nonlinear attractor in cellular dynamics

He, Chong

2018

He, C. (2018). An investigation into the occurrence of nonlinear attractor in cellular dynamics. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/75939>

<https://doi.org/10.32657/10356/75939>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**AN INVESTIGATION INTO THE
OCCURRENCE OF NONLINEAR ATTRACTOR
IN CELLULAR DYNAMICS**

CHONG HE

School of Physical and Mathematical Science
Nanyang Technological University
2018

AN INVESTIGATION INTO THE OCCURRENCE OF NONLINEAR ATTRACTOR IN CELLULAR

CHONG HE

School of Physical and Mathematical Science
Nanyang Technological University

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2018

Acknowledgments

First and foremost, I would like to devote my utmost gratitude to my supervisor CHEW Lock Yue for his wisdom and continuous encouragement across these five years and half of my doctoral study. Before my first paper got acknowledged by the referees, I was deeply depressed and in constant self-doubt. Many of a time, I couldn't see the value of either my ideas or my resulting work. It was him who always seem to be able to see the meaning and value of my work, and gave me the motivation I desperately need to write up that piece of work and submitted. I could not express how happy and thankful I am, when the referees came back to us with positive view mirroring yours. Later when I got lost and didn't know how to proceed with my research, it was also you who came up with ideas and opportunities, which eventually kicked started my second project. Looking back, I literally can not picture myself completing this journey without his guidance. Thank you for being my supervisor for all these years.

I would also like to take this opportunity to thank my lab mates. Oh, my lab mates. You guys taught me how to make coffee, how to make Matcha, how to enjoy music and spirit and lots of other things in life. As it turned out, all these things I learned from you have now become parts of my life that I can't see myself live without. I shall cherish the time we had discussing about our research together. I shall cherish the time we had enjoying life together.

With equal gratefulness, I shall mention the company that took me in when my scholarship ended, V-Key. I might have gone for work only to get to pay my tuition fee and my living expenses, so that I can finish my ph.D study without troubling my family or others. But what it showed me in return was a new passion I never knew I have.

Last but not least, I would like to express my thankfulness to my girlfriend, JieLin. After decided to leave the academy life, for quite some time, I lost my motivation to carry on with my thesis completion. Eventually, it was the prospect of keeping up with you that motivated me again and pushed me through these final months of my thesis writing. I could never know what you see in me made something like me likeable to you. But the last few months with you was definitely the most filling experience I have ever had thus far.

I struggled for quite some time trying to add my cat Tico in this list. Then again, she has done absolutely nothing for my doctoral journey here in Singapore. That being said, she is a cat. Absolutely nothing is exactly what I would like her to do for me. I love you, Tico. And thank you for staying with my parents while I could not.

Contents

1	Introduction	1
1.1	Research on Biochemical Network	2
1.2	Research on Neuronal Network	7
2	Biochemical Reactions inside a Cell	12
2.1	Irregularity of the reactions and its formulations	13
2.1.1	Gillespie Algorithm	15
2.1.2	Fokker-Planck Equation	18
2.1.3	Chemical Langevin Equation	21
2.2	Cell Size Scaling and Deterministic Dynamics	26
2.3	Biochemical Oscillation Amplitude and Its Common Stochastic Descriptors	29
3	Biochemical Oscillation	30
3.1	Introduction	30
3.2	Some Notes on the Investigation Framework	34
3.3	The amplitude distribution of oscillation models under cell size scaling	35
3.3.1	Case study: Brusselator	35
3.3.2	Discussion on scaling behavior of amplitude distribution in HBDD mechanism with weak noise strength	40
3.3.3	Discussion on scaling behavior of amplitude distribution in NDR mechanism for mild noise strength	48
3.3.4	HBDD and NDR Oscillation in High Dimensional Space	51
3.3.5	Biochemical oscillators in real-world biochemical systems	53
3.4	Simulations of p53 regulatory network and <i>Hes1</i> auto-regulatory network	55
3.4.1	p53 oscillation models	57
3.4.2	Hes1 oscillation models	59
3.4.3	Simulation method and results	59
3.5	Discussion	65
4	Applying to Experimental Data	67
4.1	Observation for P53 oscillation	68
4.2	Observation for hes1 oscillation	72
4.3	Summary	76
5	Habenula Research	77
5.1	Introduction	77

5.2	Material and Method	80
5.2.1	Experimental Method	80
5.2.2	Image Processing	80
5.3	Non-linear time series analysis	84
5.4	Spectral analysis on single cell firing pattern	91
5.5	Conclusion	93
6	Conclusion	94
7	References	95
A	Reaction networks	112
A.1	Brusselator	112
A.2	p53 Oscillatory Network: HBDD network	113
A.3	p53 Oscillatory Network: NDR network	115
A.4	Hes1 Oscillatory Network	117

An investigation into the occurrence of nonlinear attractor in cellular dynamics

by

Chong HE

Summary

High resolution observation of a cellular system of which the components are on the scale of micro-meters or smaller is typically thwarted by several aspects: the difficulty to obtain the signal from the subject itself, the high irregularity of the biological environment, and the measurement noise. Quantitative research over this subject, because of this, could not be easy. However, when we view the system from the perspective of attractors and its properties, it is possible to achieve better understanding of the dynamics of cellular processes even with the observations of limited resolution. In this thesis, we present our research on biochemical network and neuronal network from the perspective of the occurrence of nonlinear attractors in them.

In the chapter on biochemical network, we discuss the different statistical property between two attractors that induces rhythmic behavior while interacting with intrinsic noise: limit cycle and weakly attracting stable spiral. The first type of attractor gives rise to rhythmic behavior from its deterministic dynamics alone, noise only perturbs the already existing oscillation, giving it variation; the second type of attractor gives rise to rhythmic behavior only through its interaction with noise, hence both the oscillation amplitude and its variation are

closely related to the strength of the noise. Because of this difference, the statistics of the rhythmic behavior induced by these two attractors behave differently when system size changes.

Based on our theoretical result, a new method is proposed to determine which of these two mechanisms act behind an observed rhythmic behavior in biochemical system. In particular, we have applied our method in a subsequent chapter to *in vitro* hes1 data. We have uncovered that the observed hes1 rhythmic behavior is driven by Hopf bifurcation induced limit cycle in its deterministic dynamics.

Finally in the chapter on neuronal network, we explored into the possible presence of a chaotic attractor. Our results established that if a chaotic attractor do exist, it cannot possess an embedding dimension of 10 and less. Our analysis instead found that the neuronal network potentially exhibits a critical attractor that arises from the phenomenon of self-organized criticality. Ongoing research is actively being pursued to affirm this latter result definitely.

1 Introduction

It is difficult to observe a cellular system of which the components are on the scale of nano-meters or smaller. Even more so, when such systems are closely interconnected on multiple spatial and time scales, such as a living organism. Limited by the resolution of our observation and the lack of quantification means, much about how cellular processes function are still largely unknown. With the improvement in experimental techniques which lead to more data of higher resolution and the performance of more massive computation, more reliable understanding and more accurate prediction of these cellular processes become possible. In this respect, we investigate into the intrinsic mechanisms of cellular processes to decipher its underlying attractors through a dynamical system approach, based on results of computations and experimental data. In this thesis, we looked into two cellular processes: biochemical network and neuronal network. Our research gained insights into their attractors, and our findings would serve to advance existing investigation in these respective areas. In this section, we shall give a short review on the existing research on both areas.

1.1 Research on Biochemical Network

Biochemical network is an abstract graph describing the interaction network of a whole band of basic cellular components (DNA, RNA, enzymes and other proteins, metabolites, ions, etc.) when they are confined in a cellular environment provided with certain environment. The nodes of the network represent the cellular components. The edges, on the other hand, represent the physical or functional interaction between the nodes, often in the form of chemical reaction. Each node and edge is usually separately identified and measured using experimental methods and then sometimes verified computationally, especially after the maturing of bio-information technology and molecular dynamics. In the research of many cellular processes and functions, biochemical network serves as a good visual representation of the system, helping researchers combing through the constellation of molecules within a cellular environment for the relevant ones for the specific functions to their interest.

In the research on biochemical systems within cells, it has long been a challenge for researchers to collect the necessary data to understand a biochemical system of interests. Protein interactions are known to form a huge and highly interconnected network. It is not uncommon for a protein important to certain functions to have strict regulatory mechanisms to control its expression rate, its localization in the cell, its functionality and even its half-life in the cellular environment. All the controls are achieved by its DNA, mRNA and protein interacting with other proteins and complexes. One example is tumor protein p53, a transcriptional activator crucial for regulating cell cycle arrest, DNA repair and apoptosis in the event of DNA damage.¹

Tumor protein p53 is a transcriptional factor first described in 1979 as an antigen of tumor cells induced by chemicals, irradiation or viruses.² It was only later that researchers uncovered the function of p53 as an important tumor suppressor protein, directing the cell to different non-proliferative cellular responses based on a variety of stress-inducing signals received by the cell.^{1,3} The mutation of p53 is also linked

to a variety of cancers. Up to year 2008, there were over 129 genes in both the human and viral genomes found to be regulated (either activated or repressed, and in some cases situational) by p53.⁴ Most of those regulation involves p53 tetramer's specific binding on DNA response element in target gene's promoter region. In many cases, p53's binding on DNA promoter induces activation of the genes and hence elevated the expression level of the target gene.⁵ In some cases, p53 compete the DNA binding site against a more powerful activator of a target gene, hence suppressing the expression level of it.^{6,7} Tumor protein p53 also down regulates some target gene's expression by interacting with other protein, binding target gene's transcription factor,⁸ for example. Although many p53's regulatory activity is performed in nucleus, mainly as a transcription factor, it is also found that p53 translocates to mitochondria and induces apoptosis in transcription independent manner with Bcl-2, as reviewed by Vaseva, et al.⁹

For such an important regulatory protein, its own regulation mechanism is also sophisticated. In unperturbed cellular environment, p53 is kept at low level due to its short half-life of 5 – 20 minutes.¹⁰ The short half-life of the protein is due to its interaction with Mdm2. Mdm2 is one of the proteins whose expression is activated by Tumor protein p53. In return, Mdm2 ubiquitinates p53 and induces its degradation¹⁰ in the cell, inducing its short half life. Under stress, however, depending on the stress type and level, ATM, ATR, CHK2, etc. could stabilize p53 by phosphorylates the N-terminus residues and reduce its interaction with Mdm2;¹¹ ARF, on the other hand, inhibits Mdm2 directly and thus stabilizing p53.¹² Additionally, upon stabilized in nucleus, p53 promotes the expression of phosphatase Wip1, which, in turn, inhibits both p53 and ATM, by dephosphorylating them, forming a negative feedback loop.¹³

As complicated as the reaction network focusing on p53 is, the dynamics is also sensitive to quantities of reactants. For example, the ubiquitination of p53 by mdm2 can be mono-ubiquitination or poly-ubiquitination based on the abundancy of mdm2 molecules.¹³ While the latter induces the degradation of p53, the former merely export

p53 from nucleus to cytoplasm and could potentially induce apoptosis in certain conditions.^{9,14} Recently, there is another finding showing that the decision of apoptosis by p53 is made based on an ‘effective’ cumulation of p53 over time, effective here refers to the p53 concentration over a certain threshold concentration.¹⁵

For a complicated non-linear reaction network such as p53 regulatory network¹⁶ we discussed above, different reaction rate for many of the reactions involved could change the attractors and hence change the observed dynamics of the system. Obtaining the correct reaction rates is tricky for several reasons. First, obtaining pure samples of the proteins and measure the reaction time *in vitro* is in many cases infeasible. Second, *in vitro* reaction rate, even if measurable, is likely not representative to *in vivo* environment due to the localization of the molecules in the environment, the intracellular transport with cytoskeleton,¹⁷ and the macromolecular crowding effect,¹⁸ etc. As a result, in most of the modelling efforts, the reaction rate, absolute concentrations of different molecules are more inferred than actually measured,^{16,19} usually resulting a ‘sloppy’ model even after careful system calibration.²⁰ After all, number of parameters of a model almost always exceed, by large, the number of observables of a biochemical system. This lack of accuracy makes modeling those system as an effort to gain insight from model’s behavior exceedingly hard even if we could map all significant reactants and their interactions in our models. More often than not, modeling is used as a mere explanatory tool instead of a prediction tool in biochemical research.

Quantitative modeling in more detail and on larger scale is definitely ongoing and yielding better results over the years.^{15,21–24} However, researchers start to take a different perspective when viewing the biochemical systems. Instead of trying to mimic the experimental observation with the computational models and try to gain insight from the resulting time trajectories, they abstract the model and try to understand the properties of the dynamics from the perspective of their attractors. Still using the example of p53 related research, in 2009 and 2010, Ouattara and Abou-Jaoudé summarized the previous modeling efforts thus far from a theoretical level and discussed

the properties of the possible attractors from various different types of modelings.^{25,26} In 2011, Abou-Jaudé further discussed in detail on the properties of the limit cycles that some of the deterministic modelings seems to present.²⁷ In 2012, Choi, et al. revisited the p53 network from the boolean network perspective. They deduced from the attractor landscapes the feedback loops that “control the cellular response to DNA damage”.²⁸ Meanwhile, researchers more conscious about stochastic behavior in cellular environments took another step further and worked on the role noise could have played in biochemical systems.

There are many sources of noise in cellular environment, from molecular fluctuation to random changes within the overall cellular environment. Depending on the systems of interest, the source can be separated into external noise and internal noise. The effect of both has been discussed quite extensively in the biological context.^{29–33} Traditionally, noise is thought to be harmful to the biochemical system since it disrupts the signals which is important for cellular functions. While this is true in many cases, biochemical systems developed strategies to reduce the unwanted noise.²⁹ More often than not, researchers find living organisms adapting noise to their advantages. For example, phage-lambda uses the extrinsic cellular noise to decide if it stays in the lysogenic state or to go lysis and kill its host.^{34–36} In multi-cellular organisms, noise is shown to provide variation among peer cells and served as a driving force for cell differentiation.^{33,37,38} On the other hand, aside of providing variation, noise may have other interesting effects on a biochemical system. Extrinsic noise has been found to be able to alter the landscape, creating bi-stability in state distribution for biochemical system as simple as futile cycle.³⁹ And finally, there has been research showing that intrinsic molecular noise can also been used to model stable oscillation, and proposed for various gene regulatory networks.^{26,40} It means that other than supercritical Hopf bifurcation induced limit cycle, there is an alternative way to recreate the observed ultradian rhythm in biochemical networks.

Though it is indeed more convenient for us to use a more deterministic mechanism,

we should not ignore the possibility that limit cycle, as an explanation of the observed rhythmic behavior, could be triggered by noise driven oscillation. This is especially true in the context of cellular process and gene regulatory system, where copy number of the proteins involved in the process is usually low enough to let the noise induced effect stands out. Which mechanism actually drives the rhythmic phenomena observed in experiments? With the difficulty telling the two mechanisms apart from each other from the time trajectory that we get from the experiment directly, it is natural for us to adopt the analytical perspective and treat the problem using the macroscopic property of the attractor itself. In this thesis, we shall make a discussion over such oscillatory phenomena in biochemical network from a theoretical perspective.

In Chapter 2, we shall describe a framework of methods on analysing the biochemical networks analytically and with the help of simulation. After that, in Chapter 3, we shall discuss in depth on the research and findings we have with regards to distinguishing between the deterministic and the noise-driven mechanisms. An experimentally achievable method of distinguishing the two mechanisms is derived as a result of the discussion. In Chapter 4, we applies the method to data samples from two different experimentally observed biochemical oscillation, finding a promising result indicating that *hes1* oscillation in cellular environment with stable NICD expression maintained by Tet-On system is indeed driven by Hopf bifurcation induced limit cycle in its deterministic dynamics.

1.2 Research on Neuronal Network

Neurons, a type of specialized cell that receives, processes and transmits information through electrical pulses and chemical transmitters, are prevalent among the entire animal kingdom. Typically, multiple neurons connected with each other and form a network that coordinate and function as a whole. The neuron network is responsible for stimuli receiving, decision making and action controlling of the organism, making it one of the most important systems for an animal. The centralized neuronal network of some vertebrates has also been recognized to be able to have memory and problem solving abilities. Understanding how a neuronal network function could help us not only understand how each components of the brain can cooperate together, but guiding us into understanding how units as simple as neurons can interconnect and self-organized into such as to respond quickly and diversely to all the worlds of various stimuli.

A neuron is composed of a cell body (soma), multiple dendrites and one axon. Membrane of a neuron is littered with ion pumps that actively keeps the neuron at a negative resting potentials and ion channels that opens at the influence of certain neuron transmitters or whether cellular potential has surpassed threshold potential. Receptors of neuron transmitters is mostly concentrated on the dendrites, which is responsible for receiving signals. Each stimuli opens an specific ion channel that increase (Na^+ channel) or reduce (Cl^- channel) a bit of the cellular potential of the neuron. Cumulated effects of those stimuli may raise the celular potential over a certain threshold, triggering all the potential gates (Na^+ channels) sensitive to the threshold to open. This would push the local potential all the way higher comparing to the environment. We call such action firing. The firing cascade from soma all the way to the terminal of the neuron's axon, where stored neuron transmitter in synapses is released as signal for dendrites of downstream neurons to receive. This is an basic picture of how neuron processes signal and transmit the signal down to other neurons. In 1952, Alan L. Hodgkin and Andrew F. Huxley co-authored a paper describing the signal processing

process of a single neuron into a set of ordinary differential equations,⁴¹ laying the foundation of subsequent computational modeling effort for neuronal network.

Single neuron firing is relatively known and straightforward. Models of ordinary differential equations with different orders managed to model the neuron firing dynamics and simple neuron interactions *in vitro* with various success. On the other hand, how neurons interact with each other, especially as part of a neuron structure *in vivo* is much less understood. The most well known neuron interaction is synapse through neuron transmitter. However, there are multiple neuron transmitters exists in a neuronal network. And receptors for a same neuron transmitter could lead to the opening of different ion channel and/or leading to changes beyond channel opening in the neurons. The most commonly seen neuron transmitters are glutamate and GABA. Glutamate is commonly seen as an excitatory neuron transmitter, since all its variants of ionotropic receptor (receptor that directly linked to an ion channel) on the dendrites typically opens Na^+/K^+ or Ca^{2+} ion channel and raises the soma potential. However, research showed that in some of the second-order retinal neurons, ON bipolar cells, glutamate serves as an inhibitor because these cells express a metabotropic glutamate receptor (receptor that do not directly link to an ion channel, using message protein for information cascade instead), mGluR6.⁴² This particular receptor is considered inhibitory due to its coupling with the down regulation of cyclic nucleotide synthesis.⁴³ In fact, glutamate has as many as eight variants of metabotropic receptors, among which five of them are considered inhibitive.⁴³ GABA, in comparison, is usually seen as a neuron inhibitor. Since both variants of its receptors, ionotropic GABA_A and metabotropic GABA_B , are considered inhibitive.^{44–46} However, there is exception as well. In immature neuron, it is found that the activation of GABA_A induces membrane depolarization and hence excitation.⁴⁷ Also, GABA_B is known to also present presynaptically. Presynaptic GABA_B , upon activation with synaptic GABA, inhibit the release of a variety of different neuron-transmitters (glutamate and GABA included). Such activity, inhibitory in nature, has more profound significance than regulating ion

channels, because it changes the function of a neuron.⁴⁶ Aside of synapse connection, there is also electrical synapse, where two neighboring neurons forms gap junctions connecting each other.⁴⁸ This type of neuron interaction has a very short response time, allows bidirectional signal transmission, and behave inhibitory for the more depolarized neuron and excitatory for the less depolarized neuron. These properties makes it a much suitable candidate for neuron synchronization comparing to chemical synapses.⁴⁹ In comparison, the modeling efforts of a neuronal network mostly use a same form of interaction term for all neuronal interactions. A single coefficient is to differentiate the strength of the interaction. The sign of that coefficient is used to determine if the interaction is inhibitory or excitatory.^{50–52} Though these modeling efforts each yields interesting results, the lack of consideration over the vastly different types of interaction could blind us from understanding some important aspects of neuronal dynamics.

Research on neuronal network has been on going at two frontiers. Some take the reductionism approach. These researchers start from separating a single neuron from the network and research upon its property *in vitro* and try to build up networks with increasing complexity and functionality. The other route is to observe a neuronal network *in vivo* using various probing techniques such as Electroencephalography (EEG), functional magnetic resonance imaging (fMRI). The prior group of research did gained much insights on neuronal dynamics and how the simplest neuronal network can yield interesting strange attractors.^{53,54} However, neurons *in vivo* operates much different from neurons *in vitro*. For example, for the neurons *in vitro*, they are mostly in rest-state, only subject to occasional controlled stimuli from researchers and other neurons. The neurons *in vivo*, on the other hand, is typically under bombardment of signals, depolarizing them. This makes the neuron *in vivo* having a much shallower threshold before firing comparing neuron *in vitro*.⁵⁵ As a result, the neurons *in vivo* has a higher spontaneous firing rate, and their firings are more likely to be triggered by noise. The other side of the research, in comparison, provides good overview in the networks prop-

erty through its real time population-wise data over the firing statistics in functioning cerebral neuronal networks,⁵⁶ driving the empirical findings of probable neuron coding and crosstalk between cerebral structures in various scenarios. However, due to the lack of individual neuron firing reporter with high enough temporal resolution, for a long time, it has been largely left for mathematicians and computer scientist to answer from theoretical perspective on how connections between neurons could give rise to the properties observed in the firing statistics measured by EEG. A lot has been achieved so far. Yet without concrete single neuron based observation as ground work, it is difficult to confirm what theory on neuronal network might actually describes the specific cases cerebral structure to their interests. EEG data and fMRI data simply couldn't offer an observation with sufficient degree of freedom to represent the properties of a neuronal network. This limitation might be changed soon. Since 2010s, the development of ultra-sensitive protein calcium sensors, GCaMP6f, GCaMP6s, Fast-GCaMPs, etc.⁵⁷⁻⁶⁰ paves the way for single neuron firing observations *in vivo*. GCaMPs are genetically encoded Calcium reporter for neuronal cells. After transfection into embryo, the gene of GCaMPs expressed in neuronal cells. The protein fluorescents when binding Calcium in cytoplasm, making it a candidate to observe neuronal firing intensity through reporting the relative concentration of Calcium as a result of firing associated influx. In the past, the response time of GCaMPs could not provide the necessary resolution for observation in Neuronal firing due to the slow rising and decaying time for the protein's fluorescence. Even for GCaMP6, the fluorescent decaying half life is around 100ms. In latest development, however, researchers managed to reduce the reporter signal's rise and decay times to 1ms and 3ms, respectively,⁶⁰ close to the time a neuron needs to recover from a firing. This could very possibly makes it suitable for capturing even the peak firing rate of a neuronal network, enabling its use for the observation for *in vivo* single neuron firing patterns.

Rest-state brain neurons are known to fire spontaneously.^{61,62} In early days, due to the lack of apparent pattern, such spontaneous has mostly been attributed to noise.⁶³

However, there are also research showing that for cerebral neuronal networks, the balanced input could lead to chaotic dynamics in neuron firing dynamics.^{64,65} On the other hand, in recent research, it has also been shown theoretically that neuron network with balanced input with multiple different topology appears to form self-organized critical pattern such as power law behavior in avalanches.⁵² In Chapter 5, we shall present our exploration for the potentially existing attractor in habenula spontaneous activity based on rest-state habenula's calcium fluorescent data we obtained using GCaMP6f transfected zebrafish. Our goal it to find evidences if chaos attractor or critical attractor is indeed driving the spontaneous activities in habenula.

2 Biochemical Reactions inside a Cell

In this section, we shall introduce the framework for analyzing the biochemical reaction network in living cells. As introduced in Chapter. 1, a biochemical network is composed of a bunch of reactants (protein, lipid, RNA, DNA), interacting with each other in the form of biochemical reactions. These chemical reactions, albeit not different in nature from those we conduct in a test tube or other reaction chambers, take place in a very small reaction chamber. Some of those reactions take place in a cell, which is only thousands of μm^3 in size. Others might be confined in even smaller chambers like nucleus and other sub-cellular structures. Such confinement limits the number of reactants in the chamber, hence the irregularity of the reaction, though minuscule in larger scale systems, can be very important in the cellular process.

2.1 Irregularity of the reactions and its formulations

To answer what causes the irregularity of the reactions, we need to know what chemical reactions are. Though vary greatly among them, a vast majority of the reactions in the cellular process can be summarized into one or a combination of several steps of the following types of elementary reaction:⁶⁶

1. one particle changed into another type of particle or breakdown into several different particles ($A \rightarrow \text{products}$). Its reaction rate is determined by the concentration of this particle, $R = \lambda \frac{n_A}{V}$, where λ is a coefficient representing the reaction rate, n_A is the particle number of reactant A. V is the chamber size.
2. two (or rarely more) particles collide and react into something different ($A+B \rightarrow \text{products}$). Its reaction rate is determined by the concentration of both particles, $R = \lambda \frac{n_A n_B}{V^2}$, where λ is a coefficient representing the reaction rate, n_A and n_B are the particle number of reactant A and reactant B. V is the chamber size.

