On robustness in biology: from sensing to functioning

DISSERTATION

Zur Erlangung des akademischen Grades

Doctor of Philosophy (Ph.D.)

vorgelegt

dem Bereich Mathematik und Naturwissenschaften der
Technischen Universität Dresden

von

M.Sc. Mohammadreza Bahadorian

Eingereicht am 11.10.2022
Die Dissertation wurde vom 14.10.2018 bis 10.10.2022
am Center for Systems Biology Dresden (CSBD) angefertigt.

Abstract

Living systems are subject to various types of spatial and temporal noise at all scales and stages. Nevertheless, evolving under the pressure of natural selection, biology has mastered the ability of dealing with stochasticity. This is particularly crucial because these systems encounter numerous situations which require taking robust and proper actions in the presence of noise. Due to the complexity and variability of these situations, it is impossible to have a prescribed plan for an organism that keeps it alive and fully functional. Therefore, they have to be active, rather than passive, by following three essential steps: I) gathering information about their fluctuating environment, II) processing the information and making decisions via circuits that are inevitably noisy, and finally, III) taking the appropriate action robustly with organizations crossing multiple scales. Although various aspects of this general scheme have been subject of many studies, there are still many questions that remain unanswered: How can a dynamic environmental signal be sensed collectively by cell populations? and how does the topology of interactions affect the quality of this sensing? When processing information via the regulatory network, what are the drawbacks of multifunctional circuits? and how does the reliability of the decisions decrease as the multifunctionality increases? Finally, when the right decision is made and a tissue is growing with feedbacks crossing different scales, what are the crucial features that remain preserved from one subject to another? How can one use these features to understand the mechanisms behind these processes? This thesis addresses the main challenges for answering these questions and many more using methods from dynamical systems, network science, and stochastic processes. Using stochastic models, we investigate the fundamental limits arising from temporal noise on collective signal sensing and context-dependent information processing. Furthermore, by combining stochastic models and cross-scale data analyses, we study pattern formation during complex tissue growth.

Zusammenfassung

Lebende Systeme sind in allen Größenordnungen und Stadien verschiedenen Arten von räumlichem und zeitlichem Rauschen ausgesetzt. Dennoch hat die Biologie, die sich unter dem Druck der natürlichen Selektion entwickelt hat, die Fähigkeit gemeistert, mit stochastischen Fluktuationen umzugehen. Dies ist besonders wichtig, da Organismen auf zahlreiche Situationen stoßen, die es erfordern, in Gegenwart von Rauschen robuste und angemessene Maßnahmen zu ergreifen. Aufgrund der Komplexität und Variabilität dieser Situationen ist es unmöglich, einen vorgeschriebenen Plan für einen Organismus zu haben, der ihn überlebens- und funktionsfähig hält. Daher können Organismen sich nicht passiv verhalten, sondern befolgen aktiv drei wesentliche Schritte: I) Das Sammeln von Informationen über ihre dynamische Umgebung, II) Das Verarbeiten von Informationen und das Treffen von Entscheidungen über Regelnetzwerke, die unvermeidlich mit Rauschen behaftet sind, und schließlich, III) das robuste Funktionieren durch organisierte Maßnahmen, welche mehrere Größenordnungen überbrücken. Obwohl verschiedene Aspekte dieses allgemeinen Schemas Gegenstand vieler Studien waren, bleiben noch viele Fragen unbeantwortet: Wie kann ein dynamisches externes Signal kollektiv von Zellpopulationen wahrgenommen werden? Wie beeinflusst die Topologie der Interaktionen die Qualität dieser Wahrnehmung? Was sind die Nachteile multifunktionaler Schaltkreise bei der Verarbeitung von Informationen über das Regelnetzwerk? Wie nimmt die Zuverlässigkeit der Entscheidungen mit zunehmender Multifunktionalität ab? Und abschließend, wenn die richtige Entscheidung getroffen wurde und ein Gewebe wächst und dabei Rückkopplungen auf verschiedenen Größenordnungen erfährt, was sind die entscheidenden Merkmale, die von einem Versuchsobjekt zum anderen erhalten bleiben? Wie kann man diese Merkmale nutzen, um die Prozesse zu verstehen? Diese Arbeit befasst sich mit den wichtigsten Herausforderungen zur Beantwortung dieser und vieler weiterer Fragen mit Methoden aus dynamischen Systemen, Netzwerkforschung und stochastischen Prozessen.

Acknowledgements

Noise and stochasticity are ubiquitous components of life as it will be discussed in this thesis. The journey of a Ph.D. student is no exception. Navigating through this rough terrain and withstanding random perturbations was only possible because of interacting with many influential people whom I would like to thank here.

First and foremost, I must thank my supervisor, Carl D. Modes, for believing in me and giving me this opportunity. Throughout my PhD, his constructive and perpetual feedback helped me gather not only technical knowledge but also soft skills which will be useful in future. Therefore, I will always be grateful to him. On the same note, I would like to thank my thesis advisory committee, including Christoph Zechner, Marc Timme and Dora Tang because of their constructive feedback and immense knowledge that were influential in shaping my work.

I am grateful to all my close colleagues whose active help and support made this work possible. Beside Carl D. Modes who was involved in shaping every step of my projects, Christoph Zechner's ideas and help was crucial to the collective sensing project. The experimental work done by Johanna Lattner and Jacqueline Tabler provided us with the data set of the ossification patterns. Also, their feedback and patience along with the knowledge and communication skills of Adrian Lahola-Chomiak and Allyson Q. Ryan helped me in moulding the questions and discussions of the ossification project.

I would like to thank all members of the Modes group. It is their kind help and support that have made my research and life in Dresden a wonderful time. Many thanks goes Abhijeet Krishna, Felix Kramer and Alicja Szalapak, my doctoral siblings and best friends, who were kind and supportive during the good and hard times. My thanks also goes to our great postdocs, Szabolcs Horvát, Allyson Q. Ryan and Michael F. Staddon, who always made time when I needed during and outside the working hours. I must also thank the foosball team and mountain biking group in CSBD who helped me let off some steam through physical activities.

Finally, none of this could have happened without my family. I express my deepest gratitude towards my parents whose abundant support and encouragement was crucial for my entire education. I conclude by thanking Elahe whose love has made life bearable in the past seven years.