For example, for the process of the myosin walking on an actin filament⁶⁷ (see Fig. 1):

1. ATP binds to the myosin sub-domain A, sub-domain A bound with ATP dissociates from actin-filament (process 2);
2. ATP hydrolysis provides the energy for sub-domain A to swing forward (process 1);
3. phosphate releases from the myosin sub-domain, sub-domain A rebinds the actin filament (process 1);
4. ADP disassociates from the myosin sub-domain B, the myosin morphs into the state ready for the next step (1).

For single particle elementary reaction, there are generally two types of changes: existing chemical bond breaking; and new chemical bond forming. For the former case,

Fig 1. Myosin walking on actin filament, adapted from the work of Okten et al.⁶⁷

the reaction takes place when the electrons migrate to a higher energy state through thermal input. For the latter case, the reaction happens when two sub-modules getting physically close though thermal vibration. For two or more particle reaction case, the reaction happens when the particles collide with each other in thermal motion and form a chemical bond. For all the cases above, thermal activity is the basis of those reactions. If reaction happen independently between each other, and that the reaction chamber is homogeneous, then each instances of a same type of reaction takes place with a same probability that is only dictated by the thermal activity intensity and the chemical composition in the chamber. The count of arrivals for independent incidents like these reaction instances can be expressed by the Poisson distribution. Let the reaction rate being R instances per unit of time, and the distribution for reaction

instances (N) in time interval t can be expressed as

$$P_{N=n} = \frac{(Rt)^n}{n!} e^{-Rt}. \quad (2.1)$$

There are however subtleties that go beyond the above assumptions:

1. While Poisson process can be used to describe the arrival counts in certain time interval if the system is homogeneous, each instance of the reaction, when it happens, will surely change the chemical composition locally and making the system less homogeneous than before.
2. And even when we give an assumption that the redistribution of chemical is infinitely fast, the resulting chemical composition after the reaction is still different from before the reaction happening.

While partial differential equation has been used to address the spatial in-homogeneity caused by reactions in some simulation efforts of smaller scale systems,⁶⁸ most of the simulations and analytical efforts are formulated under one or more approximations that addresses the subtleties discussed above. In the following subsections, we shall discuss three of those formulations, i.e. Gillespie Algorithm, Fokker-Planck Equations and Chemical Langevin Equations.

2.1.1 Gillespie Algorithm

Since the cellular chamber's size is only hundreds to thousands of μm^3 , it is reasonable to assume that the in-homogeneity caused by instances of reaction dissipates in the reaction chamber fast enough to make it possible to use a simple reaction rate R instances per unit time to predict the timing of the next reaction. Under this assumption, we may derive the distribution of the arrival time of the next reaction instance. This distribution is actually the basis of the Gillespie Algorithm,⁶⁹ which is the most accurate simulation algorithm without considering the spatial in-homogeneity and is

heavily used in our stochastic simulations. The Gillespie Algorithm is a classic algorithm for stochastic simulation on chemical reactions. This algorithm is Monte Carlo based iterative procedure in time. Each iteration of this algorithm takes four steps to complete. For a system composed of r reactions, repeat the following process in each iteration of the simulation:

1. calculate the reaction rate (R_i) of each considered reaction type i in the system based on the current reactants number in the reaction chamber and parameters of the reaction rates. Calculate the total reaction rate $R_\Sigma = \sum_{i=1}^r R_i$. This summation is valid because of the independence of each reaction.
2. use a random number from uniform distribution (from 0 to 1) to decide the timing of the next reaction (of any reaction type) based on the next reaction arrival time's cumulative distribution function (CDF) for reaction rate R_Σ (we shall derive later).
3. use another uniformly distributed random number (from 0 to 1) to decide the specific type of reaction the next reaction is, based on the relative reaction rate among the considered reactions. A reaction of rate R_i would have a probability of $\frac{R_i}{R_\Sigma}$ to be chosen.
4. update the reactant counts for each of the reactants affected by one instance of the reaction taking place. Update the time by the amount decided by step. 2.

Other than Step. 2, other steps are quite straightforward. Now we shall derive the CDF with the total reaction rate R_Σ . For this, let's also temporarily coarse-grain the passage of time into a series of time intervals Δt . After the last instance of reaction, we note the total reaction rate for the current chemical composition to be R_Σ . Then the possibility of one reaction to happen within the time interval $n\Delta t$ and no reaction happening before it can be described as

$$p(n\Delta t) = (1 - R_\Sigma \Delta t)^{(n-1)} (R_\Sigma \Delta t) .$$

In the limit of $\Delta t \rightarrow 0$, denoting this time interval δt , the probability density of the next reaction arrival time at time $t_n = n\delta t$ is

$$\begin{aligned}\lim_{\delta t \rightarrow 0} f(t_n) \delta t &= \lim_{\delta t \rightarrow 0} (1 - R_\Sigma \delta t)^{\frac{(n-1)\delta t}{\delta t}} (R_\Sigma \delta t) \\ &= \lim_{\delta t \rightarrow 0} ((1 - R_\Sigma \delta t)^{1/\delta t})^{(n-1)\delta t} (R_\Sigma \delta t) \\ &= \lim_{\delta t \rightarrow 0, n \gg 0} e^{-R_\Sigma t_n} (R_\Sigma \delta t)\end{aligned}$$

With this, we shall also derive the probability of no reaction taking place within the time period t (CDF of the reaction arrival time distribution),

$$P(t) = 1 - \int_0^t f(\tau) d\tau = e^{-R_\Sigma t} \quad (2.2)$$

Indeed, Eq. (2.2) coincide with the Poisson distribution Eq. (2.1) as expected when $n = 0$.

Making use of the cumulative distribution function above, we can generate a uniform distribution of random number x between 0 and 1, and determine the arrival time of the next reaction by

$$t = -\frac{\ln x}{R_\Sigma}. \quad (2.3)$$

The idea behind Gillespie Algorithm is simple. It updates the system's reactant number by one reaction instance at a time. In each iteration of the algorithm, the algorithm first find out how long before next reaction happens, then it figures out which reaction it is. After deciding the reaction and its happening time, it updates the amount of each involved reactants and time accordingly.

We should note here that the simulation using Gillespie algorithm progresses at one reaction instance at a time. And for this single iteration, two random numbers needs to be generated. Due to these details, the algorithm is considerably slow. Nonetheless, since the scale and the complexity of the bio-chemical network in our consideration is relatively small, in Chapter 3, we mainly use Gillespie Algorithm for stochastic simulations on the biochemical networks.

2.1.2 Fokker-Planck Equation

Gillespie Algorithm, though powerful as a simulation tool providing reliable result with good accuracy, is not very helpful for analytical purposes. To effectively analyze a stochastic system, one ought to do the analysis from the perspective of the ensemble, while Gillespie Algorithm takes the perspective of a single trajectory. Unfortunately, to analyze the dynamics from the ensemble perspective efficiently, we will have to make a few more assumptions. In this section, we shall introduce one of the tools that can help us understand the dynamics from the ensemble perspective — the Fokker-Planck equation, since Master Equation based formulation is the most natural language in describing dynamics in the state space. In the next section, we shall introduce its equivalent, but more intuitive representation of the dynamics, Chemical Langevin Equation.

Since both Langevin Equations and Fokker-Planck Equations are defined in the ensemble of the reaction system, we need to formally define the structure of the formulation before going further in the discussion.

For a reaction system of n reactants and r reactions, the state space is defined as the n dimensional spaces, with each dimension representing the number of molecules of one reactant. At any specific time point, the system can be represented as a point in the state space, $\mathbf{X} = (X_1, X_2, \dots, X_n)$. The reactions determine how the systems migrate from one state to another in the phase space.

Since reactions are independent from each other and the reaction rate is mostly determined from the current state of the system, we regard these reactions as Markovian process. Taking reaction j for example, each instance of it taking place changes the amount of some reactants, we shall denote such changes by a vector from the original state space $\omega_j = (\omega_{j1}, \omega_{j2}, \dots, \omega_{jn})$. Recalling what we discussed at the beginning of this chapter, at each system state, due to the differences in the reactant count, the reaction rate of each reaction would also be different. We reuse the earlier notation

and denote the migration rate from one system state to another caused by reaction j as $R_j(\mathbf{X})$. With each reaction's reaction rate and state changing effects defined, a keen reader might already noticed that we are ready to define the dynamics of this biochemical reaction network with the Master Equations. Let the probability of the occupancy of system state $\mathbf{X}$ at moment t be $P_{\mathbf{X}}(t)$, we may express its dynamics under the chemical reaction set of r reactions as

$$\frac{d}{dt}P(\mathbf{X}, t) = \sum_{j=1}^r (P(\mathbf{X} - \omega_j, t)R_j(\mathbf{X} - \omega_j) - P(\mathbf{X}, t)R_j(\mathbf{X})) \quad (2.4)$$

Indeed, Eq. (2.4) is the equivalent of Gillespie Algorithm from the perspective of ensemble. However, to analyze the dynamics of an n dimensional state space, on the other hand, is no trivial matter, after all, the size of the state space is to the magnitude of $\sim \prod_i^n \nu_i$, where ν_i is the number of possible values for the molecule number of reactant i . Even for a simple reaction space of 5 reactants, the molecule count for each reactant are bounded within 0 to 999 molecules, the size of the state space would be 10^{15} , a size unwieldy for any kind of serious analysis.

To be able to make meaningful analysis over the state space, we need to remove the necessity of treating each system state as a standalone state, but a point in a continuous space, whose dynamics is governed by a differential equation. This is what we do in deriving the Fokker-Planck Equation from the original Master equation.

The first step we need to do in deriving a Fokker-Planck equation, is to find the equivalent form of Eq. (2.4) in continuous state space. We write a generic Master Equation,

$$\frac{\partial}{\partial t}P(\mathbf{X}, t) = \int P(\mathbf{X} - \mathbf{l}, t)W(\mathbf{l}, \mathbf{X} - \mathbf{l})d\mathbf{l}_1d\mathbf{l}_2 \cdots d\mathbf{l}_n - \int P(\mathbf{X}, t)W(-\mathbf{l}, \mathbf{X})d\mathbf{l}_1d\mathbf{l}_2 \cdots d\mathbf{l}_n, \quad (2.5)$$

where $W(-\mathbf{l}, \mathbf{X})$ is the transition rate from $\mathbf{X}$ to state $\mathbf{X} - \mathbf{l}$. As you shall see, it is not difficult to find the proper form of W for Eq. (2.4), with the help of Dirac Delta

function:

$$W(\mathbf{l}, \mathbf{X}) = \sum_{j=1}^r (\delta(\mathbf{l} - \omega_j) R_j(\mathbf{X})) \quad (2.6)$$

Note that for each instance of reaction, especially for elementary reactions, the change of molecule count for each reactant is usually very small, especially in comparison with the total number of the molecules of that reactant. Thus, $W(\mathbf{l}, \mathbf{X})$ is infinitesimally small for large $|\mathbf{l}|$ without any regards to the value of $\mathbf{X}$. This allow us to Taylor expand $(\mathbf{X} - \mathbf{l})$. In practice, we usually expand it to the second order, obtaining

$$\begin{aligned} \frac{\partial}{\partial t} P(\mathbf{X}, t) = & -P(\mathbf{X}, t) \int W(-\mathbf{l}, \mathbf{X}) dl_1 dl_2 \cdots dl_n + P(\mathbf{X}, t) \int W(\mathbf{l}, \mathbf{X}) dl_1 dl_2 \cdots dl_n \\ & - \int \sum_{i=1}^n \left(l_i \frac{\partial}{\partial X_i} (P(\mathbf{X}, t) W(\mathbf{l}, \mathbf{X})) \right) dl_1 dl_2 \cdots dl_n \\ & + \frac{1}{2} \int \sum_{i=1}^n \sum_{k=1}^n \left(l_i l_k \frac{\partial^2}{\partial X_i \partial X_k} (P(\mathbf{X}, t) W(\mathbf{l}, \mathbf{X})) \right) dl_1 dl_2 \cdots dl_n. \end{aligned}$$

The first two terms of this expansion cancel each other out. Let

$$a_i(\mathbf{X}) = \int l_i W(\mathbf{l}, \mathbf{X}) dl_1 dl_2 \cdots dl_n \quad (2.7)$$

$$b_{ik}(\mathbf{X}) = \int l_i l_k W(\mathbf{l}, \mathbf{X}) dl_1 dl_2 \cdots dl_n \quad (2.8)$$

Then the Taylor expanded Master equation takes the form

$$\frac{\partial}{\partial t} P(\mathbf{X}, t) = - \sum_{i=1}^n \frac{\partial}{\partial X_i} (P(\mathbf{X}, t) a_i(\mathbf{X})) + \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2}{\partial X_i \partial X_k} (P(\mathbf{X}, t) b_{ik}(\mathbf{X})). \quad (2.9)$$

Equation (2.9) is called Fokker Plank equation. From this equation, you could see that the dynamics of the distribution of the state space is determined by a_i and b_{ik} , which in turn, is determined by Eq. (2.6). Applying Eq. (2.6) to Eqs. (2.7) and (2.8), we shall have the explicit form of a_i and b_{ik} , and hence forth the Fokker-Planck equation for the bio-chemical reaction network:

$$a_i(\mathbf{X}) = \sum_{j=1}^r \omega_{ji} R_j(\mathbf{X}) \quad (2.10)$$

$$b_{ik}(\mathbf{X}) = \sum_{j=1}^r \omega_{ji} \omega_{jk} R_j(\mathbf{X}) \quad (2.11)$$

In contrary to the Gillespie Algorithm, Fokker-Planck Equation do not give you the intuitive view on how a system state transit into another. Instead, it gives a full picture of what happens in the ensemble, with its well studied form of Partial Differential Equation. Using this formulation to study the properties of a stable distribution comes quite handy, especially when its features are not immediately determined by the local flux of the system, and hence not being able to be revealed by trajectory/flux based localized analysis. In Section 3.3.2 and Section 3.3.3, we shall use the Fokker-Planck equation for analyzing the state distribution around the fixed point.

2.1.3 Chemical Langevin Equation

Fokker-Planck Equation can provide good insight into the dynamics of the state distribution in the state space. However, in many cases, it is too abstract to visualize. After all, people familiar with chemical reaction network are more used to think of the process from trajectory’s perspective instead of the perspective of the state space. It would be much more intuitive if we have a trajectory based formulation for us to describe the biochemical network analytically. Fortunately, there is an equivalent form of Fokker-Planck equation that very much resembles the familiar form of Ordinary Differential Equation, which describe the system from the perspective of trajectories. This formulation is called the Chemical Langevin Equation.

In this section, we shall use the same set notation as earlier, a reaction network composed of n reactants $\mathbf{X}$, and r reactions with reaction rates expressed as $\mathbf{R}(\mathbf{X}) = (R_1(\mathbf{X}), \dots, R_r(\mathbf{X}))$. Each instances of reaction j changes the molecule number of each reactants by vector $\omega_j = (\omega_{j1}, \omega_{j2}, \dots, \omega_{jn})$. As discussed in earlier sections, by neglecting the diffusion effect, the arrivals of the instances for reaction R_j can be described as a Poisson process, $Po(R_j\mathbf{X}(t)\tau)$, provided that during the time period τ in consideration, the chemical composition do not change significantly enough to change the reaction rate. Employing this idea, the change of $\mathbf{X}$ over a short period of time $\tau \rightarrow 0$ can be approximately expressed by the following equation set (i from 1 to

n):

$$\lim_{\tau \rightarrow 0} X_i(t + \tau) = \lim_{\tau \rightarrow 0} \left(X_i(t) + \sum_{j=1}^r \omega_{ji} Po(R_j(\mathbf{X}(t))\tau) \right). \quad (2.12)$$

Before proceeding, I should mention here that there are simulation algorithms designed based on Eq. (2.12). For system with larger size, simulation based on this method could be significantly faster than the one based on Gillespie Algorithm. However, since biochemical reaction is usually non-linear and quite stiff, in different region of the state space, the reaction rate could be significantly different, it is not trivial to find a proper τ for the entire simulation. There are active research on balancing the simulation accuracy and efficiency, like Quasi-Steady-State approximation,^{70,71} however, such discussion is out of the scope for this thesis.

In analysis, we typically ignore the higher order moment of the Poisson process. Approximating the actual dynamics with a normal distribution,

$$\begin{aligned} Po(R_j(\mathbf{X}(t))\tau) &\approx \mathcal{N}\left(R_j(\mathbf{X}(t))\tau, \sqrt{R_j(\mathbf{X}(t))\tau}\right) \\ &= R_j(\mathbf{X}(t))\tau + \sqrt{R_j(\mathbf{X}(t))\tau} \mathcal{N}_{t,j}(0, 1). \end{aligned}$$

Note that although the distribution of every reaction at any moments of time can all be described by an expression with normal distribution, $\mathcal{N}(0, 1)$, the normal distributions for one reaction at one time is not necessarily related to that for the other reaction at another time. Hence in the last line of the equation above, I use subscript t, j to denote the reaction instance distribution of reaction j at time point t to signify its stochasticity.

If the time interval is infinitesimally small, $\tau \rightarrow 0$, the equation set describing the chemical reaction system's dynamics can be written as:

$$\lim_{\tau \rightarrow 0} (X_i(t + \tau) - X_i(t)) = \lim_{\tau \rightarrow 0} \sum_{j=1}^r \left(\omega_{ji} R_j(\mathbf{X}(t))\tau + \omega_{ji} \sqrt{R_j(\mathbf{X}(t))\tau} \mathcal{N}_{t,j}(0, 1) \right).$$

Denote symbol B_j as:

$$dB_j(t) \equiv \lim_{\tau \rightarrow 0} \sqrt{\tau} \mathcal{N}_{t,j}(0, 1), \quad (2.13)$$

the dynamics can then be written into a differential equation form:

$$dX_i(t) = \sum_{j=1}^r \left(\omega_{ji} R_j(\mathbf{X}(t)) dt + \omega_{ji} \sqrt{R_j(\mathbf{X}(t))} dB_j(t) \right). \quad (2.14)$$

This is the Chemical Langevin Equation.⁷² This equation set gives the dynamics of observable $\mathbf{X}$ (particle count of each of the reactants in the reaction chamber), with its deterministic term, $\omega_{ji} R_j(\mathbf{X}(t)) dt$, and its stochastic term $\omega_{ji} \sqrt{R_j(\mathbf{X}(t))} dB_j(t)$. Comparing to the Fokker-Planck equation, this form is more familiar in the literature. Ignoring the second term, the equation set has the form of Ordinary Differential Equations that biochemical network researchers are familiar with. And the equation may be interpreted as an ODE governed dynamics with a stochastic term disrupting the system.

Note that while deriving the Langevin Equation, we have made the same set of assumptions as we did in deriving the Fokker-Planck Equation. It is natural to see that these two forms are equivalent. In fact, we can re-obtain $a_i(\mathbf{X})$ and $b_{ik}(\mathbf{X})$ in Eq. (2.10) and Eq. (2.11) using the Chemical Langevin Equation. To do this, however, we need to have a better understanding of its noise term, $dB_j(t)$.

First, from Eq. (2.13), we know that the first moment (average) of $dB_j(t)$ satisfies,

$$\langle dB_j(t) \rangle = \lim_{\tau \rightarrow 0} \sqrt{\tau} \langle \mathcal{N}_{t,j}(0, 1) \rangle = 0. \quad (2.15)$$

From the discussion at the beginning of this Chapter, we know that the elementary chemical reactions are only determined by the current chemical composition of the reaction chamber. Hence the observable $dB_j(t)$, which represents the occurrence of mutually independent chemical reactions in a unit period of time, is a Markovian process. We also know that each elementary reactions are independent of each other. Taking the arguments earlier into considerations we have

$$\langle dB_j(t) dB_k(t') \rangle = \langle \mathcal{N}_{t,j}(0, 1) \mathcal{N}_{t',k}(0, 1) \rangle \lim_{\tau \rightarrow 0} \tau = \lim_{\tau \rightarrow 0} \delta_{j,k} \delta(t - t') dt dt'. \quad (2.16)$$

Note that $\delta(a - b)$ is Dirac Delta function, while $\delta_{a,b}$ is Kronecker Delta function.

Now let's look at the form of $a_i(\mathbf{X})$ and $b_{ik}(\mathbf{X})$. By slightly changing the form of their definition, Eq. (2.7) and (2.8), we obtain

$$\begin{aligned} a_i(\mathbf{X}) &= \int l_i W(\mathbf{l}, \mathbf{X}) dl_1 dl_2 \cdots dl_n \\ &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle X_i(t + \tau) - X_i(t) \rangle \\ &= \lim_{\tau \rightarrow 0} \frac{\left\langle \int_t^{t+\tau} dX_i \right\rangle}{\tau} \end{aligned}$$

$$\begin{aligned} b_{ik}(\mathbf{X}) &= \int l_i l_k W(\mathbf{l}, \mathbf{X}) dl_1 dl_2 \cdots dl_n \\ &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle (X_i(t + \tau) - X_i(t))(X_k(t + \tau) - X_k(t)) \rangle \\ &= \lim_{\tau \rightarrow 0} \frac{\left\langle \int_t^{t+\tau} \int_t^{t+\tau} dX_i(t) dX_k(t) \right\rangle}{\tau}. \end{aligned}$$

It should be noted here that the averaging here is on the transition $\mathbf{l}$ instead of the state $\mathbf{X}$.

Let us first look at the form of $\frac{\langle dX \rangle}{dt}$. Combining the right-hand-side with Eq. (2.14),

$$\lim_{\tau \rightarrow 0} \frac{\left\langle \int_t^{t+\tau} dX_i \right\rangle}{\tau} = \sum_{j=1}^r \left(\omega_{ji} R_j(\mathbf{X}(t)) + \lim_{\tau \rightarrow 0} \omega_{ji} \frac{\sqrt{R_j(\mathbf{X}(t))} \int_t^{t+\tau} \langle dB_j(t) \rangle}{\tau} \right).$$

From Eq. (2.15), we know that the first moment (average) of $dB_j(t)$ is 0, hence from the equation above, we re-obtained the Eq. (2.10),

$$\lim_{\tau \rightarrow 0} \frac{\left\langle \int_t^{t+\tau} dX_i \right\rangle}{\tau} = \sum_{j=1}^r (\omega_{ji} R_j(\mathbf{X}(t))) = a_i(\mathbf{X}).$$

Similarly, we can re-obtain Eq.(2.11) using the Chemical Langevin Equation:

$$\begin{aligned}
& \lim_{\tau \rightarrow 0} \frac{\left\langle \int_t^{t+\tau} \int_t^{t+\tau} dX_i(t) dX_k(t') \right\rangle}{\tau} \\
&= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left\langle \int_t^{t+\tau} \int_t^{t+\tau} \sum_{j=1}^r \left(\omega_{ji} R_j(\mathbf{X}(t)) dt + \omega_{ji} \sqrt{R_j(\mathbf{X}(t))} dB_j(t) \right) \times \right. \\
&\quad \left. \sum_{l=1}^r \left(\omega_{lk} R_l(\mathbf{X}(t')) dt' + \omega_{lk} \sqrt{R_l(\mathbf{X}(t'))} dB_l(t') \right) \right\rangle \\
&= \sum_{j=1}^r \sum_{l=1}^r \omega_{ji} R_j(\mathbf{X}(t)) \omega_{lk} R_l(\mathbf{X}(t')) \lim_{\tau \rightarrow 0} \tau + \\
&\quad \lim_{\tau \rightarrow 0} \sum_{j=1}^r \sum_{l=1}^r \frac{1}{\tau} \int_t^{t+\tau} \int_t^{t+\tau} \omega_{ji} \omega_{lk} \sqrt{R_j(\mathbf{X}(t))} \sqrt{R_l(\mathbf{X}(t'))} \langle dB_j(t) dB_l(t') \rangle \\
&= \lim_{\tau \rightarrow 0} \sum_{j=1}^r \sum_{l=1}^r \frac{1}{\tau} \int_t^{t+\tau} \int_t^{t+\tau} \omega_{ji} \omega_{lk} \sqrt{R_j(\mathbf{X}(t))} \sqrt{R_l(\mathbf{X}(t'))} \delta_{j,l} \delta(t-t') dt dt' \\
&= \omega_{ji} \omega_{lk} R_j(\mathbf{X}(t)) = b_{ik}.
\end{aligned}$$

Langevin Equation is frequently used in moment analysis,⁷³ an analytical procedure to obtain the dynamics of the observables' average, standard deviation, and higher order moments by deriving the analytical form of the average of $\lim_{\tau \rightarrow 0} (\mathbf{X}(t + \tau) - \mathbf{X}(t))^n$, making use of Eq. (2.14) and properties of $dB_j(t)$ described in Eq. (2.15) and Eq. (2.16). For linear dynamics, this method is very useful, since the dynamics of the observable's higher-moment can always be expressed by the linear combination of its current value and the observable's lower moment's current value. However, for non-linear dynamics, in oppose to the linear case, the dynamic of the observable's lower moments usually depends on the current value of the higher moments, making it impossible to obtain a closed form of the dynamics without further approximation. This is why when dealing with non-linear dynamics like Chemical Langevin Equation, we usually need to make further approximations. In our discussion, we usually use local linearization to help obtain useful insights to the system, as seen in discussion in Section 3.3.2 and 3.3.3. More generic closure scheme for non-linear dynamics can be found in the work by Smadbeck, et al.,^{74,75} however, such discussion is out of the scope of this thesis.

2.2 Cell Size Scaling and Deterministic Dynamics

Since it is the limitation of reaction chamber size and the molecule count of the reactants that make the stochastic phenomena important in the discussion of biochemical networks, we shall formally formulate the effect of chamber size in the reaction network’s stochastic behavior in this section.

In macroscopic reaction system, the systems of different sizes are considered equivalent when each of the reaction components have the same concentration. This is because the subsequent trajectory of the reaction systems with the same concentration in every reactant are similar to each other. In the discussion of microscopic reaction system, since we are concerned with the effect of chamber size on the property of system’s stochastic behavior, when we discuss on “equivalent” systems of different chamber size, we are comparing the systems of different size but same concentration of each reactants.

Definition 1. *Cell size scaling operation: an operation performed on the reaction system that changes the reaction chamber size, V , without affecting the current reactant concentration or changing the number of instances of each reaction in a unit time and a unit volume.*

Symbolically, using the notation introduced in Chemical Langevin Equation, Eq. (2.14), we define the cell size scaling (or spatial scaling) operation that scales the size of the reaction chamber by factor η_c to be a scaling operation $\hat{\eta}_c$ being applied to a biochemical system that satisfies

$$V_c \equiv \eta_c V \tag{2.17}$$

$$\mathbf{X}_c \equiv \hat{\eta}_c \mathbf{X} = \eta_c \mathbf{X} \tag{2.18}$$

$$\mathbf{R}_c(\mathbf{X}_c) \equiv \hat{\eta}_c \mathbf{R}(\mathbf{X}) = \eta_c \mathbf{R}(\mathbf{X}). \tag{2.19}$$

Applying the scaling operation above to the Langevin Equation, Eq. (2.14), we

obtain the scaled dynamics

$$\begin{aligned}
dX_{ci}(t) &= \sum_{j=1}^r \left(\omega_{ji} R_{cj}(\mathbf{X}_c(t)) dt + \omega_{ji} \sqrt{R_{cj}(\mathbf{X}_c(t))} dB_j(t) \right) \\
\iff \eta_c dX_i(t) &= \sum_{j=1}^r \left(\eta_c \omega_{ji} R_j(\mathbf{X}(t)) dt + \omega_{ji} \sqrt{\eta_c R_j(\mathbf{X}(t))} dB_j(t) \right) \\
\iff dX_i(t) &= \sum_{j=1}^r \left(\omega_{ji} R_j(\mathbf{X}(t)) dt + \eta_c^{-\frac{1}{2}} \omega_{ji} \sqrt{R_j(\mathbf{X}(t))} dB_j(t) \right).
\end{aligned}$$

As expected, the scaling of the system under Definition. 1 do not change the deterministic part of the dynamics. The stochastic part of the dynamics, on the other hand, scales by a factor of $\eta_c^{-\frac{1}{2}}$. The ratio of relative noise strength between a cell size scaled system and the original system is

$$\xi \equiv \frac{\frac{\sum_j \sqrt{\omega_{cij}(\vec{x}_c)}}{\sum_j \omega_{cij}(\vec{x}_c)}}{\frac{\sum_j \sqrt{\omega_{ij}(\vec{x})}}{\sum_j \omega_{ij}(\vec{x})}} = \eta_c^{-1/2}. \quad (2.20)$$

For the same set of reactions and reactant concentrations, the smaller the reaction chamber size, i.e., $\eta_c \ll 1$, the larger the relative noise strength. On the other hand, in the limit of $\eta_c \rightarrow \infty$, the noise term become negligible, and the dynamics formulated by the Chemical Langevin Equation set regress into a set of ordinary differential equations that is symmetric to the cell size scaling operation,

$$dX_i(t) = \sum_{j=1}^r \omega_{ji} R_j(\mathbf{X}(t)) dt. \quad (2.21)$$

Here, the word “symmetric” is used to describe the property of the dynamics that the formulation of Eq. (2.21) preserves itself under the cell size scaling operation.

We should also mention here that such regression from Chemical Langevin Equation to Ordinary Differential Equation(ODE), Eq. (2.21), gives a mathematical explanation on how the stochastic nature of the chemical reactions become inconsequential in macroscopic reaction chambers, yet play a big role in some of the microscopic cellular context.

Although the deterministic dynamics only works for systems of large chamber size,

due to the maturity of efficient ODE based simulation method, Eq. (2.21) are frequently used for simulations on the biochemical networks with larger scale. Although the discussion on different simulation algorithms is out of the scope of this thesis, there is one thing with regards to the simulation I would like to discuss.

Similar to how non-linearity adds complication to the discussion of the stochastic dynamics of the system, it is also not trivial to simulate Eq. (2.21) due to the non-linearity of the term $R_j(\mathbf{X}(t))$. Since R_j is a non-linear combination of X_i , its error due to the finite time step is also non-linear.

For example, for $R_k \equiv aX_iX_j$, if the error of X_i , ΔX_i , are limited under a ratio $\frac{|\Delta X_i|}{|X_i|} < \epsilon$, where ϵ is a small positive number, the error of R_k is bounded by

$$\frac{|\Delta R_k|}{|R_k|} < \frac{\epsilon}{|X_i|} + \frac{\epsilon}{|X_j|}.$$

As observed in the equation above, the bound for the error in the reaction rate R_k is not a constant. Its variation depends on the value of X_i and X_j . In the regime where the value of X_i and X_j is larger, the error in R_k is smaller, when X_i and X_j is smaller, the error of the reaction rate increases. Such heterogeneity in the reaction rate's error in different part of the phase space eventually leads to an artificial systematic difference in the error of X_i in different part of the phase space, making the simulation extremely prone to diverge due to the cumulated simulation error, no matter how small the time step is chosen. We refer to such tendency as the stiffness of the non-linear dynamics. Since the stiffness is caused by the heterogeneity of the error when the simulation's time step is fixed, researchers typically uses "adaptive time step" based method to even the error out across the phase space. All the deterministic simulations performed in this thesis are conducted using adaptive stepsize 6th order Runge-Kutta algorithm, as described in Numerical Recipes in C.⁷⁶

2.3 Biochemical Oscillation Amplitude and Its Common Stochastic Descriptors

As we are mostly interested in the oscillatory behavior observed in the stochastic biochemical network, before closing this Chapter, I would like to introduce a few descriptors for a noisy oscillation.

To describe an oscillating biochemical system observed in the laboratory, it is necessary to construct measures to quantify the observed oscillatory behavior. By observing the number of molecules x_i of a reactant that is periodically increasing and decreasing in copy number, we can obtain the oscillation amplitude χ of such reactant.

To obtain the amplitude χ experimentally, we need to first segment the time series into a set of oscillatory periods. The one we used for the segregation marker is the first time x_i penetrate $\langle x_i \rangle + p\sigma_{x_i}$ from below after it has penetrated $\langle x_i \rangle - p\sigma_{x_i}$ from above. p is a parameter around 0.25 that can be slightly tuned based on the data we received. In our data analysis, we mostly used 0.15 for this parameter. There are many other ways to segment the time series. The guideline for choosing a specific segmentation scheme is that it can segregate the oscillating time series into a set of segments, each containing only one peak and one trough.

With the periods obtained, we may obtain the χ for each period by finding the difference between each period's peak with $\langle x_i \rangle$. Theoretically, we can also try to use each period's minimum with *average* x_i , however, experimentally, trough have a smaller readout hence less distinguishable from observation noise, it is easier and more reliable to use the peaks to represent the amplitude due to its usually sharper feature.

Two statistical features of such oscillation amplitudes are extensively discussed in our investigations: the normalized average amplitude for reactant i , $\widehat{\langle \chi_i \rangle}$, which is defined as

$$\widehat{\langle \chi_i \rangle} \equiv \frac{\langle \chi_i \rangle}{\langle x_i \rangle}; \quad (2.22)$$

and the coefficient of variation of amplitude CV_{χ_i} , which is defined as

$$CV_{\chi_i} \equiv \frac{\sigma_{\chi_i}}{\langle \chi_i \rangle}. \quad (2.23)$$

The normalized average amplitude $\widehat{\langle \chi_i \rangle}$ signifies how discernible the spikes are from the resting average copy number of molecules i , i.e., $\langle x_i \rangle$. On the other hand, the coefficient of variation CV_{χ_i} indicates how regular the amplitudes are with the presence of cellular noise. For biochemical processes where oscillations represent signals to downstream processes, the variation σ_{χ_i} in the pulsing amplitude quantifies the level of noise in the signal. Thus, a larger CV_{χ_i} means a smaller signal-to-noise ratio for the noisy signal received by the downstream processes.

Theoretically, we can obtain the stochastic measure of the oscillation period in a similar fashion. However, to our knowledge, the segregation point of the oscillation period is usually taken as the first time the trajectory of the concentration of an observed reactant piercing through a certain threshold in a specific manner^{22,77} (as you shall see in examples in Chapter. 3), which is arbitrary, to say the least. Hence in our discussion, we place less importance to the statistics of the oscillatory period as oppose to the observed oscillatory amplitude.

3 Biochemical Oscillation

3.1 Introduction

Many biological systems exhibit oscillatory behavior. At the molecular level, many biochemical networks are found to be oscillatory. For example, during embryonic development, the cellular expression level of the *Hes1* gene oscillates steadily. These *Hes1* oscillations regulate cell differentiation and tissue formation.^{78–80} The expression levels of many other genes and proteins have also been found to undergo oscillations during embryonic development. Some of them serve as internal clocks, while others regulate cell differentiation. Oscillatory behaviors are also commonly observed beyond

embryonic development. The most notable is the cell cycle, which consists of periodic divisions and duplications of cells under strict and sophisticated regulation in each of its four phases.⁸¹ Oscillatory dynamics are also observed in the stress response networks. Under ionizing radiation, p53, the protein responsible for cell fate decision upon activation by diverse stress signals, oscillates steadily with a period of 5.5h.⁸² NF- κ B, another gene regulator important in stress signal response, is also reported to show oscillatory behavior after activation, although its oscillation only lasts for a few periods.⁸³

Over the years, many models have been developed to explain these oscillatory behaviors in biochemical networks. Among the models that do not give rise to oscillations by employing explicit time delays, most employ the idea of a supercritical Hopf bifurcation in a dynamical system described by ordinary differential equations (ODEs).^{22,26,68,84–86} On the other hand, for gene regulatory network such as the p53 network which responds to stress, and the *Hes1* auto-regulatory network for embryonic development, the use of noise-driven resonance has been explored, owing to the low copy numbers of the involved reactants.^{23,26,40,87} Since the two mechanisms — one based on supercritical **H**opf **B**ifurcation in **D**eterministic **D**ynamics (HBDD), and the other based on **N**oise **D**riven **R**esonance (NDR) — are both capable of recovering the experimentally observed oscillatory behavior, we want to ask the question of whether these two mechanisms are equivalent in the biochemical context, and how to tell the difference between them if they are not equivalent.

We analyzed two statistics of noisy oscillations in our investigations: the coefficient of variation in oscillation amplitude (CV_χ , which is defined in Eq. (2.23)), and the normalized average oscillation amplitude (the average amplitude of the oscillations over the mean copy number of the observed oscillating reactant, $\langle \widehat{\chi} \rangle$, which is defined in Eq. (2.22)). We focus on CV_χ for two reasons. It gives the statistics of noisy oscillations that is directly observable from experiments,⁸² as well as information on how “regular” the oscillation amplitude is. Oscillations with a higher CV_χ is found

to possess a less regular oscillatory pattern than one with a lower CV_χ . On the other hand, while CV_χ informs us on the regularity of the oscillations, $\widehat{\langle \chi \rangle}$ indicates the discernibility of the oscillations. An oscillation in a biochemical network describes the periodic rising and falling of the copy number of one reactant or a group of reactants. In many occasions, the amplitude of the oscillations is smaller than the mean copy number of the oscillating reactant. And when the amplitude is much smaller than the mean copy number, i.e., $\widehat{\langle \chi \rangle} \ll 1$, the oscillations becomes negligible. Thus, by using the normalized amplitude, one may determine whether or not the oscillations are discernible or significant within its biochemical network. It is worth mentioning that while the coefficient of variation is experimentally observable through detecting fluorescence of tagged proteins, the same cannot be said about the normalized amplitude. Due to the unpredictable transfection rates for genes encoding fluorescence tagged proteins and background noise, fluorescence data are usually measured in arbitrary unit with arbitrary zero levels, making cell-wise comparison of normalized amplitudes from fluorescence data meaningless.

Cells of different types vary greatly in their sizes. For example, human alveolar macrophages have a typical size of $4,990 \mu\text{m}^3$,⁸⁸ while that of lymphocytes is around $210 \mu\text{m}^3$. Also, cell size is not necessarily invariant even among the same cell type. For example, a cell that is undergoing the cell cycle has its size immediately halved upon division. And before the next division, it gradually grows back to roughly its original size. Hence its size ranges, at least, from half to its original size. Because cell size variation is ubiquitous throughout the lifetime of an organism, for biochemical networks to function stably, its core functional features must be robust against cell size variations. In our work, we find that for an oscillating biochemical network of which intrinsic noise is dominant, its oscillating reactants' amplitude distribution preserves different properties under cell size scaling, depending on which mechanism underlies its oscillation. NDR driven oscillations maintain a constant CV_χ for cells of different sizes, suggesting that oscillations driven by the NDR model are able to keep the oscillation

patterns consistent for cells of various sizes. In contrast, HBDD driven oscillations hold the normalized average amplitude constant, suggesting that the oscillations in the HBDD mechanism are equally discernible for cells of different sizes.

In this Chapter, we first use the example of Brusselator give an intuitive introduction to the concept of Hopf bifurcation, limit cycle(For formal definitions, kindly refer to Steven H. Strogatz's book *Nonlinear Dynamics and Chaos*⁸⁹), noise driven resonance and how limit cycle and noise driven resonance can both drive a reliably observable oscillatory phenomena.

Next, we show how their amplitude distributions behave when cell size changes. More specifically, we will show that if the oscillations are driven by the HBDD mechanism, we can expect the normalized average amplitude to be invariant upon cell size scaling operation, while for NDR driven oscillations, the coefficient of variation will be invariant under such operation.

Then, we shall mathematically demonstrate why the statistics of oscillations through the HBDD and NDR model behave in such a way under cell size scaling. While biochemical networks are multidimensional, their dynamics still display either of the scaling behaviors stated above based on which mechanism is driving it. The mathematical explanation is again backed up by simulations conducted on HBDD and NDR driven p53 regulatory network and *Hes1* auto-regulatory network and observed the predicted scaling behaviors.

Our investigation is entirely based on the dynamics of biochemical networks with intrinsic noise, without taking into account the effects of extrinsic noise. Although extrinsic noise is important in many biochemical processes, their behavior in the context of cell size scaling is as yet undefined both theoretically and experimentally, hence rendering its investigation generally infeasible. Therefore, we shall focus on the effect of intrinsic noise on the mechanisms behind the oscillation phenomena, in the same spirit as exemplified by many other papers.^{22,23,90}

3.2 Some Notes on the Investigation Framework

The Investigation framework used in this Chapter shall be based on what has been discussed in Chapter 2. It should be noted that although the majority of the discussion in the earlier chapter was based on the assumption that the biochemical network of interest is composed of elementary reactions, this specific condition is not required. Gillespie Algorithm only requires the reaction system to be Markovian and the diffusion is fast enough that in-homogeneity do not affect the local reaction rates. Chemical Langevin Equation and its equivalent Fokker-Planck equation added one more approximation that the number of reactions in an infinitesimally short time interval can be approximated by a Gaussian distribution. For any biochemical reaction system, as long as they fit the criteria mentioned above, the mathematical framework given in Chapter 2 can be successfully applied. It is important for me to point this out, because although all reactions can be broken down into its composing elementary reactions, such breakdown is usually too clumsy and unnecessary for researchers to analyze even a moderately complicated reaction network. Instead, we usually create models with composite reactions only. Composite reactions usually merged multiple elementary reactions steps into one, hence greatly reduces the size of the models. As long as we can ensure that each of the reactions are still independent from each other, the framework developed in Chapter 2 still applies.

To simplify the problem, our investigation is performed on biochemical networks isolated from the environment. No extrinsic noise is considered in our systems. The reaction rates of our system are constants, with only the intrinsic Poisson noise from the reactions in the network giving rise to variation in the reactants concentration. Thus, any inherent differences resulting from the two different mechanisms under study are to be attributed to the underlying Poissonian noise as well as the dynamics of the oscillation models.

3.3 The amplitude distribution of oscillation models under cell size scaling

3.3.1 Case study: Brusselator

The simplest model of a chemical system that can oscillate under both HBDD and NDR mechanisms is the Brusselator. The Brusselator is defined in Eq. (3.1) as follows:

In the deterministic limit, the ODE set for this system is

$$\begin{aligned}
 \frac{d}{dt}x &= A + x^2y - x - Bx, \\
 \frac{d}{dt}y &= -x^2y + Bx.
 \end{aligned}$$

Depending on the value of B and A , the dynamics of the Brusselator can be different. The fixed point of this system is $\mathbf{P} = (A, \frac{B}{A})$. The stability of the fixed point, on the other hand, depends on the relation between B and A . If $B < A^2 + 1$, P is a stable spiral (see Fig. 2a). Trajectories from various initial condition spirals in until they get infinitely close to P . When B is increased to the point that $B = A^2 + 1$, the fixed point is no longer an attractor. Instead, there is another attractor formed surrounding its neighbourhood.

Although it is very difficult to notice this new attractor at $B = A^2 + 1$, with B increased further, $\mathbf{P}$ becomes an unstable spiral, and the newly formed attractor gets further and further away from $\mathbf{P}$. We can then see that this newly formed attractor is circular. Except for the case where the trajectory starts at point $\mathbf{P}$, in which case the trajectory stays stationary, trajectory of all other initial conditions are attracted towards the circular attractor (see Fig. 2b). Formally, we call the circular attractor “stable limit cycle”. And the process where increasing B turning the stable spiral into

(c) Attractors for $A = 1$, $B \in [1, 3]$.

Fig 2. Hopf bifurcation in the Brusselator. All results are obtained by simulations with $A = 1$. (a) displays the trajectories on the phase space when simulated with $B = 1$. The fixed point $\mathbf{P}$ of this set of simulation is a stable spiral, and all trajectories are drawn towards the fixed point. (b) displays the trajectories on the phase space when simulated with $B = 3$. For this set of simulation, the fixed point $\mathbf{P}$ is an unstable spiral. Additionally, a stable limit cycle surrounding $\mathbf{P}$ is formed and attracting almost all trajectories towards it. (a) displays the attractors during the bifurcation process. The diagram is obtained by intersecting the phase space of each B with the line $x = 1$. When $B < 2$, the attractor is $\mathbf{P}$ ($y = \frac{1}{B}$). Yet at $B = 2$, $\mathbf{P}$ is becoming unstable and a new stable limit cycle is forming around it. The process is called the supercritical Hopf bifurcation. In the regime $B > 2$, $\mathbf{P}$ is now unstable hence no longer an attractor, and the stable limit cycle is becoming the new attractor of the system with its radius increases with the increasing B .

a pair of unstable spiral and limit cycle the Hopf Bifurcation⁸⁹ (see Fig. 2c).

When the system described by Eq. (3.1) is oscillating in the limit cycle regime, its dynamics can reliably produce an oscillations (Fig. 3a). Since the limit cycle that produces the oscillatory pattern arises from the supercritical Hopf bifurcation at the deterministic limit, we refer to it as oscillations driven by the HBDD mechanism.

Interestingly, although there is no limit cycle in the regime $B < A^2 + 1$, if the

Fig 3. The trajectory in phase space (left) and time series (right) of noisy oscillations. Data is obtained from simulating Brusselator with different parameter set using Gillespie Algorithm. Parameter set of a and c are provided in Table 2. For b and d the system parameters are scaled as described in Sec. 2.2

system size is small and B approaches the bifurcation point $A^2 + 1$ from the stable spiral regime, we can still observe the system's time series to behave like an oscillator (Fig. 3c). This oscillation is induced by the Poissonian intrinsic noise of the reaction network, and maintained by the interaction between stable spiral dynamics that draws the trajectory back to the fixed point, and the noise that pumps the system away from it. Since such oscillation-like behavior is driven by noise, we refer to it as oscillations driven by the NDR mechanism.

Numerical simulations of these two Brusselator models (using Gillespie Algorithm,

	$\langle\chi\rangle$	$\widehat{\langle\chi\rangle}$	σ_χ	CV_χ
HBDD	1	0	0.5	-0.5
NDR	0.5	-0.5	0.5	0

Table 1. Scaling exponent for amplitude χ of the HBDD and NDR model

see Section. 2.1.1 for detail and Supplemental Material published at Physical Review E⁹¹ for corresponding cell size scaling script, simulation code and data analysis script) showed that HBDD and NDR driven oscillations behave differently upon cell size scaling. For the HBDD mechanism, the CV of oscillation amplitude differs between oscillations in cells of different sizes. On the other hand, its normalized average oscillation amplitude appears to be scale invariant, as shown in Fig. 3a and b. For the NDR mechanism, the scaling behavior is the opposite. Its normalized average oscillation amplitude is smaller for oscillation at the scale of larger cell size, while its CV is scale invariant (Fig. 3c and d). A more detailed comparison between the scaling exponent from the statistics of the two mechanisms is given in Fig. 4. As you shall see, in both models, The mean oscillation amplitude and the CV of the oscillation amplitude in cells of different sizes generally follow the exponents described in Table. 1, until the scale of the system become so small that the Fano Number (the grey bars in Fig. 4) of the oscillatory system approaches 1.

In the research regarding noisy oscillators, the Fano number is frequently used as an indicator of how regular and distinguishable an observed rhythmic behavior is compared to purely random Poisson process. It is defined as

$$F = \frac{\sigma_W^2}{\mu_W}, \quad (3.2)$$

where μ_W is the average number of observed oscillatory periods in a certain time window W , and σ_W^2 is the variance of it. In Fig. 4, the time window used to calculate the Fano number is 10 times the average observed period from the simulation. Typically, the larger the noise, the closer Fano number approaches to 1. When Fano number is 1, the observed trajectory is indistinguishable from a Poisson process (purely stochastic). Since the scaling exponent for both $\langle\chi\rangle$ and CV_χ fits the value in Table. 1 quite well

Fig 4. Statistics (normalized average of amplitude and CV of amplitude) of oscillation dynamics against cell size scaling. Stars are simulation result for Brusselator described by Eq. (3.1). Recorded reactant is y . Dashed lines are linear fit for log-log dataset with scale coefficient larger than 2^1 . The fitted slope in each plots are (a) 0.001, (b) -0.482 , (c) -0.466 and (d) -0.001 , closely following Table. 1. In fact, these are the typical scaling exponents for the two oscillation mechanisms. The reason behind such scaling behavior will be explained in Subsection 3.3.2 and Subsection 3.3.3. For both the HBDD and NDR mechanisms, the scaling behavior deviates from the linear fitting when scaling coefficient is small ($\eta_c < 1$). Meanwhile, we also noticed that for data points that significantly deviate from the typical scaling exponent of the mechanism, their corresponding Fano number (grey bars) for oscillation period (measured with spike recognition threshold of $0.25\sigma_y$ and time window of 10 mean period) is also significantly higher. Especially for the NDR mechanism, within typical scaling exponent, the Fano number is stabilized at around 2.5, while for data points with scaling exponent that deviates from it, their corresponding Fano numbers are close to 1. This indicates that, at these small scales, the system's behavior is closer to Poisson random walk than noise driven resonance.

until the Fano number of the observed oscillation rises to close to 1 for the specific model, we can see that the scaling exponents described in Table. 1 holds as long as the oscillation is distinguishable from the system. We should also notice from the table that the scaling exponents for HBDD mechanisms induced oscillation is strikingly

different from the NDR mechanism induced oscillation.

Specifically, if we can experimentally measure the CV_X in cells of different sizes and obtaining its scaling behavior, we might well distinguish these two oscillations generated from different mechanisms on the Brusselator.

What's more, such differences in scaling behavior between the two mechanisms is not unique to the Brusselator. In the following subsections, we shall derive the scaling behavior of the statistics of the oscillation through the HBDD and NDR mechanisms in generic biochemical networks, and observe their underlying differences. For simplicity, we shall confine our discussion to the $2D$ phase space first. Although most biochemical systems are multi-dimensional, oscillations can usually be described within a $2D$ manifold, which allows the results of the following derivations to apply.

3.3.2 Discussion on scaling behavior of amplitude distribution in HBDD mechanism with weak noise strength

Fig 5. Illustration for Morphing the Phase Space

For a system oscillating through the HBDD mechanism without noise, the oscillation amplitude should be completely determined by the shape of the limit cycle. Since deterministic dynamics (excluding the Langevin term) of a biochemical system is scale invariant (see Section. 2.2), its normalized amplitude is also scale invariant.