Contents

List of Figures

Chapter 1

Introduction

Stochasticity is present in all real-world systems and is not negligible in many cases. Yet, living systems have found ways to overcome the uncertainties arising from the stochastic components and achieve reliable functionality and reproducibility. Comprehending natural systems and designing artificial ones that can function in the presence of noise require a deep understanding of stochastic processes. To this end, powerful probabilistic methods have been developed with inspirations from statistical physics in the past few decades. However, the earliest form of a stochastic problem dates back to the late 1800s when Lord Rayleigh addressed the composition of n isoperiodic oscillators with unit amplitude and random phase at the limit of $n \to \infty$ [1]. An explicit and more famous formulation of the problem was later given by K. Pearson in 1905 [2] as following: "A man starts from a point O and walks l yards in a straight line; he then turns through any angle whatever and walks another l yards in a second straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and $r + \delta r$ from his starting point, O." This problem was solved in one dimension and with finite n by M. Smoluchowski in 1906 [3]. A more general form of this problem historically known as random flights (finite steps n and random step size l) was formulated by A. A. Markoff who also touched on the method of obtaining a general solution [4]. Converting the random flights problem to a differential equation with given boundary conditions in general was done by S. Charndrasekhar [5]. This Ref. also provides a great source for the history of the early developments in this field from which a part of the discussions here are adopted.

In parallel to the problem of random flights, another problem of significant importance, the *Brownian motion*, was established by A. Einstein in 1905 [6, 7]. Einstein's approach was to derive and solve a partial differential equation that describes the time evolution of the probability distribution of a Brownian particle. This type of equations was later known as the *Fokker-Planck equation* since deriving them was in a general and systematic way described by A. D. Fokker [8] and M. Planck [9]. A mathematically different yet physically similar approach to the Brownian motion problem was introduced by P. Langevin [10, 11].In this work, Langevin applied Newton's second law to a generic Brownian particle and derived its stochastic equation of motion.

The original form of this equation for the velocity v a Brownian particle with mass m and radius a moving in a liquid with viscosity μ reads

$$m\dot{v} = -6\pi\mu a v + \mathcal{X}, \tag{1.1}$$

in which the first term is the viscous resistance derived from Stokes' law. Here, $\mathcal{X}$ is the randoms driving force originally described as following:"About the complementary force $\mathcal{X}$, we know that it is indifferently positive and negative and that its magnitude is such that it maintains the agitation of the particle, which the viscous resistance would stop without it".This was one of the biggest landmarks of the field as it is equivalent of "$F = ma$" for stochastic processes and later called the *Langevin equation*. In the later applications of this formalism, the deterministic part of the dynamics (the drag force in 1.1) is called the drift term and the stochastic part of dynamics (the stochastic driving force in 1.1) is called the diffusion term.

Another seminal work necessary to mention is H. A. Kramers' where he attempts to formulate the deterministic reaction rates in terms of molecular parameters [12]. More importantly, this work illustrates the application of the Brownian motion to the problem of the escape of particles from a potential barrier. This type of problems are nowadays known as *noise-induced transitions* and are of significant importance in a variety of contexts such as biology, ecology, economics and climates science to name a few.

A useful and intuitive approach to model stochastic processes is to consider a probabilistic combination of different states at any given time. Then, the switching between states can be described by a transition rate matrix. This formalism was first derived by W. Pauli to describe the approach of a quantum system to statistical equilibrium [13]. A more rigorous derivation of this approach as well as its higher order considerations was introduced by L. Van Hove [14]. This framework nevertheless did not remain limited to quantum systems and soon found applications in microscopic chemical reaction systems (e.g. biological systems) [15] and countless other fields.

In many cases stochasticity hinders the functionality of the system. For example in multi-stable systems, fluctuations can cause noise-induced transitions from the desired stable state to an undesired one [16, 17].There are, however, many observations where stochasticity is utilized to achieve a certain goal such as noise-induced synchronization in chaotic oscillators [18] as well as stochastic cell fate determination to establish necessary fate diversity in some tissues such as retina [19].

1.1 Noise and stochasticity in biology

Living organisms are exposed to dynamic environments where everything from the temperature to food sources are fluctuating over space and time. Nevertheless, life

is still flourishing in nearly all conditions by robustly gaining information about the dynamic environment. However, the main tools used by cells to infer environmental cues are biochemical reaction networks with small reaction volumes that are intrinsically stochastic due to random timing of reactions also known as *intrinsic noise*. The concentration of these biochemical species also fluctuates over the intercellular space which adds extra uncertainty to external signal estimation in a single cell. Besides, cell-to-cell variability in the rates of the reactions introduces *extrinsic noise* to the system.

After sensing the dynamic and inhomogeneous environmental signal, cells again use their internal noisy chemical reaction network (i.e. regulatory network) to process the information and make decisions. Here, coping with uncertainties originating from intrinsic and extrinsic noise further increases complications. Even in the next step where cells produce proteins and other necessary components, noisy reactions result in uncertainty and variability. Despite all these stochasticities, living cells are capable of surviving and functioning properly. Multicellular organisms face yet another challenge. They even need to maintain the abundance and location of different cell types to form tissues during the development. This requires feedbacks and interactions crossing multiple scales which gives rise to complicated patterns that sometimes necessitate cross-scale analyses for studying them.

In principle, stochasticity can originate from one or more of the following phenomena: I) Lack of technology and precise measurements resulting in lack of knowledge II) Ignoring some of the information (degrees of freedom) to make the system tractable III) fundamental uncertainties at the quantum scale. Quantum phenomena are relevant in certain biological cases such as olfactory sensing, magneto-reception in birds and photosynthesis. However, in most cases the fact that biological systems are "warm, wet and noisy" causes rapid decoherence and suppression of quantum effects [20]. Therefore, throughout this thesis, we solely focus on the cases in which the uncertainties arise from lack of knowledge and/or from ignoring information for the sake of simplicity.

Living organisms differ from most other chemical systems as they utilize nonlinear regulatory pathways that can suppress stochasticities and achieve reliable functioning. Studying these mechanisms, besides deepening our knowledge into biology, can in principle help us to design novel information processing devices based on chemical reactions or similar dynamics. In the remainder of this chapter, we first review different sources of temporal fluctuations and how they affect biological processes. We then briefly discuss some examples of spatial fluctuations and challenges during development. In the end, we touch on general modeling techniques common for studying stochastic processes in biology.

1.1.1 Temporal fluctuations

Many phenomena and processes can cause temporal fluctuations in internal states of cells. A good example is the dynamic environment in which many vital resources and properties change throughout the organisms life cycle. Cells sense these changes and adjust their internal states consequently. At short time scales, these state variables also fluctuate due to the intrinsic noise of molecular machinery used by cells. In this subsection, we first discuss a few experiments showing how dynamic environments affect living systems, and then review empirical evidences of intrinsic noise influencing cells' function.