However, due to the Poisson noise of the reactions, the trajectory of the full SDE is constantly being pushed away from the limit cycle, displaying a varying amplitude (Fig. 3a). To quantify such variations, we need to understand how amplitude is measured first. Experimentally, the time series of the observed reactant is first divided

into a series of time frames, each frame containing approximately one period of the oscillation. Then, within each time frame, amplitudes can be identified by determining the local maximums. The respective values are obtained by computing the difference between the expression levels of the reactant at its local maximum with its average expression level. For a 2-reactant limit cycle, as shown on the left of Fig. 5, the local maximum measured for reactant a is at point A on the limit cycle, assuming the system has no noise. Its trajectory is perpendicular to axis a , and its amplitude is measured roughly as the distance between A and O projected on the a axis, χ_0 . However, if there is noise, the system shall reach its maximum in the vicinity of point A . If we construct a surface Σ_χ centered on A and orthogonal to the trajectory at the neighbourhood of A , the procedure is analogous to yielding the intersection in phase space between trajectories of the system and a surface of section Σ_χ (Fig. 5 Left). Note that we can also select reactant b and analyse its amplitude by a similar procedure. Nonetheless, we shall restrict our discussion to the maximum amplitude of reactant a

From Sec. 2.1.3, more specifically, Eq. (2.14), we know that for a biochemical reaction, the following formulation applies:

$$d\mathbf{X} = \mathbf{v}dt + \mathbf{L}, \quad (3.3)$$

where $\mathbf{v}$ is a vector defined as

$$v_i(\mathbf{X}) \equiv \sum_{j=1}^R \omega_{ji} R_j(\mathbf{X}(t)), \quad (3.4)$$

and $\mathbf{L}$ is the Langevin term,

$$L_i(\mathbf{X}) \equiv \sum_{j=1}^R \omega_{ji} \sqrt{R_j(\mathbf{X}(t))} dB_j(t). \quad (3.5)$$

After scaling with coefficient η_c , its dynamics becomes

$$d\mathbf{X}_c = \eta_c \mathbf{v}dt + \eta_c^{1/2} \mathbf{L}. \quad (3.6)$$

We expect the dynamics to be stable and the trajectory centered on the limit cycle. If the noise is moderate, the trajectory shall stay in the neighbourhood of the limit cycle (Fig. 5 Left). Next, we express the system in terms of coordinates centered on

the limit cycle (Fig. 5 Right). In this new representation, the dynamics of the system can be expressed as

$$\begin{aligned} dX_{c1} &= \eta_c v_1 \left(\frac{\mathbf{x}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_1 dB_1, \\ dX_{c2} &= \eta_c v_2 \left(\frac{\mathbf{x}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_2 dB_2, \end{aligned} \quad (3.7)$$

where the variables with subscripts 1 and 2 are vector components corresponding to $\hat{e}_1$ and $\hat{e}_2$ respectively. Note that $\mathbf{X}_c$ is defined in the domain

$$\begin{aligned} X_{c1} &\in [0, L_c], \\ X_{c2} &\in (-\infty, \infty), \end{aligned}$$

where L_c is the length of the limit cycle of the scaled system. If the length of the limit cycle in the original scale is L , then L_c satisfies

$$L_c = \eta_c L. \quad (3.8)$$

It should be highlighted here that the dB_1 and dB_2 in Eq. (3.7) are unit vectors on the direction of $\hat{e}_1$ and $\hat{e}_2$. They are actually the renormalized projection of the sum of all the reaction's dB_j on their respective direction, hence for this case, $\langle dB_1 dB_2 \rangle$ is not necessarily a Delta function.

In this representation, the limit cycle is the line $X_{c2} = 0$. Without noise, the system trajectory converges to the limit cycle. However, the noise given by the Langevin terms constantly pushes the system away from the limit cycle. When the two “forces” balance each other in the ensemble, a distribution of system states is formed as a belt surrounding the limit cycle. It is worth noting that although X_{c2} is not specifically bounded, the probability of a trajectory reaching regions with large X_{c2} is very unlikely if the noise is moderate.

For the convenience of further discussion, let us define surface Σ_χ as a straight line orthogonal to $\hat{e}_1$. From the self-similar property of the limit cycle over cell size operations, the surface's coordinate on $\hat{e}_1$ on scaled system is $X_{c\Sigma_\chi} = \eta_c X_{\Sigma_\chi} = \eta_c X_A$.

The amplitude distribution is represented as the distribution of flux on surface Σ_χ :

$$P(\chi_{c0} + X_{c2}) \propto \Phi_{\Sigma_\chi}(X_{c2}) , \quad (3.9)$$

where χ_{c0} is the amplitude measure of the limit cycle at the local maximum in the scaled system at A_c for reactant a . Since the limit cycle is deterministic, it is easy to show that

$$\chi_{c0} = \eta_c \chi_0 .$$

Now we estimate the scaling behavior of the amplitude distribution by calculating the analytical form of the flux $\Phi_{\Sigma_\chi}(X_{c2})$ and its scaling behavior. Note that surface Σ_χ is orthogonal to $\hat{e}_1$. If state space probability distribution $P(\mathbf{X}_c)$ is known, the flux on Σ_χ at $(X_{c1} = X_{c\Sigma_\chi}, X_{c2})$ can be simply expressed as

$$\Phi_{\Sigma_\chi}(X_{c2}) = \left(v_{c1}(\mathbf{X}_c)P(\mathbf{X}_c) - \frac{1}{2}\Gamma_1^2 \frac{\partial}{\partial X_{c1}} P(\mathbf{X}_c) - \frac{1}{2}\Gamma_1\Gamma_2 \langle dB_1 dB_2 \rangle \frac{\partial}{\partial X_{c2}} P(\mathbf{X}_c) \right) \Big|_{\mathbf{X}_c=(X_{c\Sigma_\chi}, X_{c2})} . \quad (3.10)$$

Thus we need to obtain an analytical form of the state space distribution $P(\mathbf{X}_c)$ to evaluate Eq. (3.10). We shall obtain $P(\mathbf{X}_c)$ by solving Eq. (3.7) via the method of variable separation. This can be done under the following assumptions:

1. noise is moderate and constant in a reaction network;
2. noise along the limit-cycle and orthogonal to it is independent;
3. the speed of the trajectory in phase space along the limit cycle is independent of the drift orthogonal to the limit cycle ($m_{c1} = 0$, with m_{c1} defined in Eq. (3.22));
4. retraction factor m_{c2} (it shall be properly defined and explained later with Eq. (3.22)) is a constant in the reaction network.

Our first step in solving the equations:

$$\begin{aligned} dX_{c1} &= \eta_c v_1 \left(\frac{\mathbf{X}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_1 dB_1, \\ dX_{c2} &= \eta_c v_2 \left(\frac{\mathbf{x}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_2 dB_2, \end{aligned}$$

is to separate the noise in direction $\hat{e}_1$ (along the limit cycle) and $\hat{e}_2$ (orthogonal to the limit cycle), and to reduce both noise terms to unit noise in order to simplify our later derivation. Before we do these, however, we shall simplify the form of the equations by letting y_{ci} and u_i be

$$\begin{aligned} y_{ci} &\equiv X_{ci} \eta_c^{-1/2} \Gamma_i^{-1}, \\ u_i &\equiv v_i \eta_c^{1/2} \Gamma_i^{-1}. \end{aligned} \tag{3.11}$$

With these, the two stochastic differential equations can be reduced to

$$\begin{aligned} dy_{c1} &= u_1 \left(\frac{\Gamma}{\eta_c^{1/2}} \mathbf{y}_c \right) dt + dB_1, \\ dy_{c2} &= u_2 \left(\frac{\Gamma}{\eta_c^{1/2}} \mathbf{y}_c \right) dt + dB_2. \end{aligned} \tag{3.12}$$

Let us make the following two assumptions,

1. Γ_i in Eq. (3.7) is approximated as a constant;
2. dB_1 and dB_2 are independent, i.e. $\langle dB_i dB_j \rangle = \delta(i - j)$.

With these, we can rewrite the equation set into their corresponding Fokker-Planck equation at the steady state,

$$\begin{aligned} 0 &= \mathcal{L}P(\mathbf{y}_c), \\ \mathcal{L} &= -\frac{\partial}{\partial y_{ci}} u_i + \frac{1}{2} \Delta, \end{aligned} \tag{3.13}$$

with its boundary conditions:

$$P(\mathbf{y}_c) \big|_{y_{c1}=0} = P(\mathbf{y}_c) \big|_{y_{c1}=\eta_c^{-1/2} \Gamma_1 L_c}, \tag{3.14}$$

$$P(\mathbf{y}_c) \big|_{y_{c1}=-\infty} = P(\mathbf{y}_c) \big|_{y_{c1}=\infty} = 0. \tag{3.15}$$

The flux through Σ_χ as determined by Eq. (3.12) under these conditions can be written as

$$\Phi_{\Sigma_\chi}(y_{c2}) = \frac{\Gamma_1}{\eta_c^{1/2}} \left(u_1 P(\mathbf{y}_c) - \frac{1}{2} \frac{\partial}{\partial y_{c1}} P(\mathbf{y}_c) \right) \bigg|_{\mathbf{y}_c = (y_{c\Sigma_\chi}, y_{c2})}. \tag{3.16}$$

To be able to solve Eq. (3.13), we use the linear approximation by assuming that noise is moderate and the trajectory stays within the neighbourhood of the limit cycle.

This leads us to expand $\mathbf{u}$ with respect of y_{c2} up to first-order as follows:

$$u_1(\mathbf{y}_c = \eta_c^{1/2} \Gamma^{-1} \mathbf{X}) \approx u_1 \left(\frac{\Gamma}{\eta_c^{1/2}} y_{c1}, 0 \right) + \frac{\partial}{\partial X_1} v_1 \Big|_{\mathbf{y}_c = (\eta_c^{1/2} \Gamma^{-1} X_1, 0)} y_{c2} \quad (3.17)$$

$$= u(y_{c1}) + m_{c1}(y_{c1}) y_{c2}, \quad (3.18)$$

$$u_2(\mathbf{y}_c = \eta_c^{1/2} \Gamma^{-1} \mathbf{X}) \approx \frac{\partial}{\partial X_1} v_2 \Big|_{\mathbf{y}_c = (\eta_c^{1/2} \Gamma^{-1} X_1, 0)} y_{c2} \quad (3.19)$$

$$= m_{c2}(y_{c1}) y_{c2}, \quad (3.20)$$

where

$$u(y_{c1}) \equiv u_1 \left(\frac{\Gamma}{\eta_c^{1/2}} y_{c1}, 0 \right), \quad (3.21)$$

$$m_{ci}(y_{c1}) \equiv \frac{\partial v_i}{\partial x_1} \Big|_{\mathbf{y}_c = (\eta_c^{1/2} \Gamma^{-1} X_1, 0)}. \quad (3.22)$$

We further assume that the drift orthogonal to the limit cycle ($y_{c2} \neq 0$) does not affect the speed of the trajectory along the direction of limit cycle $u_1(\mathbf{y}_c)$, i.e. $m_{c1}(y_{c1}) \approx 0$.

With this assumption, the Fokker-Planck equation at steady state takes the form

$$0 = -\frac{\partial}{\partial y_{c1}} u(y_{c1}) P(\mathbf{y}_c) - \frac{\partial}{\partial y_{c2}} m_{c2}(y_{c1}) y_{c2} P(\mathbf{y}_c) + \frac{1}{2} \Delta P(\mathbf{y}_c). \quad (3.23)$$

We can solve this final form of the dynamics (Eq. (3.23)) via variable separation:

$$P(\mathbf{y}_c) = P_{\hat{e}_1}(y_{c1}) P_{\hat{e}_2}(y_{c2}). \quad (3.24)$$

Integrating Eq. (3.24) on y_{c1} and applying boundary conditions, we have

$$\begin{aligned} 0 &= \int_0^{\eta_c^{-1/2} \Gamma_1 L_c} \left(-\frac{\partial}{\partial y_{c1}} u(y_{c1}) P_{\hat{e}_1}(y_{c1}) P_{\hat{e}_2}(y_{c2}) \right. \\ &\quad \left. - \frac{\partial}{\partial y_{c2}} m_{c2}(y_{c1}) y_{c2} P_{\hat{e}_1}(y_{c1}) P_{\hat{e}_2}(y_{c2}) + \frac{1}{2} \Delta P_{\hat{e}_1}(y_{c1}) P_{\hat{e}_2}(y_{c2}) \right) dy_{c1} \\ \iff 0 &= \frac{\partial^2}{\partial y_{c2}^2} P_{\hat{e}_2}(y_{c2}) - 2 \left(\int_0^{\eta_c^{-1/2} \Gamma_1 L_c} m_{c2}(y_{c1}) P_{\hat{e}_1}(y_{c1}) dy_{c1} \right) y_{c2} P_{\hat{e}_2}(y_{c2}). \end{aligned}$$

The solution of $P_{\hat{e}_2}(y_{c2})$ is given by

$$P_{\hat{e}_2}(y_{c2}) = \sqrt{-M/\pi} \exp(M y_{c2}^2), \quad (3.25)$$

where

$$M = \int_0^{\eta_c^{-1/2} \Gamma_1 L_c} m_{c2}(y_{c1}) P_{\hat{e}_1}(y_{c1}) dy_{c1}$$

is a constant.

Putting the solution Eq. (3.25) back into Eq. (3.24), we have

$$(-m_{c2}(y_{c1}) + M) M P_{\hat{e}_1}(y_{c1}) y_{c2}^2 + \frac{\partial}{\partial y_{c1}} P_{\hat{e}_1}(y_{c1}) - \frac{\partial}{\partial y_{c1}} (u(y_{c1}) P_{\hat{e}_1}(y_{c1})) + (M - m_{c2}(y_{c1})) P_{\hat{e}_1}(y_{c1}) = 0. \quad (3.26)$$

Since $u(y_{c1})$ and $P_{\hat{e}_1}(y_{c1})$ are both independent of solution for $P_{\hat{e}_1}(y_{c1})$, coefficient of y_{c2}^2 term, $M(M - m_{c2}(y_{c1}))$ must equal to 0 for all y_{c1} in order for Eq. (3.26) to yield a non-trivial solution for $P_{\hat{e}_1}(y_{c1})$. In addition, it is also necessary that $M \neq 0$ for non-trivial solution of $P_{\hat{e}_2}(y_{c2})$ according to Eq. (3.25). Hence, m_{c2} needs to be a constant. Let this constant be m_0 , then the condition required for non-trivial solution for $P_{\hat{e}_1}(y_{c1})$ becomes

$$m_{c2}(y_{c1}) = \text{Const.} = m_0. \quad (3.27)$$

Effectively, the term $m_{c2}(y_{c1})y_{c2}$ in Eq. (3.20) causes the trajectory that is drifting away due to noise to converge back towards the limit cycle. Hence we call m_{c2} the retraction factor. The condition given by Eq. (3.27) means that to separate variable successfully, this retraction term is required to be independent of y_{c1} , i.e. m_{c2} needs to be phase invariant. In addition, it is worth noting that from Eq. (3.22), we know that m_0 has to be scale invariant with respect to system size as well.

The solution for Eq. (3.20) based on the assumption Eq. (3.27) is given by

$$P_{\hat{e}_2}(y_{c2}) = \sqrt{-m_0/\pi} \exp(m_0 y_{c2}^2),$$

$$u(y_{c1}) P_{\hat{e}_1}(y_{c1}) - \frac{1}{2} \frac{\partial}{\partial y_{c1}} P_{\hat{e}_1}(y_{c1}) = \text{Const.}$$

Inserting these results back into Eq. (3.16), we have

$$\begin{aligned}\Phi_{\Sigma_\chi}(y_{c2}) &\propto \exp(m_0 y_{c2}^2) \\ \iff \Phi_{\Sigma_\chi}(x_{c2}) &\propto \exp\left(\frac{m_0}{\eta_c \Gamma_2^2} x_{c2}^2\right).\end{aligned}$$

Finally, putting this equation into Eq. (3.9), we obtain the amplitude distribution for HBDD mechanism:

$$P(\chi_c = \chi_{c0} + x_{c2}) = C \exp\left(\frac{m_0}{\eta_c \Gamma_2^2} x_{c2}^2\right), \quad (3.28)$$

where C is a normalization constant.

Then from Eq. (3.28), we see that

$$\langle \chi_c \rangle = \chi_{c0} = \eta_c \chi_0 \propto \eta_c, \quad (3.29)$$

$$CV_{\chi_c} = \frac{\Gamma_2}{\sqrt{m_0 \eta_c \chi_0}} \propto \eta_c^{-1/2}, \quad (3.30)$$

in accordance with our numerical results obtained earlier in Sec. 3.3.1.

Eq. (3.29) and (3.30) shows how the normalized average amplitude, $\langle \widehat{\chi_i} \rangle$, and the oscillation's coefficient of variation, CV_{χ_i} , changes in response to the cell size scaling operation. These results mirror what we have summarized in Table. 1 exactly. Although the assumptions we made appear to be too strict for an actual biochemical network, they allow us to obtain the analytical form of Eq. (3.28) by means of separation of variables. In real-life systems, the dynamics along $\hat{e}_1$ (along limit cycle) and $\hat{e}_2$ (orthogonal to limit cycle) are usually correlated, making it difficult to find precise analytical solutions. We argue that the scaling behavior of the amplitude of HBDD driven oscillation still follows Eqs (3.29) and (3.30). This solution tells us that the CV of the distribution given by Eq. (3.30) is determined by how strong the retraction term is, in comparison to the strength of noise. The retraction term is deterministic with scaling exponent 1, while the strength of noise has scaling exponent of 0.5. Hence their relative strength changes when the system size is scaled, making CV of the distribution asymmetrical to system size, with scaling exponent -0.5 as we have shown in Fig. 4b, and later on in Fig. 9c and Fig. 10a.

3.3.3 Discussion on scaling behavior of amplitude distribution in NDR mechanism for mild noise strength

Biochemical networks oscillating under the NDR mechanism behave as damped oscillations in the absence of noise. Assuming the oscillation amplitude is small, we can approximate its deterministic dynamics through a linear approximation of the attractor of the damped oscillator $\mathbf{X}_o = (X_{o1}, X_{o2})$, while using the noise strength at $\mathbf{X}_o$ to estimate the intensity of the Langevin term.

Around the attractor, the deterministic term $\mathbf{v}$ in Eq. (3.3) shall be expressed as its first order Taylor expansion:

$$\begin{aligned} v_i(\mathbf{X}) &\equiv \sum_{k=1}^r \omega_{ki} R_k(\mathbf{X}) \\ &= v_i(\mathbf{X}_o) + \sum_{j=1}^n \frac{\partial}{\partial X_j} \left(\sum_{k=1}^r \omega_{ki} R_k(\mathbf{X}) \right) \Big|_{\mathbf{X}=\mathbf{X}_o} (X_j - X_{oj}) \\ &= \sum_{j=1}^n \frac{\partial}{\partial X_j} \left(\sum_{k=1}^r \omega_{ki} R_k(\mathbf{X}) \right) \Big|_{\mathbf{X}=\mathbf{X}_o} (X_j - X_{oj}). \end{aligned}$$

Note that j is an iterator over the reactants (n of them, $n = 2$ in a two dimensional system).

Defining $M_{ij} \equiv \frac{\partial}{\partial X_j} (\sum_{k=1}^r \omega_{ki} R_k(\mathbf{X})) \Big|_{\mathbf{X}=\mathbf{X}_o}$, $y_i \equiv X_i - X_{oi}$, we have

$$v_i(\mathbf{X}) = \sum_{j=1}^n M_{ij} y_j \quad (3.31)$$

For stochastic term, to simplify the problem, we estimate the noise strength with the noise strength of the fixed point itself ($\mathbf{X}_o$). Hence the stochastic term $L_i(\mathbf{X}) \approx \omega_{ki} \sqrt{R_k(\mathbf{X}_o)} dB_j$. Let $\sigma_{ik} \equiv \omega_{ki} \sqrt{R_k(\mathbf{X}_o)}$, we have the final equation,

$$dy_i = \sum_{j=1}^n M_{ij} y_j dt + \sum_{k=1}^r \sigma_{ik} dB_j. \quad (3.32)$$

With this approximate description of NDR, we can prove the following theorem.

Theorem 1. *For a noise driven resonance depicted by Eq. (3.32), the moments of its*

distribution of the oscillatory amplitude χ satisfy the following scaling relation under cell size scaling $\hat{\eta}_c$:

$$\langle \chi_c^n \rangle = \eta_c^{\frac{n}{2}} \langle \chi^n \rangle, \quad (3.33)$$

where χ_c is the oscillation amplitude of the cell size scaled system.

Proof. Similar to the argument for the HBDD case, we may construct certain surface Γ_χ and correspond the experimentally measured amplitude point with the intersection between the system trajectory and the surface Γ_χ . Lang et al. facilitated this idea and obtained an equation to approximate the amplitude distribution of the oscillator of noise driven resonance.⁹² According to Lang's theory, the amplitude distribution of a noise driven oscillation can be approximated by

$$P(\chi) = N_\chi |v(\mathbf{y}(\chi))| P(\mathbf{y}(\chi)) d\chi, \quad (3.34)$$

where $\mathbf{y}(\chi)$ is defined as

$$\mathbf{y}(\chi) = \underset{\mathbf{y} \in \{\mathbf{y} \in \mathbb{R}^2 | \langle \mathbf{y} | \hat{e}_\chi \rangle = \chi\}}{\operatorname{argmax}} P(\mathbf{y}).$$

Note that $P(\mathbf{y})$ is the distribution of system state $\mathbf{y}$, and N_χ is the normalization factor. As before, $\hat{e}_\chi$ is the direction at which the amplitude is measured.

After the scaling operation, the amplitude distribution becomes

$$P_c(\chi_c) = N_{c\chi_c} |v_c(\mathbf{y}_c(\chi_c))| P_c(\mathbf{y}_c(\chi_c)) d\chi_c, \quad (3.35)$$

where

$$\mathbf{y}_c(\chi_c) = \underset{\mathbf{y}_c \in \{\mathbf{y}_c \in \mathbb{R}^2 | \langle \mathbf{y}_c | \hat{e}_{\chi_c} \rangle = \chi_c\}}{\operatorname{argmax}} P_c(\mathbf{y}_c).$$

The distribution of system state $\mathbf{y}$ can be obtained by corresponding Eq. (3.32) to the Fokker-Planck equation. The equivalent Fokker-Planck equation for Eq. (3.32) is

$$\frac{\partial}{\partial t} P(\mathbf{y}, t) = -\frac{\partial}{\partial y_i} [P(\mathbf{y}, t) a_i] + \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_j} [P(\mathbf{y}, t) b_{ij}], \quad (3.36)$$

where

$$a_i = \sum_{j=1}^2 M_{ij} y_j, \quad (3.37)$$

$$b_{ij} = \sum_{k=1}^r \sigma_{ik} \sigma_{jk}. \quad (3.38)$$

The solution of this Fokker-Planck equation is

$$P(\mathbf{y}) = N \exp(f(\mathbf{y})) dy^2, \quad (3.39)$$

where

$$f(\mathbf{y}) = \sum_{i,j=1}^2 m_{ij} y_i y_j,$$

$$m_{ij} = \sum_{k=2}^n 2(\mathbf{b}^{-1})_{jk} M_{ki} \left(1 - \frac{1}{2} \delta(j-i)\right),$$

with $\delta(j-i)$ and N being the Kronecker delta function and the normalization factor, respectively.

The distribution of system state for the cell size scaled network can be solved similarly. The result is

$$P_c(\mathbf{y}_c) = N_c \exp(f_c(\mathbf{y}_c)) dy_c^2, \quad (3.40)$$

where

$$f_c(\mathbf{y}_c) = \sum_{i,j=1}^2 m_{cij} y_{ci} y_{cj},$$

$$m_{cij} = \sum_{k=1}^2 2(\mathbf{b}_c^{-1})_{jk} M_{cki} \left(1 - \frac{1}{2} \delta(j-i)\right).$$

It is easy to prove that $M_{cki} = M_{ki}$, and $\mathbf{b}_c = \eta_c \mathbf{b}$. Thus we have

$$m_{cij} = \eta_c^{-1} m_{ij}, \quad (3.41)$$

$$N_c = \eta_c^{-1} N, \quad (3.42)$$

$$P_c(\mathbf{y}_c) = N \exp(f(\sqrt{\eta_c} \mathbf{y})) (d\sqrt{\eta_c} y)^2. \quad (3.43)$$

Inserting these results into Eq. (3.35), we find that if χ_c satisfies

$$\chi_c \equiv \sqrt{\eta_c} \chi, \quad (3.44)$$

we have

$$\mathbf{y}_c(\chi_c) = \sqrt{\eta_c} \mathbf{y}(\chi), \quad (3.45)$$

$$N_{c\chi_c} = N_\chi, \quad (3.46)$$

$$P_c(\chi_c) = P(\chi). \quad (3.47)$$

In other words, the probability distribution of χ_c is the same as that of $\sqrt{\eta_c}\chi$. Thus, the p th moment of χ_c :

$$\begin{aligned} \langle \chi_c^n \rangle &= \int \chi_c^n P_c(\chi_c) \\ &= \int (\sqrt{\eta_c})^n \chi^n P(\chi) \\ &= (\eta_c)^{\frac{n}{2}} \langle \chi^n \rangle. \end{aligned}$$

□

With Eq. (3.33), it is easy to see that the average of the oscillation amplitude, i.e., the first moment of χ , has a scaling exponent of 0.5. Thus normalized average amplitude $\langle \widehat{\chi} \rangle$ has a scaling exponent of -0.5 . Similarly, from Eq. (3.33), we have $\sigma_{\chi_c} = \eta_c^{1/2}$. And according to the definition given by Eq. (2.23), the coefficient of variation in amplitude CV_χ is scale invariant. These scaling exponents, again, mirrors the simulation results we obtained for Brusselator model (Table. 1).

3.3.4 HBDD and NDR Oscillation in High Dimensional Space

Although our results so far are mainly derived for 2D biochemical reaction systems, the idea is easily extensible to higher dimensional systems. Here we provides a brief discussion on the distinction between HBDD and NDR oscillations in higher dimensional systems.

HBDD and NDR are two simple mechanisms that explain the oscillatory behavior observed in biochemical networks, and both are confined in a 2D manifold. For the case of HBDD mechanism, the 2D manifold contains system's limit cycle. While for NDR mechanism, the 2D manifold contains the stable spiral. And if the trajectory converges

to the manifold quickly and do not deviate significantly from it, the system become equivalent to a 2D system and the discussion in Subsection 3.3.2 and Subsection 3.3.3 holds. Actually, because NDR mechanism is based on stable spiral, only 2 of its eigenvalues at the fixed point are close to zero (slow dynamics), hence the system must converge to the 2D manifold fast.

However, for oscillation under HBDD mechanism, the system might not necessarily converge to the 2D manifold before converging to the limit cycle. In this case, we will not be able to approximate the dynamics based on discussion made upon a 2D manifold. Nonetheless, the result is still applicable if the system satisfies a multi-dimensional version of conditions we used in our discussions for 2D.

Let us assume, like we did for 2D case, that our measurement of amplitude distribution corresponds to a measurement of flux of the system states over a $n - 1$ dimensional surface Σ_χ orthogonal to the base $\hat{e}_1$ pointing in the direction of the limit cycle. Let the intersection between limit cycle and surface Σ_χ occur at point A . By specifying the basis for measuring amplitude as $\hat{e}_2$, the amplitude of the limit cycle given by $\hat{e}_2$ is $\chi_o \equiv \overrightarrow{OA} \cdot \hat{e}_2$. For a trajectory intercepted by Σ_χ at point B during a period, the corresponding amplitude of that period is expressed as

$$\chi = \overrightarrow{AB} \cdot \hat{e}_2 + \chi_o. \quad (3.48)$$

Now, like what we have done for the 2D system, we morph the space, centering it on the limit cycle, the dynamics of the system then takes the form

$$\begin{aligned} dx_{c1} &= \eta_c v_1 \left(\frac{\mathbf{x}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_1 dB_1, \\ dx_{c2} &= \eta_c v_2 \left(\frac{\mathbf{x}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_2 dB_2, \\ &\dots \\ dx_{cn} &= \eta_c v_n \left(\frac{\mathbf{x}_c}{\eta_c} \right) dt + \eta_c^{1/2} \Gamma_n dB_n, \end{aligned} \quad (3.49)$$

which are defined in the domain

$$\begin{aligned} x_{c1} &\in [0, L_c], \\ x_{ci} &\in (-\infty, \infty), \quad \text{for } i > 1. \end{aligned}$$

The amplitude distribution, measured as flux distribution of trajectories over surface Σ_χ , is given by

$$\begin{aligned}
P(\chi_{c0} + x_{c2}) &\propto \int \cdots \int \Phi_{\Sigma_\chi}(\mathbf{x}_c) dx_{c3} dx_{c4} \cdots dx_{cn} \\
&= \int \cdots \int \left(v_{c1}(\mathbf{x}_c) P(\mathbf{x}_c) - \frac{1}{2} \Gamma_1^2 \frac{\partial}{\partial x_{c1}} P(\mathbf{x}_c) \right. \\
&\quad - \frac{1}{2} \Gamma_1 \Gamma_2 \langle dB_1 dB_2 \rangle \frac{\partial}{\partial x_{c2}} P(\mathbf{x}_c) - \frac{1}{2} \Gamma_1 \Gamma_3 \langle dB_1 dB_3 \rangle \frac{\partial}{\partial x_{c3}} P(\mathbf{x}_c) \\
&\quad \left. - \cdots - \frac{1}{2} \Gamma_1 \Gamma_n \langle dB_1 dB_n \rangle \frac{\partial}{\partial x_{cn}} P(\mathbf{x}_c) \right) \Big|_{\mathbf{x}_c \in \Sigma_\chi} dx_{c3} dx_{c4} \cdots dx_{cn},
\end{aligned}$$

where $P(\mathbf{x}_c)$ is the state distribution of the system in phase space.

Based on the assumptions

1. dB_i and dB_j are independent. That is, $\langle dB_i dB_j \rangle = \delta(j - i)$;
2. noise is moderate, trajectory stays within the neighbourhood of the limit cycle;
3. drifting from limit cycle does not affect phase speed of the trajectory;
4. retraction on direction $\hat{e}_2$ is phase independent, $m_{c2}(y_{c1}) = m_0$.

We can easily see that with the normalization condition of the state distribution, the amplitude distribution retains the form:

$$\begin{aligned}
P(\chi_{c0} + x_{c2}) &\propto \Phi_{\Sigma_{chi}}(x_{c2}) \\
&= C \exp \left(\frac{m_0}{\eta_c \Gamma_2^2} x_{c2}^2 \right).
\end{aligned}$$

Hence our conclusions for scaling coefficient CV_{χ_c} and χ_c is still valid.

3.3.5 Biochemical oscillators in real-world biochemical systems

Before further discussion, let's summarize the earlier discussion with an intuitive explanation for the differences between HBDD and NDR type of oscillation. In both the HBDD and NDR mechanisms, the variation of amplitude is determined by relative noise strength, of which the scaling exponent can be described by Eq. (2.20). On the other hand, the average of the amplitude is determined by different components

of the model. For HBDD, the oscillation is determined by the limit cycle. While noise disturbs the trajectory from it, the state distribution still centers on the limit cycle if the strength of the noise is moderate. Its average amplitude as determined by ODE dynamics of the system scales with $\sim \eta_c$. For NDR, however, its dynamics at the deterministic limit is a stable spiral, and the oscillation amplitude in this case is maintained by noise balancing with the local deterministic “force” pulling the system to the stable fixed point. In other words, the average of amplitude for NDR mechanism is determined by relative noise strength. Since relative noise strength and ODE dynamics have different scale exponents, CV_χ and normalized average amplitude $\langle \widehat{\chi} \rangle$ behave differently, as shown in Table. 1.

One issue worth noting for both HBDD and NDR oscillation is that, we have made several assumptions to simplify the form of the solution. Yet in actual biochemical systems, those conditions are not likely to be all met. For example, in the discussion on the NDR mechanism, because the oscillations could span a large portion of the phase space (see Fig. 3), especially when the system scale is small, our simplification of its deterministic dynamics with linear expansion as in Eq. (3.37) would lead to significant errors. Also for HBDD mechanism after space morphing, the dynamics is found to be non-uniform in space, where $x_2 > 0$ is denser than $x_2 < 0$ (see Fig. 5). Such unevenness in space is not addressed in our discussion. It could affect the dynamics significantly when system scale is small and the trajectory has larger probability to leave the vicinity of the limit-cycle. In fact we observe a difference in simulation result from the analytical solution for our simplified models when the system scale is extremely small. But when the system scale is relatively larger, our analytical prediction fits the simulation results very well (see Fig. 4).

Fig 6. Reaction networks of p53 oscillation. Black arrow are reactions. Black arrow with black side dots are Michaelis Menten kinetics or Hill equations, with the reactants reached by the side-dots being enzymes that catalyse the process. Red arrow indicates the translocation reactions between membranes. The p53-HBDD model is a time-scaled version of Dimitrio's model. For detailed reactions and their rate constants in p53-NDR model, refer to Appendix A.

3.4 Simulations of p53 regulatory network and *Hes1* auto-regulatory network

In order to verify our results in more realistic biochemical systems, we have conducted simulations on both the HBDD and NDR models for the p53 regulatory network and the *Hes1* auto-regulatory network.

p53 is a DNA binding protein which regulates the expression of various other proteins under the condition of DNA damaging cell stress. Based on different types and extent of DNA damage, p53 up-regulates different sets of proteins and triggers cell cycle arrest, DNA repair, or apoptosis.^{1,4,93,94} One interesting observation of p53 dynamics is that when ionizing radiation (IR) is applied to the cells, the concentration of p53, instead of staying in an activated higher level, pulses with a period of around 5.5 h in a rather noisy manner.⁸⁴ The protein Mdm2, one of p53's main regulators, goes through a similar pulsing process, with the same oscillation period as p53 and a stable delay in phase. Recently, experimental results have confirmed that a pulsing p53 level activates a different set of genes compared to p53 with a sustained level.⁹⁵ Thus a distinct signal-to-noise ratio of the pulsing amplitude is envisaged to ensure proper information transmission through the cellular signaling network.

Hes1 is also a DNA binding protein, whose function is the regulation of the expres-

Fig 7. Reaction networks of Hes1 auto-regulation. Black arrow are reactions. Black arrow with black side dots are Michaelis Menten kinetics, with the reactants reached from side-dots being enzymes that catalyse the process. $[E]$ are unspecified enzymes. Red arrow indicates the translocation process between cytoplasm and nucleus.

sion of several genes including itself. It is active in multiple tissues during embryonic development. Depending on the various environments with which it interacts, Hes1 is subjected to the regulation of the Notch signaling pathway^{78,96–98} within a self-regulatory network. Although experimental evidences show that the Notch signaling pathway is important in synchronizing and stabilizing Hes1 oscillation,^{78,96} there are also data showing that many ($\sim 56\%$) presomitic mesoderms still oscillate for over 12 h after dissociation from other cells (Notch receptor activation requires membrane contact of neighbouring cells),⁷⁸ indicating the possibility that under certain conditions, auto-regulatory network of Hes1 is enough to produce oscillatory behavior. With regards to the modelling of the oscillation induced by the Hes1 auto-regulatory network, some researchers have suggested that nonlinear degradation mechanisms make it possible for a Hopf bifurcation to arise in the ODE dynamics for certain parameter sets.⁸⁶ Conversely, there are also modelling work showing that without nonlinearity and HBDD, it is still possible for the system to oscillate relatively regularly under certain conditions⁴⁰ within the NDR mechanism.

Although our work is theoretical with the consideration of only intrinsic Poisson noise, we believe our research bears significance in elucidating the stochastic dynamics within these two networks. We make this claim based on two arguments. First, key reactants in both networks have short half-lives. For the p53 network, the half-life of p53 is $5 \sim 30$ min,¹⁰ and around 100 min after activation. Mdm2 has a half-life

of around 40 min.⁹⁹ Compared to the cycle period of 5.5 h,⁸⁷ the half-lives of these reactants are short. For the Hes1 auto-regulatory network, the half-life of Hes1 is around 22 min,¹⁰⁰ also short compared to the oscillation period of 122 min.⁷⁸ Thus, for both oscillatory networks, continuous protein generation and degradation (both are Poisson processes) are important in maintaining the cycle. Second, although the copy number of the reactants in the entire cell could be large, the copy number of gene regulators in the nucleus is moderate. Moreover, due to the large genome size, non-specific transcription factor-DNA interactions are predominant. Thus, the effective binding of transcription factors on cis-regulatory sites of DNA is low enough to show significant variations.¹⁰¹ In fact, there are research using solely intrinsic noise to model the noisy behaviors of both p53²² and Hes1 oscillatory networks.¹⁰²

The details of the reaction network models of p53 and Hes1 are described in the following two subsections.

3.4.1 p53 oscillation models

We use Dimitrio's model⁶⁸ as an example of HBDD model (p53-HBDD model, see Fig. 6a). Since the oscillation period in Dimitrio's model is too small (26.22 min) in comparison to experimental observation, we scale all reaction rates with a ratio $\eta_t = 12.586$ to make the oscillation period 5.5 h to achieve agreement with the experimentally observed period.⁸² The model (with reaction rates scaled) is stable without stress signal, and it goes through a supercritical Hopf bifurcation within a range of stress signal (from around $0.18 \mu\text{M}$ to around $8.09 \mu\text{M}$ ATM in both nucleus and cytoplasm). The limit cycle is observed as an oscillation with a period of 5.5 h. In our simulations, we use $2.7 \mu\text{M}$ as activating stress level that leads to the deterministically driven oscillation in this model (see Fig. 8(a)). To observe the signal-to-noise ratio of the oscillation amplitude of the model, we conducted a set of stochastic simulations using Gillespie algorithm on the reaction network. When simulating with this full dynamics, both the amplitude and period of the oscillations are found to vary from

its central value. Note that the central value relates to the deterministic part of the system, as exhibited in Fig. 8(b).

The NDR model (p53-NDR model) we used is also modified based on Dimitrio's model. In Dimitrio's model, active ATM phosphorylates p53 in both the nucleus and cytoplasm. However, ATM mainly localized in the nucleus of a cell, since the nuclear localisation signal (a section on the protein) is present on ATM and it plays an important role in ATM localisation in the nucleus. Both the inactive dimer form of ATM and the activated monomer form of ATM are associated with chromatin, though the nature of their interactions with chromatin differ between the two forms.¹⁰³ On the other hand, even if cytoplasmic ATM in the model is interpreted as an activator of p53 in general, they do not have significant presence in the cytoplasm either, i.e., the protein responsible for T18 site phosphorylation on p53 (the only site phosphorylation site responsible for the lessened affiliation of p53 to its regulator Mdm2), Chk2, is also localized and activated in the nucleus.^{104, 105} Based on these evidences, we believe that making kinase of p53 presents in cytoplasm as a stress signal does not make biological sense. Hence in our modified model, we removed ATM along with its associated reactions in the cytoplasm. The resulting reaction network (Fig. 6b) is described in detail in Appendix A. It is observed that the deterministic part of the new network no longer supports a stable oscillation at stress level $2.7 \mu\text{M}$ of ATM in the nucleus, as shown in Fig. 8(c), but the full stochastic dynamics over the same stress condition still does (Fig. 8(d)). A closer observation shows that the deterministic part of the model yields a damped oscillation at $2.7 \mu\text{M}$ stress level, while the noise of the system serves to maintain the amplitude of the oscillation to a certain level. It is also observed that the distribution of the concentration of all reactants involved in the oscillation is Gaussian (figure not shown). Thus the oscillation is noise driven.^{26, 87, 92}

3.4.2 Hes1 oscillation models

In the case of Hes1, the HBDD model of the auto-regulated oscillation is extracted from the full regulatory network presented by Barton *et al.*¹⁰⁶ The model possessed a mechanism of component degradation which is nonlinear (see Fig. 7a), and according to Morant *et al.*,⁸⁶ it is possible to achieve oscillations through a Hopf bifurcation. The model description and parameter of choice is described in Appendix A.

Without the presence of non-linearity in the degradation mechanism (see Fig. 7b), it is not possible to achieve Hopf bifurcation in the ODE dynamics of the self-repressed gene regulatory network. However, according to Wang *et al.*,⁴⁰ oscillatory behavior of relatively regular frequency can be observed by simulating the full stochastic differential equations within a certain parameter range. In our work, we have linearized the degradation mechanism of the HBDD model. We have also changed the transcription rate accordingly, up to the point where the Fano number of the oscillation period drops below ~ 0.3 . The details of the resulting Hes1 NDR model is also described in Appendix A.

3.4.3 Simulation method and results

For all the four models used to simulate the p53 and hes1 oscillatory networks, we use a same simulation method:

- For the simulation of stochastic dynamics, we used the Gillespie algorithm (see Section. 2.1.1 for more detail) directly applied to each of the reaction network described in Appendix A.
- For the simulation of deterministic dynamics, we used an automated program to dynamically generate the ODEs from each reaction network in Appendix A. Then we applied the adaptive step size 6th order Runge-Kutta algorithm to perform the simulation.⁷⁶

The full implementation of the biochemical model handling logic, the simulation algorithms and the simulations for each models described in this chapter are all packaged as the Supplemental Material of a paper published on Physical Review E.⁹¹

Fig 8. Time trajectory of Mdm2 and phosphorylated p53 in a p53-HBDD model with time scaled up by $\eta_t = 12.586$; b SDE version of p53-HBDD model; c deterministic simulation of p53-NDR model; d full stochastic simulation of p53-NDR model. For all plots shown above, cell sizes are scaled down by $\eta_c = (\frac{1}{2})^8$ when the simulation is performed.

In our simulation, the default cell size is $3,401.7 \mu\text{m}^3$ as is measured for the human fibroblast cell grown on a 60 mm dish by Russell *et al.*,¹⁰⁷ and the nucleus-cytoplasm volume ratio we used is 1 : 10.

At the proper simulation condition (*ATM* level at $2.7 \mu\text{M}$), the p53-HBDD model displays stable oscillations during ODE simulation,⁶⁸ while the p53-NDR model shows damped oscillations. In full stochastic simulation, however, the HBDD model displays oscillations with a very small coefficient of variation ($CV = 0.01$) on its amplitude. On the other hand, the NDR model displays steady oscillations with comparably larger coefficient of variation ($CV = 0.39$).

With cell volume scaled down by $\eta_c = (\frac{1}{2})^8$, the CV of the oscillation amplitude obtained from the p53-HBDD model is dramatically increased to around 0.2, while the CV of the oscillation amplitude from the p53-NDR modified model remains largely the same ($CV = 0.35$). The time trajectory of our SDE simulation for cell size scaled by $(\frac{1}{2})^8$ is shown in Fig. 8.

To further confirm our result, we performed a series of simulations with different scaling coefficient (Fig. 9), just like what we did for the Brusselator. During cell size scaling, normalized average oscillation amplitude for the HBDD model is virtually invariant of system size scale, $\widehat{\langle \chi_c \rangle} \sim \eta_c^{-0.02} \widehat{\langle \chi \rangle}$ (Fig. 9a), with $\widehat{\langle \chi_c \rangle}$ being the normalized average amplitude of the system scaled by coefficient η_c , as discussed before. For the NDR model, we found from our numerical results that $\widehat{\langle \chi_c \rangle} \sim \eta_c^{-0.43} \widehat{\langle \chi \rangle}$, with the scaling exponent being around -0.5 (Fig. 9b). Additionally, the standard deviation of amplitude σ_χ behaves the same for both mechanisms, with scaling exponent close to 0.5 . Also observed from numerical simulation is that the coefficient of variation of amplitude CV (defined by Eq. (2.23)) is scale invariant for the NDR model (Fig. 9d), while it has a scaling exponent of -0.5 for the HBDD model (as shown in Fig. 9c).

The results above indicate that, like what we have found in 2D systems, the HBDD and NDR mechanisms of p53 oscillation appear different when compared among cells with size varied in a range rather than being a constant. During cell size scaling, HBDD driven oscillations retain its normalized amplitude, thus the spikes are equally discernible across cells of different sizes. The same cannot be said, however, for NDR driven oscillations. Due to the -0.5 scaling exponent, the spikes are more visible for smaller cells, and less so for larger cells. On the other hand, NDR driven oscillations maintain its CV in amplitude constant. Thus during cell size scaling, the signal-to-noise ratio in pulsing remains the same for NDR models. On the contrary, the oscillations of the HBDD model in smaller cells is less regular than those in the larger cells, because of the scaling exponent of -0.5 in its CV of amplitude.

Similar results have also been observed when comparing the oscillations of the Hes1 auto-regulatory network between the HBDD model and NDR model. As shown in Fig. 10, the normalized average amplitude of Hes1 for the HBDD model has a scaling exponent of -0.03 , while that for the NDR model has a scaling exponent of -0.52 . This indicates that the normalized average amplitude of the Hes1 HBDD model is scale invariant, just like the p53 HBDD model that we have seen earlier. Conversely,

the normalized average oscillation amplitude of the NDR model is observed to be insignificant when η_c becomes very large. With regards to the *CV* of amplitude, consistent with that obtained earlier in p53 and in 2D systems, we again uncovered that the *CV* for the HBDD model has a scaling exponent of -0.45 , while the NDR model is scale invariant.

In summary, our simulation results have further affirmed the conclusions obtained via our Brusselator simulation and mathematical analysis performed earlier. For the HBDD oscillatory mechanism, the normalized average oscillation amplitude remains unchanged under cell size scaling operation. However, the coefficient of variation in oscillation amplitude becomes large when the system is sized down. On the other hand, the oscillations of the NDR mechanism maintain its coefficient of variation in amplitude during cell size scaling, but its normalized average amplitude becomes small when the system is sized up.

Fig 9. The results of simulation of p53 oscillatory network with cell size scaled by η_c . Each + is evaluated from a simulation of the model scaled by scaling coefficient η_c . a and b are with respect to the standard variation σ_χ and relative average $\langle \chi \rangle$ of the oscillatory amplitude for p53, respectively. a is for HBDD model, $\sigma_\chi \sim \eta_c^{0.51}$, $\langle \chi \rangle \sim \eta_c^{-0.02}$; b is for NDR model, $\sigma_\chi \sim \eta_c^{0.57}$, $\langle \chi \rangle \sim \eta_c^{-0.43}$. c and d relate to the coefficient of variation (CV) for the oscillation amplitude of p53 and Mdm2. c is the result for p53-HBDD model, $CV_{p53} \sim \eta_c^{-0.46}$, $CV_{mdm2} \sim \eta_c^{-0.46}$; d is the result for p53-NDR model, $CV_{p53} \sim \eta_c^{0.025}$, $CV_{mdm2} \sim \eta_c^{-0.04}$. We have used an ensemble size of 20 to estimate the statistics of the amplitude distribution at each size scale. Simulation time for each run is $T = 3.6 \times 10^5$ min. The first 6.0×10^4 min is used for system relaxation. The number of spikes sampled for estimation is around 1×10^3 .

Fig 10. These plots compare the CV and normalized average $\langle \hat{\chi} \rangle$ of oscillation amplitude in the Hes1-NDR and the Hes1-HBDD model. For CV of amplitude, $CV \sim \eta_c^{-0.45}$ in the HBDD model, and $CV \sim \eta_c^{-0.03}$ in the NDR model (shown in a). For the normalized average of amplitude, $\langle \hat{\chi} \rangle \sim \eta_c^{-0.02}$ for the HBDD model, and $\langle \hat{\chi} \rangle \sim \eta_c^{-0.52}$ for the NDR model (shown in b). Further downscaling of the cell size is also conducted. However in the regime of extremely low copy numbers, the oscillation of both models break down. For the NDR model, when $\eta_c < \frac{1}{2}^{-9}$, the Fano number of observed oscillations increases drastically. For the HBDD model, when $\eta_c < (\frac{1}{2})^{-5}$, the variation of the amplitude becomes large enough to make it difficult to distinguish the oscillation from noise. We have used an ensemble size of 20 to estimate the statistics of the amplitude distribution at each size scale. Simulation time for each run is $T = 2.0 \times 10^5$ sec. The first 4.0×10^4 sec is used for system relaxation. The number of spikes sampled for estimation in each simulation is around 1×10^3 .

3.5 Discussion

A difficulty faced in the building of biochemical network models is that, due to the scarcity of experimental data and a large number of degrees of freedom in the system, even a mechanism well fitted against the data can still be incorrect. In our work, we have demonstrated that two types of oscillatory mechanisms, both having been shown in the literature to be capable of representing pulsing phenomena observed in experiments,^{22,26,40,68,84–87} behave differently under cell size variation. The HBDD mechanism is found to maintain a constant normalized amplitude of the oscillation, while the NDR mechanism possesses an invariant coefficient of variation of the amplitude as cell size changes. Since the two mechanisms are not equivalent in the biochemical context, it poses the question of which mechanism is actually behind the oscillations in a specific biochemical network. In this context, our results have provided a theoretical basis for the conduct of experimental probes to elucidate the mechanisms underlying the oscillatory behavior observed in cells. For this, the use of the coefficient of variation of the amplitude CV_χ as an experimental measure against cell size changes would enable the inference of whether the HBDD or NDR mechanism is the correct mechanism underlying the oscillatory behavior. Although the normalized average amplitude $\widehat{\langle\chi\rangle}$ is not as convenient and appropriate an experimental measure as compared to CV_χ , its determination may be possible with the development of new experimental techniques in the near future.

Let us next evaluate the consequence of our main results from the perspective that oscillatory patterns are experimentally observed to directly contribute to downstream signaling in biochemical systems.⁹⁵ Since neither of the HBDD and NDR mechanisms hold both CV_χ and $\widehat{\langle\chi\rangle}$ constant simultaneously during cell size scaling, it seems to suggest that regardless of which oscillation mechanism is found to be the right one, robust oscillations for signaling against cell size changes would require an efficient size control mechanism for cells.¹⁰⁸

One of the limitations of our work is that we have not considered the effects of extrinsic noise. We have focused on intrinsic noise of the biochemical system, which basically arises from Poisson noise. The strength of Poisson noise is determined by the copy number of reactants, i.e., the system size. This has made it directly relevant for us to investigate the effect of cell size scaling on biochemical networks. Although cellular noise consists generally of both intrinsic and extrinsic noise, the latter does not have a clear relation with system size due to its diverse origins. Thus in order to experimentally apply our results to certain biochemical systems, it is necessary to either demonstrate the system has dominant intrinsic noise, or to separate the stochastic behavior due to intrinsic noise from that of extrinsic noise.