Living in a dynamic environment

Survival of living organisms depends on the information they can get about their dynamic environment as much as it depends on energy [21]. Studies in this area can be divided into three general categories: I) when time scales of environmental changes are significantly shorter than species' lifetime, II) when the time scale of the changes is comparable to the life time, and III) when time scale of changes is much longer than the life time. In the third scenario which is the simplest one, the classic evolution is the dominant view: seemingly constant environment favors better adopted species enabling them to flourish by using their maximum potential while the other species decline. In the second case with the intermediate time scale, bet-hedging strategy is shown to be the favored strategy in which cell-to-cell variability is utilized for encountering various environmental conditions [22]. However, this is a population level strategy and individuals also have to deal with notable changes during their life time. A simple yet important examples is the winter dormancy of plants. The changes in temperature and daylight duration enables them to sense the start of cold season and begin hibernation process, accordingly [23, 24]. At a shorter scale, single cell organisms also experience changing environment and use their regulatory network to react. For instance, *Escherichia coli* bacterial cells ideally grow consuming glucose, but when exposed to only lactose, they can express Lac protein to catabolise it [25]. When the food source fluctuates, depending on the time scale of the fluctuations relative to the life-time, these cells can incorporate different strategies to maximize their growth rate [26, 27].

Many theoretical studies have focused on determining the fundamental limits of sensing environmental cues carried out via chemical reaction networks at the steady state by a single cell [28, 29]. Similarly, sensing a dynamic environmental signal by a single cell have been studied in [30, 31]. Limitation of collective sensing of an environmental signal is also partially explored, mostly focusing on the steady-state concentration sensed through simplified mean-field interactions [32]. However, combination of dynamic signal and collective sensing requires is not well understood and will be discussed in chapter 2.

Random timing of bio-chemical reactions

Many crucial biochemical reactions regulating cell functions occur at such a low rate that the abundance of the reactants and products follow fluctuating trajectories. Let us demonstrate this by a simplified example. Consider the DNA transcription process which is the first step of gene expression. Generally, in this step, an mRNA molecule is synthesized from a segment of the DNA using an RNA polymerase macro-molecule, and translated into the desired protein in the next step. The ploymerase molecules get recruited at so called promoter regions of the DNA and then transcribe the downstream coding segment. As a specific system, we consider bacteria *E. coli* which is a prokaryote meaning there's no membrane bound nucleus and transcription can happen anywhere in the entire cell volume. The cell volume of E. coli is of order of magnitude of $1\mu m^3$ [33]. In this volume, there are approximately 3000 RNA polymerase [34], $20-30\%$ of which are active [35]. Given that each polymerase is around $10nm^3$ in volume [36], only 0.0001% of the cell is occupied by RNA polymerase. Diffusion coefficient of these molecules is measured to be of the order of magnitude of $0.1\frac{\mu m^2}{s}$ [37]. Given that the DNA molecule, because of its size, is diffusing much slower than the polymerase, one can calculate the typical time that the polymerase molecules need to explore the entire cell volume and find a given promoter site. If we ignore the probability of losing the polymerase to other promoter sites, this time will equal the time needed for filling the cell volume V with sum of spheres that are explored by polymerase molecules. The volume that a molecule explores due to diffusion is a sphere whose radius equals the Mean Square Displacement (MSD) which is $6D\tau$ in three dimension. Therefore, for N molecules, we have $N\frac{4}{3}\pi\text{MSD}^{3/2} = V$, in which the typical time τ equals to $0.04s$ by substituting the values mentioned above. Since the probability of each reaction event in an infinitely small time interval is constant and independent of each other, the reaction events are Poisson point process and the waiting times follow an exponential distribution with average of $\tau = 0.04s$. It should be noted that this is a hand waving argument to show the relevance of stochasticity and the actual problem is much more complicated. Many high-order effects such as the shape of the cell and volume of the polymerase molecules are ignored here. Also, there are complicated mechanisms employed by cells to regulate this typical time that are beyond the scope of this simple calculations. For example, attachment of activating transcription factors to the promoters and can increase its effective volume and consequently boosting the probability of recruiting a polymerase molecule, or attachment of inhibitor molecules can prevent the recruitment of other molecules.

In reality, cells usually produce even fewer mRNA molecules. Consider mRNA production in yeast, as an example. Although directly measuring the reaction events is technically very challenging, if possible, one can infer the kinetic parameters of the

transcription-translation models from the protein abundance over time as demonstrated in Ref. [38]. Here, the authors show that in their specific case, on average around six mRNA molecules are produced per minute and their half-life is 10 minute. This implies that processes as simple as mRNA production should indeed be considered stochastic. This stochasticity gets amplified in the next step of protein production in which mRNA gets translated into protein. As a result, the production of proteins obtains bursting dynamics where periods of high reaction events are followed by long waiting periods between them [39]. This temporal stochasticity in gene expression can explain many variations observed in clonal populations with the same genome which is experimentally observed in Ref. [40] and theoretically discussed in Ref. [41].

1.1.2 Spatial noise

Bursting dynamics, discussed earlier, can result in cell-to-cell variability. Moreover, low copy number of biochemical species combined with slow diffusion can cause spatial fluctuations in the abundance of typical biochemical species. In this section, we briefly review these phenomena and cases where their effects are observed.

Diffusion-limited spatial fluctuations

Much of the cells' information processing and functioning is done via their enormous regulatory networks. Elements (nodes) of this network are biochemical species produced stochastically and diffused in the cell volume stochastically. Most studies solely focus on the temporal fluctuation of these processes and assume well-mixed intracellular space which is achievable in the regime of fast diffusion. However, experiments on the mobility of proteins in the cytoplasm show the relevance of spatial fluctuations even in the small E. coli bacterial cells [42]. During proliferation, inhomogeneous distribution of small number of proteins results in random partitioning of these molecules and causes cell-cell variation in the daughter cells. These variations are significant and their effect is comparable to that of temporal fluctuations of gene expression [43]. In the extracellular environment, crucial signals are also in the form of biochemical species at low copy number which is subject to fluctuations due to slow diffusion. Cells in some communities secret signaling molecules and sense the pool of the molecules to communicate with each other. Inhomogeneity of distribution of these molecules due to low copy number and slow diffusion is another complication making the effective communication challenging.

Cell-to-cell variability

Inhomogeneous partitioning of proteins during cell division causes cell-to-cell variability also known as extrinsic noise [43, 44]. This variability can be distinguished

from intrinsic noise by measuring concentration of two gene products that are controlled by the same transcription factor [45]. This variability can in principle be suppressed by utilizing phase separation where dilute and dense phase coexist and their concentration is determined by the thermodynamics of the system. The extra protein is then partitioned into the dense phase (droplets whose size can vary) keeping local concentrations constant [46]. However, in many cases, the variability is present and affecting cells' functions. For example, in communicating cells estimating the cell density (quorum sensing), the internal master regulators show a big variability especially at low cell density [47]. Similar to other processes, here, internal and external feedbacks can be utilized to optimize quorum sensing despite the cell-to-cell variability [48].

Development of multicellular organisms is a prime example of achieving organization and reproduciblity despite various stochasticities. During this process, different cell types are established at specific location. The positional information required for this are usually provided by so called morphogens gradient [49]. Target cells that are adjacent to the secreting domains can sense the morphogen concentration and express the necessary genes for the desired cell type. Variability in the target cells results in spatial fluctuations of the effective diffusion coefficient as well as degradation rate of the morphogen. These fluctuations in turn limit the precision of the positional information provided by the morphogen gradient [50]. However, cells can circumvent this limitation by transducing the signal through their complex and redundant regulatory network [51]. Some developmental processes such as skull bone growth, involves cell differentiation and and proliferation (resulting in tissue growth) at the same time and with the same time scales which is known as *progressive differentiation*. This adds extra complexity necessitating non-local interactions to achieve organization and reproducibility. These tissues form non-trivial patterns that require high-order and non-local analyses for characterization. For example, when embryo size is varied, non-local feedbacks (e.g. expansion-repression mechanism to scale morphogen gradients) are responsible for maintaining proper development and scaling of the tissues [52, 53].

1.2 Modeling temporal noise in biology

Before discussing robustness of biological processes, one needs to model the temporal stochasticity present in different processes and reaction volumes. We start by construction of Chemical Master Equation (CME) in a well-mixed setup. We then discuss challenges when dealing with CME's. Finally, we discuss how to construct a Langevin equation and its corresponding Fokker-Planck equation at large reaction volume regime resulting in small fluctuations.

1.2.1 Chemical reaction networks

Suppose that N chemical species are connected via m reaction channels as following:

$$\sum_{i=1}^{N} a_{i,j} X_i \xrightarrow{f_j(\mathbf{x})} \sum_{i=1}^{N} b_{i,j} X_i \qquad (j = 1, \ldots, m)\,, \tag{1.2}$$

where X_i represents species i whose abundance is given by x_i, and $a_{i,j}$ and $b_{i,j}$ are the stoichiometric coefficients of reactants and products, respectively which acquire nonnegative values. In a deterministic view, $f_j(\mathbf{x})$ is the rate of the reaction j which is in general a function of $\mathbf{x} := (x_1, x_2, \ldots\ x_N)^\top$. However, in a more rigorous stochastic setup, $f_j(\mathbf{x}(t))\,\mathrm{d}t$ is regarded as the probability of reaction j taking place between time t and $t + \mathrm{d}t$. In a well mixed system (i.e. neglecting the spatial fluctuations) one can simulate stochastic trajectories of abundance (in terms of copy number) of each species over time for any given reaction network using the Gillespie algorithm based on a rigorously derived Monte Carlo procedure [54].

For a probabilistic treatment of the reaction network described in Eq. 1.2, one can write the chemical master equation [15] as:

$$\frac{\partial}{\partial t} P(\mathbf{x}, t|\mathbf{x}_0, t_0) = \sum_{j=1}^{m} \left[f_j\left(\mathbf{x} - \boldsymbol{\nu}_j\right) P\left(\mathbf{x} - \boldsymbol{\nu}_j, t|\mathbf{x}_0, t_0\right) - f_j(\mathbf{x})\, P(\mathbf{x}, t|\mathbf{x}_0, t_0) \right] \tag{1.3}$$

where $P(\mathbf{x}, t|\mathbf{x}_0, t_0)$ is the probability of having the system at state $\mathbf{x}$ at time t given that the system has been initiated from $\mathbf{x}_0$ at time t_0. Moreover, $\boldsymbol{\nu}_j$ is a vector corresponding to the change in the state variable $\mathbf{x}$ when reaction j occurs. Its components are the difference between the stoichiometric coefficients on two sides of Eq. 1.2 i.e. $\nu_{i,j} = b_{i,j} - a_{i,j}$.

Given that for every value of $\mathbf{x}$ there is an Ordinary Differential Equation (ODE) coupled to those with different $\mathbf{x}$, treating these master equations analytically for most practical cases is nearly impossible. However, useful insight can be achieved from the moments of the distribution. In principle, one can multiply both sides of Eq. 1.3 by the desired powers of x_i and integrate over all values, but because of the term $f_j(\mathbf{x})$ on the r.h.s., any given moment equation is coupled to higher-order ones when the rates are polynomial functions. Therefore, one cannot truncate the system of equation to solve them. It becomes even worse in the case of non-polynomial rate functions as they produce moments of non-integer order. The issue with the polynomial rate functions, also known as the moment closure problem, can be circumvented by certain assumptions about the distributions which enables one to rewrite higher-order moments in terms of the lower-order ones. Although this approach offers some insight to the statistics of the system, it does not provide a complete description. Additionally, the assumptions needed here limit the applicability.

1.2.2 The chemical Langevin equation

Assuming that the reaction volume is large enough, one can consider the abundance of each species as continuous variables with small fluctuations over time. Then, the random trajectories of the system can be realized via the chemical Langevin equation that is rigorously derived from the chemical master equation in Ref. [55]. For the reaction network described in eq. 1.2, we have

$$\frac{\partial}{\partial t}x_i\left(t\right)=\sum_{j=1}^{m}\nu_{i,j}f_j\left(\mathbf{x}\left(t\right)\right)+\sum_{j=1}^{m}\nu_{i,j}f_j^{1/2}\left(\mathbf{x}\left(t\right)\right)\Gamma_j\left(t\right), \tag{1.4}$$

where $\Gamma_j\left(t\right)$ are temporally uncorrelated white Gaussian noise with zero mean and and unit variance, i.e. $\left\langle \Gamma_j\left(t\right)\Gamma_j'\left(t\right)\right\rangle =\delta_{j,j'}$. In this setup, the probability distribution of the system starting from $\mathbf{x}_0$ at t_0 is governed by a Fokker-Planck equation as

$$\begin{aligned}\frac{\partial}{\partial t}P\left(\mathbf{x},t|\mathbf{x}_0,t_0\right)=&-\sum_{i=1}^{N}\frac{\partial}{\partial x_i}\left[\left(\sum_{j=1}^{m}\nu_{i,j}f_j(\mathbf{x})\right)P\left(\mathbf{x},t|\mathbf{x}_0,t_0\right)\right]\\&+\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2}\left[\left(\sum_{j=1}^{m}\nu_{i,j}^2f_j(\mathbf{x})\right)P\left(\mathbf{x},t|\mathbf{x}_0,t_0\right)\right]\\&+\sum_{\substack{i,i'=1\\i<i'}}^{N}\frac{\partial^2}{\partial x_i\partial x_{i'}}\left[\left(\sum_{j=1}^{m}\nu_{i,j}\nu_{i',j}f_j(\mathbf{x})\right)P\left(\mathbf{x},t|\mathbf{x}_0,t_0\right)\right].\end{aligned} \tag{1.5}$$

This formalism is widely used for modeling reaction networks at the biologically relevant scales. The first term in Eq. 1.4 is called the *drift* term which describes the deterministic part of the dynamics. The second term describing the stochastic part of the dynamics is referred to as the *diffusion* term. Since the diffusion term for chemical reaction networks depends on the state variables, it is categorized as a multiplicative noise. Fixed points of the system can be determined from the drift term while the diffusion term drives the system away from the trajectories generated by the deterministic dynamics.