Another assumption made in our work is that cell size does not change significantly during the oscillation, as reaction chamber size is treated as a constant during the simulation. This is true for oscillatory networks with a period of a smaller time scale compared to that of cell division. However, for oscillations with a period comparable to that of cell division, such a limitation would jeopardize the usefulness of our method in distinguishing between the oscillation mechanisms. This is because cell size is no longer a constant in these cases. Possible ways to overcome this limitation is to halt cell division at various stages with chemical agents or to mutate certain genes that do not affect the oscillatory system directly. If we can keep cell size constant using these approaches, we would be able to distinguish between the two mechanisms based on the data collected from cells of different sizes.

In the next Chapter, we test our theoretical proposal on experimental data collected from Ultradian rhythm in gene-regulatory networks, i.e. p53 networks and hes1 networks, in which intrinsic noise is very likely to be dominant. Our aim is to identify the underlying mechanism of these observed rhythmic behavior using the method we discussed in this Chapter (also published on Physical Review E⁹¹), by comparing CV in amplitude for cells of different sizes.

4 Applying to Experimental Data

In this Chapter, we shall describe our effort in testing our proposed method that statistics of coefficient of variation in oscillation amplitude from cells of different sizes can be used to determine whether an observed oscillation is induced by noise driven resonance or caused by deterministic limit cycle (see Chapter 3). We applied our method and analyzed p53 oscillatory data provided by Lahav Lab⁸⁷ and hes1 oscillatory data provided by Ryoichiro's lab from Tokyo University.⁷⁹ Result indicates that hes1 oscillation is caused by deterministic limit cycle. Yet for p53, the result is not conclusive.

4.1 Observation for P53 oscillation

The experimental data for p53 oscillation we obtained from Lahav lab. It is gathered from a clone of the cell lines, MCF7.

MCF7 cell line is a cancer cell line first isolated in 1970 from the breast tissue of a 69 year old Caucasian woman.¹⁰⁹ Originally, the cells of MCF7 contained 85 chromosomes, which was later reduced to 69 in modern MCF7 cell line. MCF7 cell has a cell cycle period of around $20 \sim 23$ hours,¹¹⁰ and usually considered to be easy to propagate. While MCF7 cells generally grow in epithelial like monolayer *in vitro*, the cells are prone to overlap each other and forming domes, making it hard to discern the cell boundaries under a microscope. The MCF7 clone used in their experiment is stably transfected with CFP fused P53 gene (p53-CFP),⁷⁷ and YFP fused MDM2 gene, where CFP (cyan fluorescent protein) and YFP (yellow fluorescent protein) are fluorescent proteins, modified upon GFP from jelly fish *Aequorea victoria* for a much shorter half life (~ 5 min for CFP).¹¹¹ CFP and YFP with a short half-life has minimal cumulation effect. Since cells with CFP fused P53 gene express CFP and p53 proportionally, with YFP fused MDM2 gene express YFP and mdm2 proportionally, making CFP and YFP ideal reporters to increase and decrease of p53/mdm2 expression rate in the cellular dynamics.⁷⁷

A stable p53 oscillation in CFP-p53 transfected MCF7 clones is triggered using γ -irradiation. The CFP and YFP expression level in the cell are then tracked in separate channels using time-lapse fluorescent microscope with an interval of 20 min between each subsequent frames. The level of p53 in the cell is then inferred through the expression rate of mdm2, represented by the yellow fluorescent intensity by YFP. A specific portion of nucleus from each individual cell is used to obtain the fluorescent readout from data frame. The original data images along with their fluorescent readout over the 72-hour observation is shared by Dr. Lahav's lab.

Oscillatory portion from each cell's fluorescent time trajectory is extracted using

spike recognition threshold of $0.15\sigma_y$ and time window of 4 mean period. The extracted oscillation is then linearly detrended to get rid of the slow shift in the observed fluorescent strength. The resulting time trajectories are shown in Fig. 11.

Fig 11. Extracted p53 oscillation from the MCF7 cells. Each sub-figure is a different cell. In each sub-figure, x axis indicates time with unit of minutes, y axis indicates fluorescent intensity of a cell with arbitrary unit. Black thin line is the original time trajectory. Blue dashed overlay is the detected oscillation, red solid line is the detrended oscillation. Oscillation statistics like coefficient of variance is to be extracted with the red line.

To obtain the cell size of the observed cells is tricky. One possible method is using the nucleus size to represent the cell size. However, although the data we obtained is treated with Cy5 dye, the majority of the cells do not have their nucleus sufficiently dyed such that a clear nucleus boundary is discernible (see Fig.12a). Instead, we measured the cell's area by measuring the area enclosed by cell boundaries through phase contrast microscope, as shown in Fig. 12b (thanks to Lahav Lab for providing

the images). This is tricky for two reasons. First, when cells are crawling over the substrate, its surface area is constantly changing due to its flexible shape. Second, MCF7 cells tend to overlap each other, forming domes (cell 11, 12, 1 in Fig. 12b, for example). This not only makes the measurement of area of overlapping cells next to impossible, it also makes the relation between measured area and cell volume more complicated. To counter these two issues, we only measure the area of the non-overlapping cells; and we represent the size (volume) of the cell by averaging the readout of each cell's area (in number of pixels) across the first 145 frames.

As shown in Fig. 12c, scaling relation between oscillation's coefficient of variation and the cell size is inconclusive for p53 oscillation in *MCF7* cell. Several aspects of the experiment could be improved before a conclusion can be drawn. Due to the genetic variability of MCF7 cell line and a potential radiation induced variability among the observed cells, a larger sample size is needed to separate the variation induced by different cell size and variation induced by other factors. On the other hand, since MCF7 cells tend to overlap each other, the overlapping cells could make the fluorescent readout unreliable. It is preferable for an isolation mechanism to be employed so that the overlapping of fluorescent cells can be minimized. Although the p53 oscillation period and cell cycle period of MCF7 cell is of a comparable time scale, we are not concerned by the potential influence of cell cycle over the cell size or the p53 oscillation, because the cells that are undergoing p53 oscillation are radiation treated and has entered growth arrest.⁸²

(a)

(b)

(c)

Fig 12. (a) displays the result of nucleus markings by Cy5 dye. (b) displays the cells under phase contrast microscope. Numbers overlaying the images are identified cells. Note that not all identified cells have discernible oscillation. Time trajectory of the cells with discernible oscillation are shown in Fig. 11. (c) is the result of the scaling relation between cell size and CV of oscillation amplitude for each of the different cells, x axis is the cell size of arbitrary unit; y axis is the measured CV of the cell; the overlay number of each point is the ID of the cell.

4.2 Observation for *hes1* oscillation

The experimental data for *hes1* oscillation was obtained from Ryoichiro’s lab in the University of Tokyo. It is gathered from a clone of the cell lines, NIH 3T3.

NIH 3T3 is a cell line obtained from mouse embryo tissue using the ‘3T3’ protocol.¹¹² The cell cycle period of each individual cells varies from 13 to 23 hours,¹¹³ with an average of around 20 hours. Unlike MCF7 cells we have dealt with earlier, the 3T3 cells which were grown on 2D substrate *in vitro* under moderate concentration, do not have the tendency to overlap each other. This proves to be helpful in determining both the cell boundary and the fluorescent level of the cell reliably. Although *Hes1* oscillation can occur in the absence of Notch signaling,⁷⁸ it is likely that some factors is still needed for *hes1* oscillation to occur. To quote Ryoichiro, “serum can induce *Hes1* oscillation, but this oscillation is rather unstable, probably because factors in serum are unstable”. The *hes1* oscillatory data we received from Ryoichiro’s lab is observed in NIH 3T3 cells with a stable NICD (cleaved Notch intra-cellular domain, the activated form of the Notch receptor) expression maintained by the Tet-On system.

The *hes1* expression is reported by a modified firefly luciferase (Ub2-Luc). Firefly luciferase can illuminate in the presence of luciferin. However, to use it as a reporter protein, it needs to have a shorter half-life than *hes1* in cellular environment. Ryoichiro’s team showed that with ubiquitination, luciferase’s half life can be shortened to 10min with single ubiquitination (UB1-Luc), and 6min with double ubiquitination (Ub2-Luc), both are qualified to be *hes1*’s reporter protein. They then use a 2.5-kb *hes1* promoter to drive the ubiquitinated luciferase’s expression (Hes1-Ub2-Luc), and transfected the Hes1-Ub2-Luc gene into NIH 3T3 cells as *hes1* expression’s reporter.⁷⁸ The illumination of the reporter protein is then captured using time-lapse microscope.

Similar to p53 data, oscillatory portion from each cell’s fluorescent time trajectory is extracted using spike recognition threshold of $0.15\sigma_y$ and time window of 3 mean

period. The extracted oscillation is then linearly detrended to get rid of the slow shift in the observed fluorescent strength. The resulting time trajectories are shown in Fig. 13b - j.

To evaluate the location of those cells, we made use of the H2B-mCherry staining, which was used for pinpointing the nucleus location (see Fig. 13a). We encircle the boundary of stained oval area of each cell, and count the pixel number of the encircled area as an indicator of the nucleus size. We shall use the nucleus size as an indicator of the cell size of these observed cells.

With cell size and oscillation statistics measured, we summarized the relation between the oscillation amplitude's CV and the cell size in Fig. 14. Interestingly, despite the small sample data, the result can be fitted with an exponent of -0.68 , which is very close to the HBDD paradigm, suggesting that the *hes1* oscillation stimulated by stably expressed NICD is indeed driven by a deterministic limit cycle.

(a) H2B-mCherry stained nucleus

(b) Cell 20

(c) Cell 35

(d) Cell 43

(e) Cell 52

(f) Cell 202

(g) Cell 246

(h) Cell 719

(i) Cell 821

(j) Cell 909

Fig 13. (a) is the nucleus of NIH 3T3 cells with H2B-mCherry staining. The three images are snapshots at the beginning of the observation(0 minute mark), 705 minute mark and 1410 minute mark, respectively. Each cell is encircled in different color and marked with an ID. (b) - (j) are time trajectories of cells from which we can extract stable oscillations. For these plots, x axis is time in minute, y axis is illumination in arbitrary unit. Black dashed line is the original trajectory; blue dashed overlay is the detected oscillation; red line is the detrended oscillation result.

Fig 14. The scaling relation between cell size and hes1 oscillation amplitude's coefficient of variation. x axis is the cell size in arbitrary unit; y axis is the coefficient of variation. Each red dot represents a cell, which is labeled with their respective cell ID. The blue line is the fitted result using linear least-squares regression over the log of both oscillation amplitude CV and cell size. The fitted scaling exponent is -0.68 .

4.3 Summary

Although the lack of sample size could make our results less reliable than desired, it shows the promise of our methods in using the scaling statistics of oscillation amplitude CV in cells of different sizes to infer the mechanism behind the oscillation. Cells, especially cell lines that get frequent use by experimentalists, vary greatly in sizes. With reliable nucleus staining, it is possible to measure the nucleus of each cell to represent the cell size. For the diverse interesting oscillatory processes like p53 and hes1 oscillation, experimentalists have shown to be capable of obtaining reliable oscillation statistics using reporter proteins.^{77,78} With sufficient data of both oscillation statistics and cell size for each individual cell, our approach provides a way to determine the underlying mechanism behind the oscillations. We are hopeful to see our method being applied by experimentalists to advance their understandings on the oscillatory behavior of the systems of interest to them.

5 Habenula Research

5.1 Introduction

The dynamics of the brain has always been a fascinating subject to research. As the central control system of an animal, it needs to process and respond to a good variety of external/internal stimuli in a short time. This requires the brain to maintain both the complexity within the neuronal network and to possess a good information transmission rate. In order to achieve such functionality, brain tissue is highly metabolically active. As a structure composed of only 2% of body weight, the brain of an average adult human accounts for 20% of calories burned by the body.¹¹⁴ Since decades ago, it was already established experimentally that brain metabolism is highly constant regardless of whether the subject is mentally active.¹¹⁵ And the high rest state metabolism was attributed to spontaneous activities of the neuronal network.¹¹⁴ Nonetheless, with the complexity and high plasticity of the neuronal network (let alone to the sheer number of neuronal cells in the system) how a brain turn the highly constant metabolism into its complex yet prompt responses would elude researchers for a long time.

Spontaneous activities in neuronal network has long been thought to be purely stochastic in nature. However, last two decades saw an increased interests on the source of such spontaneous activity and a shift from regarding spontaneous activities as simple noise to considering it as an active functionality of the brain. Several pieces of theoretical research showing that with balanced excitatory and inhibitory input, a chaotic behavior can be observed in simulated neuronal network.^{64,65} On the other hand, experimentally, we consistently find that the spectral analysis of spontaneous activity follows power law distributions,^{116–118} suggesting the likelihood of critical state.^{119–121} With the advancing understandings of balanced excitatory and inhibitory input, researchers found that network with balanced input could also achieve the $1/f$ phenomena observed in various spectrum analysis.⁵² These findings, along

with the recognition that neuronal network at critical state has the maximized information capacity and transmission,¹²² seem to suggest a model of how cortical neuronal network converts the elevated metabolism into the functionality of prompt response to various inputs, i.e. with the mechanism of balanced input, cortical neuronal network have strong likelihood of maintaining a self-organized critical state.

That being said, it also worths noting that observing the $1/f$ phenomena in spectral analysis does not necessarily indicates a critical state in the neuronal network. Bedard et al. shown from theoretical perspective that it is possible for non-critical dynamics to display some type of $1/f$ behavior in power spectral analysis.¹²³ When facing the spontaneous activity of a specific cerebral neuronal network, it is necessary to explore all the three possibilities: there exists chaotic attractor, there exists critical attractor, or the phenomena is driven by pure noise.

The habenula is a highly conservative brain structure from the evolution perspective. It is present in virtually all vertebrate species. Some of its efferent and afferent connections are also present in some of the the most primitive vertebrate (lamprey) to mammals.^{124–126} It has been shown that habenula is vital for stress evasion, value-based decision-making, reproductive and maternal behavior in vertebrate species. Moreover, the abnormality of the habenula has also been linked with psychiatric disorders in human.¹²⁷ Over the years, a lot of experiments has been conducted on the physiological and/or neuron connectivity perspective on various model animals such as lamprey, zebrafish, rats and macaque monkeys.^{124, 126–128} Specifically, zebrafish has been extensively used to observe the habenula connectivity and signal transmission pathway because they are transparent during early development and are genetically accessible.¹²⁵

Zebrafish habenula has been confirmed to be connected to a variety of other brain structures. It receives signals from olfactory bulb, subpallium and preoptic areas, ventral Entopeduncular nucleus, parapineal and posterior hypothalamic lobe. There are also observation indicating that thalamus also projects to habenula.^{126, 128, 129} The

processed output of habenula are also diverse. Medial (dorsal) habenula is known to be connected to and sending signals to interpeduncular nucleus; lateral (ventral) habenula has also been found to project to raphé nuclei, hypothalamus, ventral tegmental area, thalamus, etc.^{125,130} Accounting also the multiple roles habenula takes in various processes, we believe that habenula as a neuronal network is likely to be in the critical state to handle the various signal processing with effectiveness.

To test our hypothesis, we chose to analyze the spontaneous neuron activities of habenula in transgenic (GCaMP6f) zebrafish embryo after clamping it down and placing it in the dark environment. In our research, we conducted analysis using nonlinear time trajectory analysis to find out the existence of chaotic attractor. Results show that the potential chaotic attractor cannot be of low dimensions. We also find that the spectrum analysis of the habenula cells display a distinct power law pattern of exponent -1.5 , suggesting a critical state

In the following section, we shall describe each part of this piece of research in detail.

5.2 Material and Method

5.2.1 Experimental Method

To record the spontaneous firing patterns of habenula at individual cellular level, we employed two-photon laser-scanning confocal fluorescence microscope to observe. A series of observations are conducted over the rest state habenula on transgenic zebrafish line expressing the calcium indicator GCaMP6f, with the zebrafish clamped down in dark environment for an extensive period of time. To obtain the highest temporal resolution possible, we conducted single focal plane scanning on the horizontal plane that captures the largest number of habenula cells. Frequency of the focal plane scanning is 7Hz. Duration of our observation is around 15min. It is difficult to conduct experiment for a period of time longer than 15min due to a likely uncorrectable relative movement between fish and microscope orthogonal to the observation plane over a prolonged period of time. The cause of such movement could be due to the thermal effect of the equipment.

Each observation yields a stack of fluorescent data images recording neuronal activities of the habenula cells intersected by the observed focal plane. The image stack is taken at 7 frame per second over a period of around 15min(see Fig. 15a). The experimental data was obtained from Dr. Jesuthasan’s lab of the Institute of Molecular and Cell Biology of A-Star.

5.2.2 Image Processing

Since the shape of neuronal cells is constant over time, and that the relative position between the habenula neuronal cells stay static, it is conventional to use static cell masks to overlay the image stack to obtain the fluorescent signal from each individual cell.^{131,132} In this section, we shall discuss the image processing method we employed for our dataset in order to obtain the time trajectory of each cell’s fluorescent level.

Motion Correction The necessity of motion correction arises from the observation that for the duration of the recordings, the fish habenula’s relative position to the camera usually goes through a significant drift over the observation plane.¹³³ To represent each neuron by its corresponding static cell mask, it is vital to conduct sufficient motion correction to compensate the displacement.

Since the size of image of our data set is relatively small (256x128 pixel), the movement between consecutive frames could be as little as a fraction of pixel. Hence the use of cross correlation method on spatial domain directly would not yield meaningful result. Trying to estimate the drift over two frames with significant time distance in between, on the other hand, will also fail to yield good result. Since different neurons fire at different timings, two frames with large time interval between each other tend to have a vastly different firing patterns. A confused correlation based algorithm would then yield unreliable, and sometimes nonsensical result for the motion correction.

Given the consideration stated above, we employed a Fourier based sub-pixel image registration algorithm.¹³⁴ This algorithm is cross correlation method implemented on image pair’s frequency domain, to estimate the displacement on observation plane between each pair of consecutive image frame, then by accumulating the displacement over time to obtain the required translation to align any image with a pre-chosen reference frame.

The result of the estimation of displacements using Fourier based sub-pixel registration algorithm is satisfying. However, correcting the displacement using spatial displacement’s Fourier transform tend to contain multiple abnormal bright dots in resulting motion corrected images, reducing data’s reliability. Hence after obtaining the cumulated displacement of each image relative to the reference frame, we round the displacement to nearest integer number of pixels, then directly translate the image in spatial domain. This would introduce some random shaking in the resulting motion corrected images. However, since the remaining displacement is within 1 pixel, while a cell’s diameter is usually over 10 pixels, ignoring such small displacement errors is

acceptable.

(a) Habenula Single Layer: Full Image

(b) Habenula Single Layer: Masked Image

(c) Habenula Single Layer: extracted ROI

Fig 15. Single Layer Image of Zebrafish habenula. (a) original image (motion corrected); (b) image after manually removing the neuropil, skull, and non-habenula region to get rid of interference; (c) extracted ROI using STICA algorithm.

ROI finding and fluorescent data extraction Taking advantage of the temporal activities of the neuron, we use Spatial-Temporal Independent Component Analysis (STICA) to capture neuronal cell with the representation of the regions of interests (ROI) across the image stack to keep track of each individual cell's fluorescent level over time (see Fig. 15c). The time trajectory of fluorescent level of each ROI then help us to derive the firing activities of neuron those ROI represents.

In order to reduce the background noise that interfere with the result, we manually masked out the regions that are not clearly distinguishable with the habenula

neuronal cells (see Fig. 15b) before applying STICA. After applying STICA method, the automatically generated ROIs are then manually sifted before use (see Fig. 15c).

After the ROIs are found, we use each ROI as a cell mask, extracting its fluorescent level at each time frame using the average of pixel value within the mask.

5.3 Non-linear time series analysis

Under the assumption that the spontaneous activities of habenula being chaotic rather than simply noisy, we performed a quantitative analysis into its chaotic nature by means of the approach of non-linear time series analysis.

The basis of non-linear time series analysis is that the property of a multi-dimensional attractor can be reconstructed using a single observer dimension by employing the time-delay embedding theorem.¹³⁵

For recordings $x_t, x_{t+1}, \dots, x_{t+i}, \dots$, a time series of m embedded dimension can be constructed as

$$\vec{x}_t = (x_t, x_{t+\tau}, \dots, x_{t+m\tau}), \quad (5.1)$$

where τ is a number larger than 0, referred to as the time lag. With this method, an m -dimensional phase space can be reconstructed from a single dimensional time trajectory (see Fig. 16a, 16b for an example of deterministic chaos obtained from Lorenz system). It is worth noting that both m and τ requires careful consideration for the reconstruction to represent the original dynamics. If τ is too small, the two neighboring dimensions could be so correlated to each other that the vector $\vec{x}$ are clustered around the diagonals in the $\mathbb{R}^m$ state space. On the other hand, if τ is too large, each of the dimensions could behave independent from each other, resulting in a unrecognizable “cloud” in the state space. Hence it is preferable for τ to be as small as possible, yet big enough to de-correlate the neighboring dimensions. As described in,¹³⁶ the most suitable time lag τ is the first zero of the autocorrelation function,

$$c_\tau = \frac{\langle x_t x_{t-\tau} \rangle - \langle x_t \rangle^2}{\sigma_{x_t}^2}, \quad (5.2)$$

on the signal’s time trajectory.

With a properly chosen time delay, we made our attempt to find the dimension of the potential chaos attractor in the time trajectory of habenula’s single cell fluorescence recordings. We used correlation dimension for dimension measuring. This measure

is defined with the shrinking of the size of neighborhood ϵ by measuring how the number of the “neighboring states”, $C(\epsilon, N)$ decreases in a state space of N consecutive frames.¹³⁶

$$d(N, \epsilon) = \frac{\partial \ln C(\epsilon, N)}{\partial \ln \epsilon}, \quad (5.3)$$

where $C(\epsilon, N)$ is defined with Heaviside step function $\Theta(x)$ as follows:

$$C(\epsilon, N) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\epsilon - |\vec{x}_i - \vec{x}_j|), \quad (5.4)$$

$$\Theta(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1, & \text{if } x > 0. \end{cases} \quad (5.5)$$

Since correlation dimension d is defined on the reconstructed state space $\mathbb{R}^m$, the quality of the reconstruction would affect the effectiveness of finding the correct correlation dimension from the original dynamics. For obvious reasons, by analyzing a $m = 2$ reconstructed state space $\mathbb{R}^m$, one can only observe a correlation dimension up to 2. And if the actual dynamics has correlation dimension larger than 2, one needs a higher embedding dimension in order to capture it. According to Kantz, et al.,¹³⁶ if m is larger than twice the box counting dimension D_F , regardless of how high the dimensionality of the original state space is, the attractor formed by $\vec{x}_t$ is equivalent to the original attractor, and hence the measurement of correlation dimension is approximately equivalent to the box counting dimension in the original state space (see Fig. 16c for an example of deterministic chaos from Lorenz system).

For noisy time trajectory with limited sample size, it could be difficult to tell if the observed dimension is from the actual dynamics or induced by anomaly due to limited sample size. To distinguish them, a different measure is required, and we use the (ϵ, τ) -entropy per unit time, defined as

$$h_q = \sup_{\mathcal{P}} \lim_{m \rightarrow \infty} \frac{1}{m} H_q(m, \mathcal{P}_\epsilon) \quad (5.6)$$

(a) Original State Space

(b) Reconstructed State Space

(c) Correlation Dimension

(d) Entropy per unit time

Fig 16. We display here the results of nonlinear time series analysis for the Lorenz system to serve as an example. (a) shows a trajectory of the Lorenz system with $\sigma = 10$, $\beta = \frac{8}{3}$, $\rho = 28$. (b) is the reconstructed state space based on the variable x at each time point using embedded dimension, with τ in this plot being chosen as 100 units. (c) displays the correlation dimension obtained using Eq. (5.3) based on measuring correlation against the neighborhood size ϵ . From bottom to top, the lines are obtained by increasing the embedding dimensions from 1. The correlation dimension is measured to be around 2. (d) displays the entropy per unit time of this chaotic system. From top to bottom, the lines are obtained by increasing the embedding dimensions from 1. For embedding dimension 1 and 2, the entropy per unit time increase linearly against a decreasing $\ln \epsilon$, because such low embedding dimension is not enough to capture the deterministic chaos with correlation dimension of 2. Starting from embedding dimension 3 and above, the measured entropy per unit time start to behave as the one expected in deterministic chaos. For a range of ϵ , the entropy per unit time is bounded below around 0.96, which is the sum of the positive Lyapunov exponents of the system. At small ϵ , on the other hand, the entropy per unit time increases again due to measurement noise.

for this purpose. Note that the entropy $H_q(m, \mathcal{P}_\epsilon)$ is given by

$$H_q(m, \mathcal{P}_\epsilon) = \frac{1}{1-q} \ln \left(\sum_{i_1, i_2, \dots, i_m} p_{i_1, i_2, \dots, i_m}^q \right). \quad (5.7)$$

The parameter m here is the number of time intervals, and $H_q(m, \mathcal{P}_\epsilon)$ measures the entropy of the dynamical system over the period of time $m\tau$ by cover $\mathcal{P}_\epsilon$. Hence h_q is called the entropy per unit time τ with cover bounded by size ϵ .