1.3 Modeling spatial noise in biology

Spatial noise inside a cell, at low copy number and slow diffusion regime, can be modeled by a master equation similar to 1.3, but with expanding the state vector to incorporate the position of each molecule. This results in an infinite-dimensional system of equations due to continuity of space which rendering this approach useless. Certain assumptions can be made to simplify the system that we discuss in this section. First, we discuss compartmentalization of intracellular space and reaction-diffusion at the sub-cellular scale. We then go beyond cellular scale and consider cell-population level where a mean-field approximation can be used to make the

continuous system tractable as well as the network approach in which cells constitute units (nodes) and all the interactions between them can be condensed and considered as pairwise connections (links).

1.3.1 Spatial fluctuations in the intracellular volume

A rigorous framework to study a dynamic population of compartments with arbitrary interactions is developed in Ref. [56]. This formalism can be used to model spatial fluctuations in the intracellular space. Suppose there are N compartments inside a cell that can in principle represent membraneless condensates, organelles with membrane or even the cytoplasmic space containing the other compartments. Content of compartment i is represented by its D-dimensional state vector $\mathbf{x}_i$ whose components correspond to the number of D chemical species encapsulated and they acquire non-negative integers. Such a system can then be treated as a counting process similar to a chemical reaction system discussed in Sec. 1.2.1 and the full state of the system can be represented by a compartment number $\mathbf{n}_\mathbf{x}$ enumerating the compartments with a given content $\mathbf{x}$. This approach can incorporate spatial fluctuations in terms of variations in the compartment contents, but it cannot capture any positional information about the compartments.

In order to capture the positional information about the fluctuations, i.e. spatial distribution of high or low concentration of a certain species, one can employ a similar method discussed in Ref. [57]. In this approach, one can divide the entire cell volume into N subvolumes that are small enough to assume a constant concentration of given species. Then, the diffusion of a molecule to the neighboring subvolumes can be considered as an event similar to the chemical reactions. The authors here introduce an efficient Monte Carlo algorithm in which the expected time for the next reaction is only recalculated for the subvolumes when they undergo a change due to a reaction or a diffusion event.

1.3.2 Networked dynamics

A large number of elements interacting with each other in a pairwise manner can be modeled as a network where each node represents a subunit, and the links represent the interaction between the nodes. This type of models are widely used in physics, biology and other fields due to their power and generality. Each node can in principle have more than one state variables corresponding to the internal states of interest, but for the sake of simplicity, in this section, we only discuss the scenario where each node has one state variable x_i. The dynamics of such a networked system with N units can be represented as

$$\dot{x}_i = F_i\left(x_i\right) + \sum_{j=1}^{N} a_{i,j} G\left(x_i, x_j\right), \tag{1.6}$$

where $a_{i,j}$ is an element of the adjacency matrix, and equals the strength of the interaction between node i and j, and it can have real value. $F_i\left(x_i\right)$ is the self dynamics of node i, and $G\left(x_i, x_j\right)$ determines the form of the interactions. Variations in cell parameters can be captured in the first term on the r.h.s of Eq. 1.6, and the second term is capable of capturing variations in neighborhood and interactions. Even a diffusion term capturing the temporal fluctuations can be added here to construct a Langevin equation and incorporate temporal fluctuations if needed. Using this type of models, one can thus go beyond mean-field approximations and study effect of variations, separately. Additionally, advanced topological measures are designed to capture local and non-local network features that makes studying role of network structures on the functionality of the system possible [58, 59].

It should be noted that applicability of network dynamics in biology is not limited to multicellular level. This framework is also proven useful for modeling intracellular processes such as gene regulatory network [60, 61].

1.4 Analyzing fluctuations to uncover the underlying processes

So far we have only discussed the modeling approaches for studying fluctuations. However, this type of approaches require a priori knowledge or assumptions about the underlying mechanisms. For example, when building a stochastic model of a specific gene regulatory network, one needs information about gene expression processes such as the intermediate steps (i.e. transcription and translation), an estimate of copy numbers to determine the noise strength, and the relevant dynamics. Also, a set of possible interactions (within the regulatory circuits and its interaction with other genes) is needed to be able to construct a model and study effects of each component.

A prior knowledge or plausible assumptions might not be feasible in some cases at least at the length- or time-scale of interest. However, in such cases, useful information can still be obtained by repeating the measurement over space or time, and analyzing the stochastic and potentially nonlinear data sequence in space or in time (time-series). Various analyses are developed to shed light on the stochastic or deterministic components of the system [62].

1.4.1 Non-parametric model construction

More information such as the dynamical model of a stochastic system can be achieved when adequate empirical data is accessible. For example, little is known about underlying dynamics of electroencephalographic (EEG) recordings as they are readouts of a sum of activity of thousands or millions of neurons throughout the inhomogeneous and complex scalp tissue. Especially, under pathological conditions such as epilepsy, it is challenging to make useful models based on the known mechanisms

of neural activities. Nevertheless, when long time-series are available, one can construct a Langevin equation and its corresponding Fokker-Planck equation for EEG recordings by calculating Kramers–Moyal coefficients [63]. Besides shedding light into the process, these dynamics parameters have also proved useful for diagnostic purposes. A similar pipeline is applied to beat-to-beat fluctuations of the heart rates of healthy subjects and those affected by Congestive Heart Failure (CHF) [64]. In this case, again, a Langevin equation is constructed without assumptions about the underlying mechanisms and its drift and diffusion terms are both capable of distinguishing healthy subjects from those with CHF. Ref. [65] provides a review of this approach in many fields while Ref. [66] portraits a more comprehensive overview on the conditions and generalizations.

It should be noted that in this framework, determining the dynamics is possible solely due to existence of fluctuations. In fact, the stochastic components of the systems under study act as random perturbations and recovery from them is the key to understanding systems dynamics. Considering these fluctuations to be random noise and removing them renders such methods futile.

1.4.2 Combining models with fluctuation analyses

Although non-parametric model construction provides very powerful tools for studying complex systems, they typically require hundreds of thousands of sequential measurements which is not possible in many practical cases. For example, in the context of developmental biology where measurements may be invasive and the processes of interest happen over such a short period of time that given the limited time-resolution, only few measurements are possible. Therefore, constructing a stochastic model as discussed in Sec. 1.3, and fitting the parameters to the empirical data is more practical. As an example, one can consider cell sheet folding in the development of green alga Volvox during which the spherical embryos turn themselves inside out. Here, the combination of theoretical modeling and analyzing shape variability reveals how two separate and temporally uncorrelated mechanisms are involved in this shape change process [67]. In a different context, yet conceptually similar case, analyzing fluctuations in abundance of gene products reveals the kinetic parameters of a regulatory network [68].

Chapter 2

Collective sensing of dynamic environmental signals

Biological systems being exposed to dynamic environments need to gain information about their surrounding and adjust their internal states and functions. Here, like many other cases, collaboration and communication may improve the efficiency and accuracy of sensing. However, this is not well understood in the case of cell populations communicating through arbitrary topologies to estimate a dynamic signal. In this chapter, we construct a robust framework for collective signal estimation via continuous time Markov chains. We then study the effects of communication in the identical cell populations as well as nonidentical ones. We furthermore investigate the role of topological properties of the network on the quality of estimation at the single cell and population level. The majority of the material presented in this chapter is adopted from our article published in 2020 (see Ref. [69]).

2.1 Introduction

Information about the environment is crucial for cells' survival and functionality. Slow changes in environmental conditions are dealt with at the population level and across generations, but fast changes needs reactions at short time scales compared to the life cycle. For example, optimally using limited resources requires bacterial populations to sense the cell density and adjust the proliferation rate. Also in developing embryos, cells need to sense the morphogens and acquire the appropriate fate in a limited time in order to properly form tissues.

Majority of biological signals are in the form of concentration of biochemical species and therefore the ability of a cell to estimate external concentrations determines its fitness and potential to make reliable decisions. Therefore, much effort has been devoted to studying physical limits of concentration sensing in various setups. The pioneering work by Berg and Purcell [28] introduced a lower bound for error of sensing concentration by single cell at steady state and showed that the concentration sensitivity of E. coli is surprisingly close to the optimal value. Later studies

addressed the accuracy of sensing a concentration gradient in a single cell [70, 71]. Also, sensing an external concentration increasing in time by single cells was addressed in Ref. [72]. A more general case in which the external signal can fluctuate in time was studied by Zechner et. al. in Ref. [30] where they develop an optimal estimation framework to address the problem of sensing a dynamic signal by single cells. However, the inference of environmental signals can in principle be made collectively and the topology of interactions between cells may play a crucial role on the quality of estimation.

In other fields, the distributed estimation across communities has already been a central focus for years [73, 74, 75]. However, the first attempts to study the effect of communication on the quality of estimation in cell populations have been made recently. Fancher and Mugler [32], for instance, modeled concentration sensing in communicating cells by directly exchanging transmembrane molecules or through secreting and sensing diffusive signalling molecules.However, even in the first case, this analysis is limited to the mean-field approach where neighborhood is assumed to be identical for all cells. Besides, assuming constant concentration of the external signal limits applicability of the model and obscures any transient effects arising from communications.

In this chapter, we develop a framework capable of addressing the accuracy of sensing a dynamic external signal by cell populations interacting through arbitrary and complex networks. Our formalism rigorously captures temporal fluctuations due random timing of reactions in the sensory and communication channels as well as cell-to-cell variability. We then study the effect of network topology (i.e., neighborhood inhomogeneity) on the sensing accuracy in identical and nonidentical cells. We first investigate these effects in two biologically inspired case studies: I) bacterial-like all-to-all communication (fully-connected networks), II) epithelial-like two-dimensional hexagonal lattice. Then, in order to better shed light on the effects of the interaction topology, we construct random networks and study the interplay between the sensing accuracy (at the single cell and population level) and the local and global topological properties.

2.2 Collective sensing of dynamic external signals

Let us consider a cell population of size N that is supposed to sense a dynamic external signal $Z(t)$. This signal $Z(t)$ can, for example, be the concentration of a signaling molecule acting on the cell population. For the sake of simplicity we consider $Z(t)$ to be a one-dimensional birth-death process as

$$\emptyset \xrightarrow{\rho} Z \xrightarrow{\varphi} \emptyset, \tag{2.1}$$

where ρ and φ are the birth-rate and and death-rate of the molecule, respectively. In our model cells are supposed to estimate the signal up to proportionality constant

γ, i.e. $\gamma Z(t)$. For simplicity, one can consider that cell i is sensing the signal $Z(t)$ through a single catalytic reaction channel:

$$Z \xrightarrow{\gamma c_M} Z + M^{(i)}. \tag{2.2}$$

Here, γc_M is the rate constant where c_M and γ are the sensor rate and the enhancement factor of the cell (i.e. the proportionality constant of the estimation), respectively. When a sensor reaction happens inside cell i, one molecule $M^{(i)}$ is produced which can then regulate the downstream processes. It should be noted that the reaction times of the sensors here provide a noisy and indirect readout of the external signal $Z(t)$ to the cell since higher values of $Z(t)$ results in more frequent firings of sensor reaction and vice versa. Moreover, the enhancement factor γ controls the informativeness of these sensor reactions. Large γ values result in more frequent firing of the sensor reaction and following the abundance of $Z(t)$ more closely.

For this system, Zechner et. al. in Ref. [30], construct a stochastic differential equation governing the time evolution of the so called *filtering distribution*. This distribution determines the probability of $Z(t) = z$ given the observed timings of the sensing reaction in 2.2. They then showed that the mean of this distribution which minimizes the error of estimation [76] can be approximated by the sensor molecule $M^{(i)}$ if it also undergoes the following birth-death:

$$\emptyset \xrightarrow{\gamma\rho} M^{(i)} \xrightarrow{\varphi + c_M} \emptyset, \tag{2.3}$$

when an estimation of $\gamma Z(t)$ is required. We refer to this setup as a *Poissonian estimation* since in the absence of the sensor reaction, the abundance of the molecule $M^{(i)}(t)$ exhibits Poissonian statistics at stationarity. However, in the presence of extrinsic noise arising from cell-to-cell variability or additional chemical steps in the dynamics of estimator, the circuit can exhibit *super-Poissonian* statistics. We consider a super-Poissonian estimator by introducing a random mismatch between the birth rate of the estimator in cell i and the optimal birth rate $\gamma\rho$. More precisely, we assume cell i is equipped by a different birth rate $\gamma\rho^{(i)} = \gamma(\rho + \Delta\rho^{(i)})$ where $\Delta\rho^{(i)}$ for all $i = 1, \ldots, N$ are uncorrelated, zero-mean random variables (i.e., $\mathbb{E}[\Delta\rho^{(i)}] = \mathbb{E}[\Delta\rho^{(i)}\Delta\rho^{(j)}] = 0$) with variance $\mathbb{E}[(\Delta\rho^{(i)})^2] = \sigma^2$.

We extend this model to allow communication between cells. In particular, we consider the scenario in which estimator molecules can diffuse back and forth between two neighboring cells with a fixed rate constant, i.e.,

$$M^{(j)} \underset{\alpha_{ji}}{\overset{\alpha_{ij}}{\rightleftharpoons}} M^{(i)} \tag{2.4}$$

where α_{ij} is the rate constant of transport from cell j to i. It should be noted that $\alpha_{ij} = 0$ if cells i and j are not neighbors. For simplicity, we assume symmetric interactions, i.e. $\alpha_{ij} = \alpha_{ji} = \alpha$ for every connected i and j. Such interactions will equalize the

estimation values and suppress cell-to-cell differences. This is achieved by a net flux of estimator molecules from cells containing more towards neighboring cells with fewer estimator molecules.