An early work published by Gaspard, et al.¹³⁷ showed that (ϵ, τ) -entropy per unit time behaves differently between stochastic system and system at deterministic chaos when ϵ varies. For stochastic system,

$$h_q = -\log(\epsilon) + h^{(c)}(m). \quad (5.8)$$

On the other hand, for chaotic system, entropy per unit time is bounded by the sum of the dynamics' positive Lyapunov exponent λ_i , according to Pesin theorem.¹³⁷

From the observational point of view (see Fig. 16d), entropy per unit time tend to increase linearly with regards to a decreasing $\log \epsilon$ for purely stochastic system. For deterministic chaos, on the other hand, While entropy per unit time increases close to linearly with regards to a decreasing $\log \epsilon$ initially, it eventually reaches a certain plateau bounded by $\sum_i \lambda_i$. Note that even for deterministic chaos, the entropy per time measured is small epsilon will increase again with decreasing ϵ , as if the system is stochastic. This increment is caused by the measurement noise.

With the knowledge above, we conducted a non-linear time series analysis over the fluorescent data. The analysis is performed on each of the observed neuronal cell. As shown in Fig. 18 and 19, none of the observed neuronal cells in the rest state habenula displayed any sign of deterministic chaos in any of the reconstructed state space up to 10 embedding dimensions. We are not able to reconstruct any meaningful higher dimensional state space due to experimental limitations.

Although it is still possible that the spontaneous firing within the habenula is purely stochastic, the latest findings in our team's population analysis seems to reject this possibility. We can also rule out the factor that the measurement noise and the

(a) Cell 8

(b) Cell 21

(c) Cell 28

(d) Cell 29

Fig 17. Observed trajectory and 3D reconstructed state space for fluorescent signal of 4 cells with ID 8, 21, 28, 29, respectively from one of our observed fish.

(a) Cell 8

(b) Cell 21

(c) Cell 28

(d) Cell 29

Fig 18. Extracted correlation dimension from the four cells (ID 8, 21, 28, 29 from one of our observed fish). For each cell, we attempt to capture the correlation dimension from the observed fluorescence signal by embedding the signal into a reconstructed state space that is up to 20 dimensions, the color code for observed correlation dimension for each embedded dimension from low to high is a cyclic sequence of blue, green, red, cyan, purple, yellow and black. For each embedded dimension, we search through the neighborhood size ϵ from a maximum data interval in the reconstructed state space to $1/1000$ of that value, and we calculate the correlation dimension using Eq. (5.3). As we can see from the figure, correlation dimension does not converge as embedded dimension increases. With increasing embedded dimension, we also notice that data points become too scarce to obtain a valid readout with a relatively large ϵ). This indicates that the observed data is not sufficient to determine the observed spontaneous activity's correlation dimension, if there is any.

Fig 19. The figure above displays the relation between ϵ and entropy per unit time h for habenula's spontaneous activity. It shows that $h \sim -\ln \epsilon$ for the entirety of the scanned ϵ . Also, h is not bounded as it should be if there is deterministic chaos in the system.

noise in the dynamics itself make the deterministic chaos unobservable, since evidence of critical behavior of the network can already be observed with this and even more sparsely sampled data set.

There could be two possible reasons for the lack of convergence in the measured attractor dimension for rest-state habenula. First, the dynamics of habenula is under constant perturbation from other brain structures, such that the system of interests never actually converge to its attractor. Second, the dimension of the system's strange attractor may be too high to be captured by the limited experimental data (as is supported by another observation that the false nearest neighbors¹³⁶ of reconstructed state space is larger than 10% up to 12 embedding dimensions for most of the cells observed in our experimental data).

5.4 Spectral analysis on single cell firing pattern

We then started our exploration on the possibility that the observed spontaneous activities reflect a resting habenula in its critical state. Since the key feature of a neuronal network being in a critical state is revealed through its spectrum, we first performed spectral analysis over the single cell firing patterns observed from our samples. Because we do not have many cells in each observation, we have not performed population analysis. To avoid additional arbitrary error from the deconvolution process, we calculated the spectrum from the original fluorescent data for all the observed cells. Result of the analysis (see Fig. 20) displays a distinct power law behavior of exponent -1.49 ± 0.20 . The result appears promising, as numerous independent research on critical neuronal network has also obtained a power law of exponent -1.5 in the slow dynamics,^{52,119–121} albeit such power law exponents are usually reported on avalanche size in the population analysis.

Fig 20. Spectral analysis result of the fluorescent time trajectories of 4 cells (ID 8, 21, 28, 29) in one of our sample fish(fish 204) observed in the rest state. The spectral data has been binned by every 50 points in frequency space. The resulting spectrum appears to follow the power law. The exponents for the four cells are -1.39 , -1.53 , -1.65 , -1.52 , respectively.

One continuation of this work would be to confirm that the rest state habenula is indeed in a critical state by conducting multi-layer scanning of fluorescent data in order to obtain more sample cells for population analysis. Although the scanning frequency would be as low as 1Hz if multi-layer scanning is attempted, the reduced

scanning frequency (from current 7Hz scanning rate) should not obtained from affect the result of our analysis. According to our analysis on fluorescent time trajectory, the power law persists in the lower frequency realm. Indeed, the latest population analysis of zebrafish habenula by our team displayed an power law exponent of -1.5 in avalanche size and an exponent of -2.0 in avalanche time duration, in accordance to the exponents obtained theoretically from the critical neuronal network⁵² (result is yet to be published as of the time of this writing). The result also suggests a potential connection between the noise spectrum of single neuron fluorescent time trajectory and the avalanche size in terms of the number of spontaneous cell firings.

5.5 Conclusion

To unravel the property of habenula's rest state dynamics,^{124–128,130} we performed a series of experimental data analysis on its spontaneous activities. Based on the existing research, we hypothesized that the spontaneous activities could be driven by either a chaotic attractor^{64,65} or via self-organized criticality.^{52,64,116–120,138} In our effort to find the chaotic attractor of the rest-state habenula using non-linear time trajectory analysis,¹³⁶ we observe an outcome that do not converge up to 10 embedded dimensions. This indicates that either the dynamics of the rest state network do not possess a low-dimensional attractor, or the constant perturbation from other brain structures prohibit the habenula from converging to its attractor. On the other hand, our analysis using spectral analysis yields a power law behavior with exponent -1.5 , which indicates that the habenula is potentially exhibiting a critical behavior. However, more research is still required to verify if the power law distribution we observe reflects the one observed in a neuronal network of balanced input that operates at critical state.⁵² In the next phase of our research, my colleagues are performing extensive pre-processing on the obtained data and creating data analytic tools to perform avalanche analysis. Our goal is to obtain sufficient evidence to determine if a spontaneously active habenula is in a critical neuronal state.

6 Conclusion

Attractor analysis in cellular processes could help us gain insights on the underlying dynamical system when observables are limited in degree of freedom or resolution. In this thesis, attractors in two different cellular context is discussed. In the context of biochemical network, we discussed on how two different attractors interact differently with intrinsic noise. This difference leads to distinct scaling behavior in the statistics of the amplitude of oscillation induced by these two attractors with respect to the cell size. Following the discussion on two mechanisms: HBDD and NDR, that could possibly generate an observable rhythmic behavior, we proposed an experimental approach to determine the correct mechanism behind a specific observed rhythmic behavior of interests, and applied the approach to actual experimental data. With the limited sample size and control on experimental conditions, we managed to obtain promising result from *in vitro* hes1 data (NIH 3T3 cell clones with stable NICD expression maintained by Tet-On system), where we demonstrated that the observed hes1 rhythmic behavior is driven by Hopf bifurcation induced limit cycle in its deterministic dynamics. In the last part of the thesis, we presented our exploration into the data that depicts the neuronal rest-state of the zebrafish habenula, in search of the possible presence of a chaotic attractor or critical attractor. We confirmed that there is no observable low dimensional chaotic attractor in the system. Nonetheless, we found promising leads that suggest the presence of critical behavior in the habenula dynamics.

7 References

- [1] Elizabeth A Slee, Daniel J O'Connor, and Xin Lu. To die or not to die: how does p53 decide? *Oncogene*, 23(16):2809–2818, apr 2004.
- [2] A B DeLeo, G Jay, E Appella, G C Dubois, L W Law, and Lloyd J Old. Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proceedings of the National Academy of Sciences*, 76(5):2420–2424, may 1979.
- [3] Jack T. Zilfou and Scott W. Lowe. Tumor Suppressive Functions of p53. *Cold Spring Harbor Perspectives in Biology*, 1(5):a001883–a001883, nov 2009.
- [4] Todd Riley, Eduardo Sontag, Patricia Chen, and Arnold Levine. Transcriptional control of human p53-regulated genes. *Nature Reviews Molecular Cell Biology*, 9(5):402–412, may 2008.
- [5] George Farmer, John Colgan, Yoshiro Nakatani, James L Manley, and Carol Prives. Functional interaction between p53, the TATA-binding protein (TBP), and TBP-associated factors in vivo. *Molecular and Cellular Biology*, 16(8):4295–4304, aug 1996.
- [6] Vishwanie Budhram-Mahadeo, Peter J. Morris, Martin D. Smith, Carol A. Midgley, Linda M. Boxer, and David S. Latchman. p53 Suppresses the Activation of the Bcl-2 Promoter by the Brn-3a POU Family Transcription Factor. *Journal of Biological Chemistry*, 274(21):15237–15244, may 1999.
- [7] K C Lee, a J Crowe, and M C Barton. p53-mediated repression of alpha-fetoprotein gene expression by specific DNA binding. *Molecular and cellular biology*, 19(2):1279–88, feb 1999.
- [8] Claes Ohlsson, Nikolai Kley, Haim Werner, and Derek LeRoith. p53 Regulates Insulin-Like Growth Factor-I (IGF-I) Receptor Expression and IGF-I-Induced

Tyrosine Phosphorylation in an Osteosarcoma Cell Line: Interaction between p53 and Sp1. *Endocrinology*, 139(3):1101–1107, mar 1998.

- [9] Angelina V. Vaseva and Ute M. Moll. The mitochondrial p53 pathway. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1787(5):414–420, may 2009.
- [10] Amato J. Giaccia and Michael B. Kastan. The complexity of p53 modulation: emerging patterns from divergent signals. *Genes & Development*, 12(19):2973–2983, oct 1998.
- [11] M F Lavin and N Gueven. The complexity of p53 stabilization and activation. *Cell death and differentiation*, 13(6):941–950, jun 2006.
- [12] Christopher L Brooks and Wei Gu. p53 Ubiquitination: Mdm2 and Beyond. *Molecular Cell*, 21(3):307–315, feb 2006.
- [13] Sathyavageeswaran Shreeram, Oleg N. Demidov, Weng Kee Hee, Hiroshi Yamaguchi, Nobuyuki Onishi, Calvin Kek, Oleg N. Timofeev, Crissy Dudgeon, Albert J. Fornace, Carl W. Anderson, Yasuhiro Minami, Ettore Appella, and Dmitry V. Bulavin. Wip1 Phosphatase Modulates ATM-Dependent Signaling Pathways. *Molecular Cell*, 23(5):757–764, sep 2006.
- [14] Susan Erster, Motohiro Mihara, Roger H Kim, Oleksi Petrenko, and Ute M Moll. In vivo mitochondrial p53 translocation triggers a rapid first wave of cell death in response to DNA damage that can precede p53 target gene activation. *Molecular and Cellular Biology*, 24(15):6728–6741, jul 2004.
- [15] Mengqiu Wu, Hui Ye, Zhiyuan Tang, Chang Shao, Gaoyuan Lu, Baoqiang Chen, Yuyu Yang, Guangji Wang, and Haiping Hao. p53 dynamics orchestrates with binding affinity to target genes for cell fate decision. *Cell Death and Disease*, 8(10):e3130, oct 2017.

- [16] Lan Ma, John Wagner, John Jeremy Rice, Wenwei Hu, Arnold J Levine, and Gustavo A Stolovitzky. A plausible model for the digital response of p53 to DNA damage. *Proceedings of the National Academy of Sciences*, 102(40):14266–14271, oct 2005.
- [17] Adeela Kamal and Lawrence SB Goldstein. Connecting vesicle transport to the cytoskeleton. *Current Opinion in Cell Biology*, 12(4):503–508, aug 2000.
- [18] R. John Ellis. Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences*, 26(10):597–604, 2001.
- [19] Hoda Eydgahi, William W Chen, Jeremy L Muhlich, Dennis Vitkup, John N Tsitsiklis, and Peter K Sorger. Properties of cell death models calibrated and compared using Bayesian approaches. *Molecular Systems Biology*, 9(1):644–644, apr 2014.
- [20] Ryan N. Gutenkunst, Joshua J. Waterfall, Fergal P. Casey, Kevin S. Brown, Christopher R. Myers, and James P. Sethna. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Computational Biology*, 3(10):e189, 2007.
- [21] Marc Sturrock, Alan J. Terry, Dimitris P. Xirodimas, Alastair M. Thompson, and Mark A J Chaplain. Influence of the Nuclear Membrane, Active Transport, and Cell Shape on the Hes1 and p53–Mdm2 Pathways: Insights from Spatio-temporal Modelling. *Bulletin of Mathematical Biology*, 74(7):1531–1579, jul 2012.
- [22] Jae Kyoung Kim and Trachette L Jackson. Mechanisms That Enhance Sustainability of p53 Pulses. *PLoS ONE*, 8(6):e65242, jun 2013.
- [23] Bing Liu, Divesh Bhatt, Zoltán N. Oltvai, Joel S Greenberger, and Ivet Bahar. Significance of p53 dynamics in regulating apoptosis in response to ionizing

radiation, and polypharmacological strategies. *Scientific Reports*, 4:6245, sep 2014.

- [24] Ján Eliaš, Luna Dimitrio, Jean Clairambault, and Roberto Natalini. The p53 protein and its molecular network: Modelling a missing link between DNA damage and cell fate. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(1):232–247, jan 2014.
- [25] Wassim Abou-Jaoudé, Djomangan A Ouattara, and Marcelle Kaufman. From structure to dynamics: Frequency tuning in the p53-Mdm2 network I. Logical approach. *Journal of Theoretical Biology*, 258(4):561–577, 2009.
- [26] Djomangan a. Ouattara, Wassim Abou-Jaoudé, and Marcelle Kaufman. From structure to dynamics: Frequency tuning in the p53-Mdm2 network. II. Differential and stochastic approaches. *Journal of Theoretical Biology*, 264(4):1177–1189, jun 2010.
- [27] Wassim Abou-Jaoudé, Madalena Chaves, and Jean-Luc Gouze. A theoretical exploration of birhythmicity in the p53-Mdm2 network. *PLoS ONE*, 6(2):e17075, 2011.
- [28] Minsoo Choi, Jue Shi, Sung Hoon Jung, Xi Chen, and Kwang-Hyun Cho. Attractor landscape analysis reveals feedback loops in the p53 network that control the cellular response to DNA damage. *Science Signaling*, 5(251):ra83, nov 2012.
- [29] Jonathan M Raser. Noise in Gene Expression: Origins, Consequences, and Control. *Science*, 309(5743):2010–2013, sep 2005.
- [30] Vahid Shahrezaei and Peter S. Swain. The stochastic nature of biochemical networks. *Current Opinion in Biotechnology*, 19(4):369–374, aug 2008.
- [31] Lev S. Tsimring. Noise in biology. *Reports on Progress in Physics*, 77(2), 2014.

- [32] Quincey A. Justman. An Explicit Source for Extrinsic Noise. *Cell Systems*, 1(5):308–309, nov 2015.
- [33] Marzo Forment and Guillermo Rodrigo. Molecular noise can minimize the collective sensitivity of a clonal heterogeneous cell population. *Journal of Theoretical Biology*, 416(September 2016):38–44, 2017.
- [34] I Herskowitz and D Hagen. The Lysis-Lysogeny Decision of Phage lambda: Explicit Programming and Responsiveness. *Annual Review of Genetics*, 14(1):399–445, dec 1980.
- [35] J Hasty, J Pradines, M Dolnik, and J J Collins. Noise-based switches and amplifiers for gene expression. *Proceedings of the National Academy of Sciences*, 97(5):2075–2080, feb 2000.
- [36] Ertugrul M Ozbudak, Mukund Thattai, Iren Kurtser, Alan D Grossman, and Alexander van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature genetics*, 31(1):69–73, 2002.
- [37] Marc Sturrock, Andreas Hellander, Anastasios Matzavinos, and Mark A J Chaplain. Spatial stochastic modelling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. *Journal of The Royal Society Interface*, 10(80):20120988–20120988, jan 2013.
- [38] Robert Ahrends, Asuka Ota, Kyle M. Kovary, Takamasa Kudo, Byung Ouk Park, and Mary N. Teruel. Controlling low rates of cell differentiation through noise and ultrahigh feedback. *Science*, 344(6190):1384–1389, jun 2014.
- [39] Michael Samoilov, Sergey Plyasunov, and Adam P Arkin. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences*, 102(7):2310–2315, feb 2005.

- [40] Jingkui Wang, Marc Lefranc, and Quentin Thommen. Stochastic Oscillations Induced by Intrinsic Fluctuations in a Self-Repressing Gene. *Biophysical Journal*, 107(10):2403–2416, nov 2014.
- [41] A L HODGKIN and A F HUXLEY. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500–44, aug 1952.
- [42] Urs Gerber. Metabotropic glutamate receptors in vertebrate retina. *Documenta ophthalmologica. Advances in ophthalmology*, 106(1):83–7, jan 2003.
- [43] J.-P. Pin and R. Duvoisin. The metabotropic glutamate receptors: Structure and functions. *Neuropharmacology*, 34(1):1–26, jan 1995.
- [44] Jean-Marc Fritschy and Hanns Mohler. GABAA-receptor heterogeneity in the adult rat brain: Differential regional and cellular distribution of seven major subunits. *The Journal of Comparative Neurology*, 359(1):154–194, aug 1995.
- [45] Erwin Sigel, F Anne Stephenson, Cleanthi Mamalaki, and Eric A Barnard. A gamma-aminobutyric acid/benzodiazepine receptor complex of bovine cerebral cortex. *The Journal of biological chemistry*, 258(11):6965–71, jun 1983.
- [46] David Mott. The metabotropic GABAB receptors. In *Cellular and Molecular Neurophysiology (Fourth edition)*, chapter 11, pages 245–267. Academic Press, 2015.
- [47] Marianna Leonzino, Marta Busnelli, Flavia Antonucci, Claudia Verderio, Michele Mazzanti, and Bice Chini. The Timing of the Excitatory-to-Inhibitory GABA Switch Is Regulated by the Oxytocin Receptor via KCC2. *Cell Reports*, 15(1):96–103, apr 2016.
- [48] Sheriar G. Hormuzdi, Mikhail A. Filippov, Georgia Mitropoulou, Hannah Monyer, and Roberto Bruzzone. Electrical synapses: a dynamic signaling system

that shapes the activity of neuronal networks. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1662(1-2):113–137, mar 2004.

- [49] Michael V.L. Bennett and R.Suzanne Zukin. Electrical Coupling and Neuronal Synchronization in the Mammalian Brain. *Neuron*, 41(4):495–511, feb 2004.
- [50] Igor Belykh, Enno De Lange, and Martin Hasler. Synchronization of bursting neurons: What matters in the network topology. *Physical Review Letters*, 94(18):1–4, 2005.
- [51] Christian Meisel and Thilo Gross. Adaptive self-organization in a realistic neural network model. *Physical Review E*, 80(6):061917, dec 2009.
- [52] F. Lombardi, H. J. Herrmann, and L. de Arcangelis. Balance of excitation and inhibition determines 1/f power spectrum in neuronal networks. *Chaos*, 27(4), 2017.
- [53] Attila Szücs, Pablo Varona, Alexander R. Volkovskii, Henry D. I. Abarbanel, Mikhail I. Rabinovich, and Allen I. Selverston. Interacting biological and electronic neurons generate realistic oscillatory rhythms. *NeuroReport*, 11(3):563–569, feb 2000.
- [54] Mikhail Rabinovich, Ramón Huerta, Maxim Bazhenov, Alexander K. Kozlov, and Henry D. I. Abarbanel. Computer simulations of stimulus dependent state switching in basic circuits of bursting neurons. *Physical Review E*, 58(5):6418–6430, nov 1998.
- [55] Alain Destexhe, Michael Rudolph, and Denis Paré. The high-conductance state of neocortical neurons in vivo. *Nature Reviews Neuroscience*, 4(9):739–751, sep 2003.

- [56] Michael D. Fox and Marcus E. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–711, sep 2007.
- [57] Christine Grienberger and Arthur Konnerth. Imaging Calcium in Neurons. *Neuron*, 73(5):862–885, mar 2012.
- [58] J. Akerboom, T.-W. Chen, T. J. Wardill, L. Tian, J. S. Marvin, S. Mutlu, N. C. Calderon, F. Esposti, B. G. Borghuis, X. R. Sun, A. Gordus, M. B. Orger, R. Portugues, F. Engert, J. J. Macklin, A. Filosa, A. Aggarwal, R. A. Kerr, R. Takagi, S. Kracun, E. Shigetomi, B. S. Khakh, H. Baier, L. Lagnado, S. S.-H. Wang, C. I. Bargmann, B. E. Kimmel, V. Jayaraman, K. Svoboda, D. S. Kim, E. R. Schreiter, and L. L. Looger. Optimization of a GCaMP Calcium Indicator for Neural Activity Imaging. *Journal of Neuroscience*, 32(40):13819–13840, 2012.
- [59] Aleksandra Badura, Xiaonan Richard Sun, Andrea Giovannucci, Laura A. Lynch, and Samuel S.-H. Wang. Fast calcium sensor proteins for monitoring neural activity. *Neurophotonics*, 1(2):025008, 2014.
- [60] Nordine Helassa, Borbala Podor, Alan Fine, and Katalin Török. Design and mechanistic insight into ultrafast calcium indicators for monitoring intracellular calcium dynamics. *Scientific Reports*, 6(1):38276, dec 2016.
- [61] R L Cowan and C J Wilson. Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology*, 71(1):17–32, jan 1994.
- [62] M. Hausser. The Beat Goes On: Spontaneous Firing in Mammalian Neuronal Microcircuits. *Journal of Neuroscience*, 24(42):9215–9219, oct 2004.
- [63] Carson C. Chow and John A. White. Spontaneous action potentials due to channel fluctuations. *Biophysical Journal*, 71(6):3013–3021, 1996.

- [64] C. v. Vreeswijk, H. Sompolinsky, C van Vreeswijk, and H. Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science (New York, N.Y.)*, 274(5293):1724–6, dec 1996.
- [65] C van Vreeswijk and H Sompolinsky. Chaotic balanced state in a model of cortical circuits. *Neural computation*, 10(6):1321–71, aug 1998.
- [66] Miloslav Nič, Jiří Jiráť, Bedřich Košata, Aubrey Jenkins, and Alan McNaught, editors. *IUPAC Compendium of Chemical Terminology*. IUPAC, Research Triangle Park, NC, jun 2009.
- [67] Zeynep Ökten, L. Stirling Churchman, Ronald S. Rock, and James A. Spudich. Myosin VI walks hand-over-hand along actin. *Nature Structural & Molecular Biology*, 11(9):884–887, sep 2004.
- [68] Luna Dimitrio, Jean Clairambault, and Roberto Natalini. A spatial physiological model for p53 intracellular dynamics. *Journal of Theoretical Biology*, 316:9–24, jan 2013.
- [69] D T Gillespie. Exact Stochastic Simulation of Coupled Chemical-Reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [70] Jae Kyoung Kim, Krešimir Josić, and Matthew R. Bennett. The Validity of Quasi-Steady-State Approximations in Discrete Stochastic Simulations. *Biophysical Journal*, 107(3):783–793, aug 2014.
- [71] Philipp Thomas, Arthur V Straube, and Ramon Grima. The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC Systems Biology*, 6(1):39, 2012.
- [72] Daniel T Gillespie. The chemical Langevin equation. *The Journal of Chemical Physics*, 113(1):297–306, jul 2000.