2.3 The quality of sensing in cell population

We assume that each cell's internal volume as well as the environment are well-mixed, individually. In other words, we consider the spatial fluctuations of concentration of chemical species to be negligible in every subvolume of the system. Therefore, we can describe the time-evolution of the environmental signal $Z(t)$ and each cell's estimator molecule $M^{(i)}(t)$ by continuous-time Markov chains [77, 78]. In particular, we employ a counting process formalism, where $Z(t)$ and $M^{(i)}(t)$ are described by a system of stochastic integral equations with independent unit Poisson processes counting the occurrence of each reaction (see appendix A.1).

We can then derive the time evolution of the expected estimator $\mathbb{E}[M^{(i)}(t)]$ as described in App. A.1 which reads:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[M^{(i)}(t)] =&\gamma\rho - (\varphi + c_M)\mathbb{E}[M^{(i)}(t)] + \gamma c_M \mathbb{E}[Z(t)] \\ &+ \sum_j \big(\alpha_{ij}\mathbb{E}[M^{(j)}(t)] - \alpha_{ji}\mathbb{E}[M^{(i)}(t)]\big).\end{aligned} \tag{2.5}$$

Note that taking an average over an ensemble of cell populations results in vanishing the effect of the parameter mismatch since by definition $\mathbb{E}[\Delta\rho^{(i)}] = 0$, and consequently both Poissonian super-Poissonian's expected estimator behave similarly over time. Moreover, $\mathbb{E}[Z(t)]$ is the expected environmental signal whose time evolution is governed by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[Z(t)] = \rho - \varphi\mathbb{E}[Z(t)]. \tag{2.6}$$

It is worth noting that Eq. 2.5 has the same form as Eq. 1.6, and therefore, it is capable of capturing spatial fluctuations in the neighborhood as an interactions topology.

2.3.1 Sensing bias in cell populations

We define the Mean Error (ME) $\mathbb{E}\left[e_i(t)\right] = \mathbb{E}\left[Z(t) - {}^{M^{(i)}(t)}/_{\gamma}\right]$ to assess the bias of each cell's estimator. By subtracting the Eq. 2.6 from Eq. 2.5, one easily gets the ordinary differential equation governing the dynamics of expected error of estimation in cell a:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[e_i(t)] = -(\varphi + c_M)\mathbb{E}[e_i(t)] + \sum_j \big(\alpha_{ij}\mathbb{E}[e_j(t)] - \alpha_{ji}\mathbb{E}[e_i(t)]\big). \tag{2.7}$$

By simultaneously solving the set of equations in 2.7 for given initial conditions, one arrives at the time-evolution of the expected error of the estimation. Typical

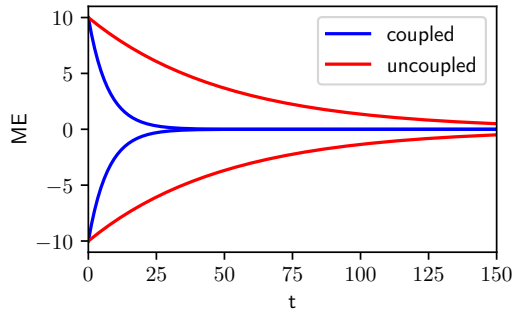


FIGURE 2.1: **Evolution of expected error of estimation** in two cases of coupled (blue curve) and uncoupled (red curve) where parameters are chosen as $c_M = 0.01$, $\rho = 0.1$, $\varphi = 0.01$ and $\alpha = 0.04$.

trajectories of ME for a two cell system starting from opposite biases are depicted in Fig. 2.1. Here, we compare two cases with coupling (in blue) and without coupling (in red). As one can see in this figure, the coupling or communication between cells helps them to achieve zero ME much faster than in the uncoupled case, but in either case ME vanishes at the stationary state. This is due to the construction of estimators.

To demonstrate the performance of these estimators, we simulate the trajectories of the environmental signal as well as a population of three Poissonian estimators using the Gillespie algorithm [79]. Fig. 2.2 depicts a comparison of two fully-connected and uncoupled estimators with a typical set of parameters ($c_M = 0.5$, $\rho = 1$, $\varphi = 0.1$,$\gamma = 1$ and $\alpha = 0.8$). Figs. 2.2a and 2.2c show the transient dynamics of coupled and uncoupled populations, respectively. As one can see here, coupled populations of estimators approach the ground truth signal $Z(t)$ faster. However, both populations of estimators can estimate not only the correct steady state value of $Z(t)$, but also resolve its fluctuations (see 2.2b and 2.2d).

2.3.2 Sensing accuracy in cell populations

The value of the mean error always approaches zero over time, the Mean Squared Error (MSE) $\mathbb{E}\left[e_i^2(t)\right] = \mathbb{E}\left[\left(Z(t) - M^{(i)}(t)/\gamma\right)^2\right]$ represents a better measure for the accuracy of estimation. Using Ito's lemma for counting processes, we derive a set of ODE's describing the dynamics of the MSE of estimation in populations of Poissonian and super-Poissonian estimators as descussed in App. A.2. The generality of our derivation allows studying the quality of estimation in cell populations communicating through any arbitrary network.

Communication effect on sensing accuracy in Poissonian estimators

We first consider a simple mean-field interaction between population of N Poissonian estimators. More specifically, we assume a fully-connected network which can capture, for example, the interactions between cells in a bacterial population through secreting and sensing fast diffusing signaling molecules. This assumption simplifies