- [73] D T Gillespie. *Markov processes: an introduction for physical scientists*. Elsevier, 1991.
- [74] P Smadbeck and Y N Kaznessis. Efficient moment matrix generation for arbitrary chemical networks. *Chemical Engineering Science*, 84:612–618, dec 2012.
- [75] Patrick Smadbeck and Yiannis N Kaznessis. A closure scheme for chemical master equations. *Proceedings of the National Academy of Sciences*, 110(35):14261–14265, aug 2013.
- [76] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, nov 1985.
- [77] Galit Lahav, Nitzan Rosenfeld, Alex Sigal, Naama Geva-Zatorsky, Arnold J Levine, Michael B Elowitz, and Uri Alon. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature Genetics*, 36(2):147–150, feb 2004.
- [78] Yoshito Masamizu, Toshiyuki Ohtsuka, Yoshiki Takashima, Hiroki Nagahara, Yoshiko Takenaka, Kenichi Yoshikawa, Hitoshi Okamura, and Ryoichiro Kageyama. Real-time imaging of the somite segmentation clock: Revelation of unstable oscillators in the individual presomitic mesoderm cells. *Proceedings of the National Academy of Sciences*, 103(5):1313–1318, jan 2006.
- [79] Itaru Imayoshi, Akihiro Isomura, Yukiko Harima, Kyogo Kawaguchi, Hiroshi Kori, Hitoshi Miyachi, Takahiro Fujiwara, Fumiyoshi Ishidate, and Ryoichiro Kageyama. Oscillatory Control of Factors Determining Multipotency and Fate in Mouse Neural Progenitors. *Science*, 342(6163):1203–1208, dec 2013.
- [80] Yukiko Harima, Itaru Imayoshi, Hiromi Shimojo, Taeko Kobayashi, and Ryoichiro Kageyama. The roles and mechanism of ultradian oscillatory expression of the mouse Hes genes. *Seminars in Cell & Developmental Biology*, 34:85–90, oct 2014.

- [81] Marcos Malumbres and Mariano Barbacid. Cell cycle, CDKs and cancer: a changing paradigm. *Nature Reviews Cancer*, 9(3):153–166, mar 2009.
- [82] Naama Geva-Zatorsky, Nitzan Rosenfeld, Shalev Itzkovitz, Ron Milo, Alex Sigal, Erez Dekel, Talia Yarnitzky, Yuvalal Liron, Paz Polak, Galit Lahav, and Uri Alon. Oscillations and variability in the p53 system. *Molecular Systems Biology*, 2:2006.0033, jun 2006.
- [83] Savaş Tay, Jacob J Hughey, Timothy K Lee, Tomasz Lipniacki, Stephen R Quake, and Markus W Covert. Single-cell NF- κ B dynamics reveal digital activation and analogue information processing. *Nature*, 466(7303):267–271, jul 2010.
- [84] Carole J Proctor and Douglas A Gray. Explaining oscillations and variability in the p53-Mdm2 system. *BMC Systems Biology*, 2(1):75, 2008.
- [85] Ruiqi Wang, Kaihui Liu, Luonan Chen, and Kazuyuki Aihara. Neural fate decisions mediated by trans-activation and cis-inhibition in Notch signaling. *Bioinformatics*, 27(22):3158–3165, nov 2011.
- [86] Pierre-Emmanuel Emmanuel Morant, Quentin Thommen, François Lemaire, Constant Vandermoëre, Benjamin Parent, and Marc Lefranc. Oscillations in the Expression of a Self-Repressed Gene Induced by a Slow Transcriptional Dynamics. *Physical Review Letters*, 102(6):068104, feb 2009.
- [87] Naama Geva-Zatorsky, Erez Dekel, Eric Batchelor, Galit Lahav, and Uri Alon. Fourier analysis and systems identification of the p53 feedback loop. *Proceedings of the National Academy of Sciences*, 107(30):13550–13555, jul 2010.
- [88] F. Krombach, S. Münzing, A. M. Allmeling, J. T. Gerlach, J. Behr, and M. Dörger. Cell size of alveolar macrophages: an interspecies comparison. *Environmental health perspectives*, 105 Suppl(9):1261–1263, sep 1997.

- [89] Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. CRC Press, 2014.
- [90] Akshit Arora, Saurav Gera, Tanuj Maheshwari, Dhvani Raghav, Md Jahoor Alam, R K Brojen Singh, and Subhash M Agarwal. The Dynamics of Stress p53-Mdm2 Network Regulated by p300 and HDAC1. *PLoS ONE*, 8(2):e52736, feb 2013.
- [91] Chong He, Keng-Hwee Chiam, and Lock Yue Chew. Comparison of cellular oscillations driven by noise or deterministic mechanisms under cell-size scaling. *Physical Review E*, 94(4):042425, oct 2016.
- [92] Moritz Lang, Steffen Waldherr, and Frank Allgöwer. Amplitude distribution of stochastic oscillations in biochemical networks due to intrinsic noise. *PMC Biophysics*, 2(1):10, 2009.
- [93] Clare H McGowan and Paul Russell. The DNA damage response: sensing and signaling. *Current Opinion in Cell Biology*, 16(6):629–633, dec 2004.
- [94] X.-P. Zhang, Feng Liu, Zhang Cheng, and Wei Wang. Cell fate decision mediated by p53 pulses. *Proceedings of the National Academy of Sciences*, 106(30):12245–12250, jul 2009.
- [95] Jeremy E Purvis, Kyle W Karhohs, Caroline Mock, Eric Batchelor, Alexander Loewer, and Galit Lahav. p53 Dynamics Control Cell Fate. *Science*, 336(6087):1440–1444, jun 2012.
- [96] Miguel Maroto, J Kim Dale, Mary-Lee Dequeant, Anne-Cecile Petit, and Olivier Pourquie. Synchronised cycling gene oscillations in presomitic mesoderm cells require cell-cell contact. *The International Journal of Developmental Biology*, 49(2-3):309–315, 2005.
- [97] Mark E. Fortini. Notch Signaling: The Core Pathway and Its Posttranslational Regulation. *Developmental Cell*, 16(5):633–647, may 2009.

- [98] David del Álamo, Hervé Rouault, and François François Schweisguth. Mechanism and significance of cis-inhibition in notch signalling. *Current Biology*, 21(1):40–47, jan 2011.
- [99] Mikael S Lindström, Aiwen Jin, Chad Deisenroth, Gabrielle White Wolf, Yanping Zhang, M. S. Lindstrom, Aiwen Jin, Chad Deisenroth, Gabrielle White Wolf, and Yanping Zhang. Cancer-associated mutations in the MDM2 zinc finger domain disrupt ribosomal protein interaction and attenuate MDM2-induced p53 degradation. *Molecular and cellular biology*, 27(3):1056–1068, feb 2007.
- [100] Hiromi Hirata, Shigeki Yoshiura, Toshiyuki Ohtsuka, Yasumasa Bessho, Takahiro Harada, Kenichi Yoshikawa, and Ryoichiro Kageyama. Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop. *Science (New York, N.Y.)*, 298(5594):840–843, oct 2002.
- [101] S Ramsey, A Ozinsky, A Clark, K D Smith, P de Atauri, V Thorsson, D Orrell, and H Bolouri. Transcriptional noise and cellular heterogeneity in mammalian macrophages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1467):495–506, mar 2006.
- [102] Manuel Barrio, Kevin Burrage, André Leier, and Tianhai Tian. Oscillatory Regulation of Hes1: Discrete Stochastic Delay Modelling and Simulation. *PLoS Computational Biology*, 2(9):e117, sep 2006.
- [103] David B. Young, Jyoti Jonnalagadda, Magtouf Gatei, David a. Jans, Stephen Meyn, and Kum Kum Khanna. Identification of domains of ataxia-telangiectasia mutated required for nuclear localization and chromatin association. *Journal of Biological Chemistry*, 280(30):27587–27594, jul 2005.
- [104] D P Teufel, M Bycroft, and A R Fersht. Regulation by phosphorylation of the relative affinities of the N-terminal transactivation domains of p53 for p300 domains and Mdm2. *Oncogene*, 28(20):2112–2118, may 2009.

- [105] Franck Toledo and Geoffrey M Wahl. Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. *Nature reviews. Cancer*, 6(12):909–923, dec 2006.
- [106] A Barton and A.J. Fendrik. Sustained vs. oscillating expressions of Ngn2, Dll1 and Hes1: A model of neural differentiation of embryonic telencephalon. *Journal of Theoretical Biology*, 328:1–8, jul 2013.
- [107] JAMES D. RUSSELL, WILLIAM S. WITT, and James D Russell. CELL SIZE AND GROWTH CHARACTERISTICS OF CULTURED FIBROBLASTS ISOLATED FROM NORMAL AND KELOID TISSUE. *Plastic and Reconstructive Surgery*, 57(2):207–212, feb 1976.
- [108] Y.-H. M. Chan and Wallace F. Marshall. How Cells Know the Size of Their Organelles. *Science*, 337(6099):1186–1189, sep 2012.
- [109] A. S. Levenson and V C Jordan. MCF-7: the first hormone-responsive breast cancer cell line. *Cancer research*, 57(15):3071–8, aug 1997.
- [110] S. Cos, J. Recio, and E.J. Sánchez-Barceló. Modulation of the length of the cell cycle time of MCF-7 human breast cancer cells by melatonin. *Life Sciences*, 58(9):811–816, jan 1996.
- [111] Diane Longo and Jeff Hasty. Dynamics of single-cell gene expression. *Molecular Systems Biology*, 2, nov 2006.
- [112] G J TODARO and H GREEN. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *The Journal of cell biology*, 17(6):299–313, may 1963.
- [113] M Hitomi and D W Stacey. Cellular ras and cyclin D1 are required during different cell cycle periods in cycling NIH 3T3 cells. *Molecular and cellular biology*, 19(7):4623–32, jul 1999.

- [114] M. E. Raichle and D. A. Gusnard. Appraising the brain’s energy budget. *Proceedings of the National Academy of Sciences*, 99(16):10237–10239, aug 2002.
- [115] Louis Sokoloff, Renward Mangold, Richard L. Wechsler, Charles Kennedy, and Seymour S. Kety. THE EFFECT OF MENTAL ARITHMETIC ON CERE-BRAL CIRCULATION AND METABOLISM 1. *Journal of Clinical Investiga-tion*, 34(7 Pt 1):1101–1108, jul 1955.
- [116] E. Novikov, A. Novikov, D. Shannahoff-Khalsa, B. Schwartz, and J. Wright. Scale-similar activity in the brain. *Physical Review E*, 56(3):R2387–R2389, sep 1997.
- [117] P. A. Robinson, C. J. Rennie, J. J. Wright, H. Bahramali, E. Gordon, and D. L. Rowe. Prediction of electroencephalographic spectra from neurophysiology. *Physical Review E*, 63(2):021903, jan 2001.
- [118] Garrett T. Neske. The Slow Oscillation in Cortical and Thalamic Networks: Mechanisms and Functions. *Frontiers in Neural Circuits*, 9(January):1–25, 2016.
- [119] Woodrow L. Shew, Hongdian Yang, Thomas Petermann, Rajarshi Roy, and Di-etmar Plenz. Neuronal Avalanches Imply Maximum Dynamic Range in Cortical Networks at Criticality. *Journal of Neuroscience*, 29(49):15595–15600, dec 2009.
- [120] L. de Arcangelis and H. J. Herrmann. Learning as a phenomenon occurring in a critical state. *Proceedings of the National Academy of Sciences*, 107(9):3977–3981, 2010.
- [121] Daniel Millman, Stefan Mihalas, Alfredo Kirkwood, and Ernst Niebur. Self-organized criticality occurs in non-conservative neuronal networks during ‘up’ states. *Nature Physics*, 6(10):801–805, oct 2010.

- [122] W. L. Shew, H. Yang, S. Yu, R. Roy, and D. Plenz. Information Capacity and Transmission Are Maximized in Balanced Cortical Networks with Neuronal Avalanches. *Journal of Neuroscience*, 31(1):55–63, jan 2011.
- [123] C Bédard, H Kröger, and Alain Destexhe. Does the 1/f frequency scaling of brain signals reflect self-organized critical states? *Physical review letters*, 97(11):118102, sep 2006.
- [124] Julián Yáñez and Ramón Anadón. Afferent and efferent connections of the habenula in the larval sea lamprey (*Petromyzon marinus* L.): An experimental study. *Journal of Comparative Neurology*, 345(1):148–160, jul 1994.
- [125] Hitoshi Okamoto, Masakazu Agetsuma, and Hidenori Aizawa. Genetic dissection of the zebrafish habenula, a possible switching board for selection of behavioral strategy to cope with fear and anxiety. *Developmental Neurobiology*, 72(3):386–394, mar 2012.
- [126] Katherine J. Turner, Thomas A. Hawkins, Julián Yáñez, Ramón Anadón, Stephen W. Wilson, and Mónica Folgueira. Afferent Connectivity of the Zebrafish Habenulae. *Frontiers in Neural Circuits*, 10(April):1–18, 2016.
- [127] Okihide Hikosaka. The habenula: from stress evasion to value-based decision-making. *Nature reviews. Neuroscience*, 11(7):503–513, 2010.
- [128] Bai bing Zhang, Yuan yuan Yao, He fei Zhang, Koichi Kawakami, and Jiu lin Du. Left Habenula Mediates Light-Preference Behavior in Zebrafish via an Asymmetrical Visual Pathway. *Neuron*, 93(4):914–928.e4, 2017.
- [129] Suresh Jesuthasan. The thalamo-habenula projection revisited. *Seminars in Cell & Developmental Biology*, pages 8–11, aug 2017.

- [130] Lucas Lecourtier and Peter H. Kelly. A conductor hidden in the orchestra? Role of the habenular complex in monoamine transmission and cognition. *Neuroscience & Biobehavioral Reviews*, 31(5):658–672, jan 2007.
- [131] Patrick Kaifosh, Jeffrey D Zaremba, Nathan B Danielson, and Attila Losonczy. SIMA: Python software for analysis of dynamic fluorescence imaging data. *Frontiers in neuroinformatics*, 8(September):80, 2014.
- [132] Eftychios A. Pnevmatikakis, Yuanjun Gao, Daniel Soudry, David Pfau, Clay Lacefield, Kira Poskanzer, Randy Bruno, Rafael Yuste, and Liam Paninski. A structured matrix factorization framework for large scale calcium imaging data analysis. *arXiv*, 1409.2903:1–21, 2014.
- [133] Eftychios A. Pnevmatikakis and Andrea Giovannucci. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods*, 291:83–94, 2017.
- [134] Manuel Guizar-Sicairos, Samuel T. Thurman, and James R. Fienup. Efficient subpixel image registration algorithms. *Optics Letters*, 33(2):156, 2008.
- [135] LYLE NOAKES. THE TAKENS EMBEDDING THEOREM. *International Journal of Bifurcation and Chaos*, 01(04):867–872, dec 1991.
- [136] Holger Kantz and Thomas Schreiber. *Nonlinear Time Series Analysis*, 2004.
- [137] Pierre Gaspard and Xiao-Jing Wang. Noise, chaos, and (ϵ, τ) -entropy per unit time. *Physics Reports*, 235(6):291–343, dec 1993.
- [138] Sophie Denève and Christian K Machens. Efficient codes and balanced networks. *Nature Neuroscience*, 19(3):375–382, feb 2016.

Appendix

A Reaction networks

A.1 Brusselator

Both the NDR and HBDD oscillation scheme can be obtained from Brusselator with different parameter set. Our HBDD oscillation model is chosen with a parameter set that gives a limit-cycle. Our NDR oscillation model is chosen with a parameter set on the edge of the Hopf bifurcation. Table. 2 gives the detail of the parameter used in our simulations.

Index	Reaction Type	Reaction	NDR Reaction Rates	HBDD Reaction Rates
R_1	order-0	$\phi \longrightarrow x$	$k_{R1} = 79.27$	$k_{R1} = 141.4$
R_2	order- 3	$2x + y \longrightarrow 3x$	$k_{R2} = 0.0001$	$k_{R2} = 0.0001$
R_3	order-1	$x \longrightarrow \phi$	$k_{R3} = 1.0$	$k_{R3} = 1.0$
R_4	order-1	$x \longrightarrow y$	$k_{R4} = 1.395$	$k_{R4} = 8.0$

Table 2. Brusselator model. The parameters are at the scale of original size.

A.2 p53 Oscillatory Network: HBDD network

We have translated Dimitrio’s ODE model⁶⁸ into a reaction network as our HBDD model. Since the oscillation frequency of their model is too small, we have scaled all reaction rates with a ratio $\eta_t = 12.586$ to make the oscillation period 5.5 h. The reaction network, along with their rate constants, is given in Table. 4. Other than the reactants that stay constant in the network (as shown in Table. 3), the initial conditions for all variables are zero. The unit for the amount of reactants is the number of molecules. The unit of time is minute.

Reactant Name	ATM_c	ATM_n	$wip1$	DNA
Initial value(molecules)	558696	55870	558696	2

Table 3. Initial value for fixed reactant. ATM value here is 30% of the maximum ATM level ($8.09 \mu\text{M}$) used by Dimitrio⁶⁸ within the Hopf bifurcation regime.

Index	Reaction Type	Reaction	Reaction Rates
R_{c1}	MM kinetics	$p53_n + mdm2_n \rightleftharpoons p53_n \cdot mdm2_n \longrightarrow mdm2_n$	k = 7.95e+00 ; K = 1.88e+05
R_{c2}	MM kinetics	$p53_n + ATM_n \rightleftharpoons p53_n \cdot ATM_n \longrightarrow ATM_n + p53p_n$	k = 7.95e-02 ; K = 1.86e+04
R_{c3}	MM kinetics	$p53_c + ATM_c \rightleftharpoons p53_c \cdot ATM_c \longrightarrow ATM_c + p53p_c$	k = 7.95e-02 ; K = 1.86e+05
R_{c4}	MM kinetics	$p53p_n + wip1 \rightleftharpoons p53p_n \cdot wip1 \longrightarrow wip1 + p53_n$	k = 2.65e-03 ; K = 9.31e+03
R_{c5}	MM kinetics	$p53_c + mdm2_c \rightleftharpoons p53_c \cdot mdm2_c \longrightarrow mdm2_c$	k = 7.95e+00 ; K = 1.88e+06
R_{p1}	order-0	$\phi \longrightarrow p53_c$	k = 2.22e+03
R_{p2}	order-1	$p53_c \longrightarrow \phi$	k = 1.59e-02
R_{p3}	order-1	$DNA \longrightarrow DNA + RNAmdm2_n$	k = 3.70e+01
R_{p4}	Hill equation	$DNA + 4p53p_n \rightleftharpoons DNA \cdot 4p53p_n \longrightarrow DNA + RNAmdm2_n + 4p53p_n$	k = 7.40e+03 ; K = 1.86e+04 ; H = 4.0
R_{p5}	order-1	$RNAmdm2_c \longrightarrow mdm2_c$	k = 7.95e-02
R_{p6}	order-1	$RNAmdm2_n \longrightarrow \phi$	k = 7.95e-06
R_{p7}	order-1	$mdm2_n \longrightarrow \phi$	k = 1.27e-02
R_{p8}	order-1	$RNAmdm2_c \longrightarrow \phi$	k = 7.95e-06
R_{p9}	order-1	$mdm2_c \longrightarrow \phi$	k = 1.27e-02
R_{t1}	order-1	$p53_n \longrightarrow p53_c$	k = 6.59e-02
R_{t2}	order-1	$RNAmdm2_n \longrightarrow RNAmdm2_c$	k = 6.59e-02
R_{t3}	order-1	$mdm2_n \longrightarrow mdm2_c$	k = 3.18e-02
R_{t4}	order-1	$p53_c \longrightarrow p53_n$	k = 6.59e-03
R_{t5}	order-1	$p53p_c \longrightarrow p53p_n$	k = 6.59e-03
R_{t6}	order-1	$mdm2_c \longrightarrow mdm2_n$	k = 3.18e-03

Table 4. The set of reactions for p53-NDR model

A.3 p53 Oscillatory Network: NDR network

This model is built based on the HBDD model. We have removed ATM from cytoplasm, as described in Section 3.4. The reaction network, along with their rate constants, is given in Table. 6. Other than the reactants that stay constant in the network (as shown in Table. 5), the initial conditions for all variables are zero. The unit for the amount of reactants is the number of molecules. The unit of time is minute.

Reactant Name	<i>ATM</i>	<i>wip1</i>	<i>DNA</i>
Initial value(molecules)	1862320	558696	2

Table 5. Initial value for fixed reactant. *ATM* value here (equivalent to $8.09\mu\text{M}$) is the maximum level of *ATM* used by Dimitrio *et al.*⁶⁸ In our stochastic simulation, we have used 30% of this amount of molecules for a stressed cell.

Index	Reaction Type	Reaction	Reaction Rates
R_{c1}	MM kinetics	$p53_n + mdm2_n \rightleftharpoons p53_n \cdot mdm2_n \longrightarrow mdm2_n$	k = 7.95e+00 ; K = 1.88e+05
R_{c2}	MM kinetics	$p53_n + ATM \rightleftharpoons p53_n \cdot ATM \longrightarrow ATM + p53p_n$	k = 7.95e-02 ; K = 1.86e+04
R_{c3}	MM kinetics	$p53p_n + wip1 \rightleftharpoons p53p_n \cdot wip1 \longrightarrow wip1 + p53_n$	k = 2.65e-03 ; K = 9.31e+03
R_{c4}	MM kinetics	$p53_c + mdm2_c \rightleftharpoons p53_c \cdot mdm2_c \longrightarrow mdm2_c$	k = 7.95e+00 ; K = 1.88e+06
R_{p1}	order-0	$\phi \longrightarrow p53_c$	k = 2.22e+03
R_{p2}	order-1	$p53_c \longrightarrow \phi$	k = 1.59e-02
R_{p3}	order-1	$DNA \longrightarrow DNA + RNAmdm2_n$	k = 3.70e+01
R_{p4}	Hill equation	$DNA + 4p53p_n \rightleftharpoons DNA \cdot 4p53p_n \longrightarrow DNA + RNAmdm2_n + 4p53p_n$	k = 7.40e+03 ; K = 1.86e+04 ; H = 4.0
R_{p5}	order-1	$RNAmdm2_c \longrightarrow mdm2_c$	k = 7.95e-02
R_{p6}	order-1	$RNAmdm2_n \longrightarrow \phi$	k = 7.95e-06
R_{p7}	order-1	$mdm2_n \longrightarrow \phi$	k = 1.27e-02
R_{p8}	order-1	$RNAmdm2_c \longrightarrow \phi$	k = 7.95e-06
R_{p9}	order-1	$mdm2_c \longrightarrow \phi$	k = 1.27e-02
R_{t1}	order-1	$p53_n \longrightarrow p53_c$	k = 6.59e-02
R_{t2}	order-1	$RNAmdm2_n \longrightarrow RNAmdm2_c$	k = 6.59e-02
R_{t3}	order-1	$mdm2_n \longrightarrow mdm2_c$	k = 3.18e-02
R_{t4}	order-1	$p53_c \longrightarrow p53_n$	k = 6.59e-03
R_{t5}	order-1	$mdm2_c \longrightarrow mdm2_n$	k = 3.18e-03

Table 6. The set of reactions for p53-NDR model

A.4 Hes1 Oscillatory Network

The HBDD model used in this paper is extracted from Barton’s model for the Notch signaling pathway.¹⁰⁶ We have converted all reaction unit into number of instances of reactions per minute, and all reactant amount unit into number of molecules. Transcription rate of Hes1’s RNA is adjusted so that the parameter set falls into the Hopf bifurcation regime. The reactions and reaction rates are shown in Table. 7. After 10,000 min of simulation to relax the system, we then proceed to perform measurement. The NDR model is revised from the HBDD model by replacing the nonlinear degradation reactions with their linear counterpart. Transcription rate of NDR model is adjusted to the point that Fano number of oscillation period is lower than 0.3. The reaction and parameter used are shown in Table. 8. For both the HBDD and NDR models of Hes1 oscillation, all reactants have zero initial molecules except for the gene number which has an initial molecule number of one.

order-1	$RNA_{hes1} \longrightarrow RNA_{hes1} + hes1_c$	$k = 0.184$
MM kinetics	$hes1_n + [E] \rightleftharpoons [complex] \longrightarrow [E] + \phi$	$v = k[E] = 9.647 ; K = 24.21$
MM kinetics	$hes1_c + [E] \rightleftharpoons [complex] \longrightarrow [E] + \phi$	$v = k[E] = 507.1 ; K = 264.4$
MM kinetics	$RNA_{hes1} + [E] \rightleftharpoons [complex] \longrightarrow [E] + \phi$	$v = k[E] = 119.2 ; K = 193.7$
order-1	$hes1_c \longrightarrow hes1_n$	$k = 0.00067$
Hill equation	$DNA \cdot 2hes1_n \rightleftharpoons DNA + 2hes1_n \longrightarrow RNA_{hes1} + 2.0hes1_n + DNA$	$k = 158.6 ; K = 29.24 ; H = 2.0$

Table 7. The HBDD model of Hes1 self-regulatory network. RNA, as in Barton’s model, is directly calculated in cytoplasm. Thus, in the Hill equation, the factor of 10 comes from the volume ratio of cytoplasm and nucleus, 10 : 1

order-1	$RNA_{hes1} \longrightarrow RNA_{hes1} + hes1_c$	$k = 0.184$
order-1	$hes1_n \longrightarrow \phi$	$k = 0.1319$
order-1	$RNA_{hes1} \longrightarrow \phi$	$k = 0.04754$
order-1	$hes1_c \longrightarrow \phi$	$k = 0.06989$
order-1	$hes1_c \longrightarrow hes1_n$	$k = 0.00067$
Hill equation	$DNA \cdot 2hes1_n \rightleftharpoons DNA + 2hes1_n \longrightarrow RNA_{hes1} + 2.0hes1_n + DNA$	$k = 119.2; K = 29.24; H = 2.0$

Table 8. The NDR model of Hes1 self-regulatory network.