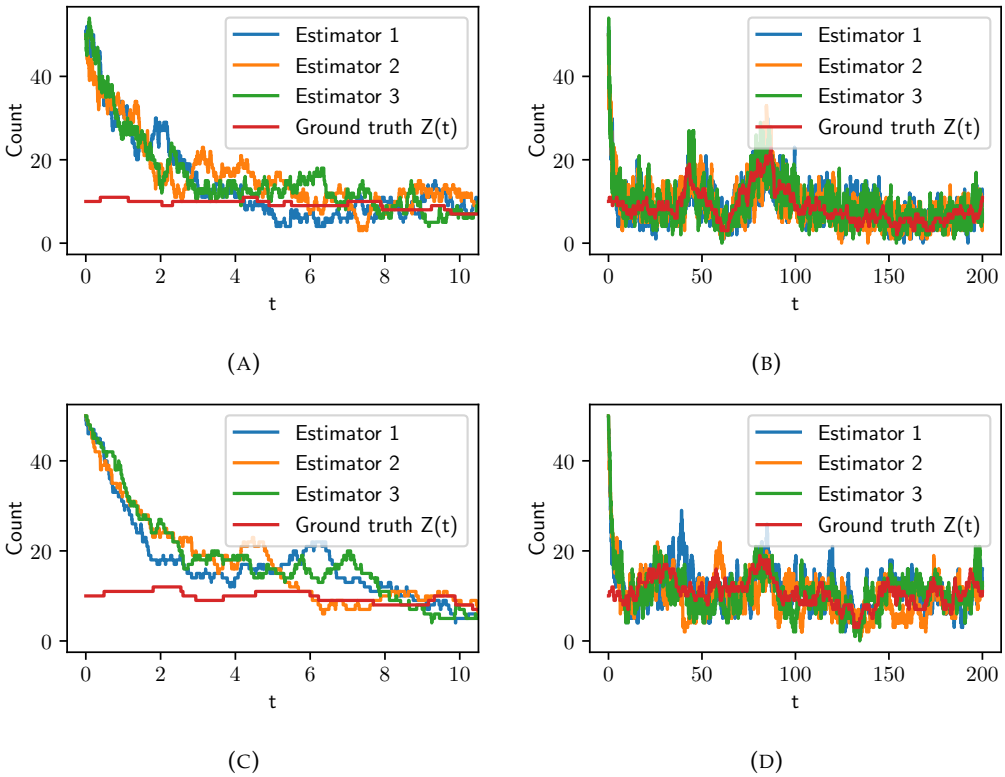


FIGURE 2.2: **Typical trajectories of the environmental signal and its Poissonian estimators** simulated using the Gillespie algorithm [79]. **(A)** The transient part of the estimation in fully-connected estimators. **(B)** The trajectories over a longer period of time. **(C)** and **(D)** also show the transient and longer trajectories in the case of uncoupled estimators. In these simulations the parameters are chosen as $c_M = 0.5$, $\rho = 1$, $\varphi = 0.1$,$\gamma = 1$ and $\alpha = 0.8$.

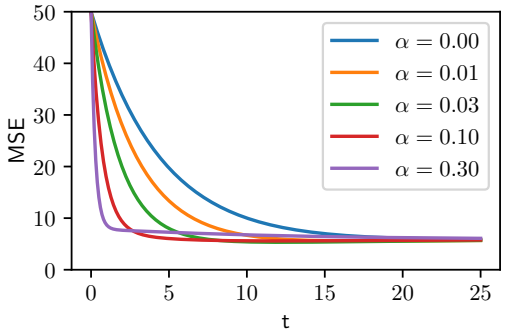


FIGURE 2.3: **Dynamics of MSE of a Poissonian estimator in a fully-connected network** with 5 neighbors and different coupling strength α and following set of parameters: $\rho = c_M = 0.1$, $\varphi = 0.01$, and $\gamma = 2$.

the system significantly rendering it analytically tractable. With this assumption, we rewrite the general equations for time-evolution of MSEs derived in App. A.2 as following:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[e_i^2(t)\right] =& 2\left(\varphi\left(1+\gamma\right)+c_M+n\alpha\right)\frac{\rho}{\gamma\varphi} - 2\left(\varphi+c_M+n\alpha\right)\mathbb{E}\left[e_i^2(t)\right] \\ &+ 2n\alpha\mathbb{E}\left[e_i(t)e_j(t)\right] + \frac{1}{\gamma}\left(\varphi+c_M n\alpha\left(1+\gamma\right)\right)\mathbb{E}\left[e_i(t)\right], \end{aligned} \tag{2.8}$$

for all i and j's. Here, $\mathbb{E}\left[e_i(t)e_j(t)\right]$ is the correlation between the error of estimations in cell i and j. Their dynamics can also be simplified as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[e_i e_j(t)\right] =& 2\rho\left(1-\frac{\alpha}{\varphi\gamma}\right) - 2\left(\varphi+c_M+\alpha\right)\mathbb{E}\left[e_i e_j(t)\right] \\ &+ 2\alpha\mathbb{E}\left[e_i^2(t)\right] + 2\alpha\mathbb{E}\left[e_i(t)\right]. \end{aligned} \tag{2.9}$$

Similarly, the dynamics of ME is governed by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[e_i(t)\right] = -\left(\varphi+c_M\right)\mathbb{E}\left[e_i(t)\right]. \tag{2.10}$$

Solving these equations simultaneously, one can uncover the full dynamics of the system. Fig. 2.3 depicts typical trajectories of MSE for a cell in a fully-connected population with five neighbors and various strength α. As one can see here, starting from the same biased estimation and consequently large initial MSE, the estimators with stronger coupling will approach the steady state MSE faster. However, the steady state values appear be independent from α. One can find these steady state values analytically. The dynamics of ME in equation 2.10 is independent of the other two state variables and can be solved separately, resulting in an exponential decay with exponent $-\left(\varphi+c_M\right)$, i.e. $\mathbb{E}\left[e_i(t)\right] = \mathbb{E}\left[e_i(0)\right]e^{-(\varphi+c_M)t}$. Therefore, at the steady state, ME vanishes and one can solve the following equations for the steady-state

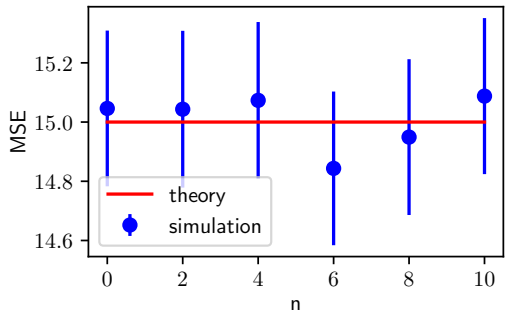


FIGURE 2.4: **MSE vs. number of neighbors $n = N - 1$ in fully-connected networks of Poissonian estimators** with enhancement factor $\gamma = 1$. The average (blue circles) and the standard error (shown with the error bars) of MSE are the result of 10^5 realizations of stochastic simulations.

MSE from stationary A.13 and A.15:

$$\begin{aligned} 0 &= -(\varphi + c_M + n\alpha)\mathbb{E}[e_i^2] + n\alpha\mathbb{E}[e_i e_j] + \frac{1}{\gamma}\big(\varphi(1+\gamma) + c_M + n\alpha\big)\frac{\rho}{\varphi} \\ 0 &= -(\varphi + c_M + \alpha)\mathbb{E}[e_i e_j] + \alpha\mathbb{E}[e_i^2] + \big(1 - \frac{\alpha}{\varphi\gamma}\big)\rho \end{aligned} \tag{2.11}$$

where $n = N - 1$ is the number of neighbors of each cell and $j \neq i$. This system of algebraic equations is easily solvable, but the solution is surprisingly independent of the coupling strength α and the number of neighbors n:

$$\begin{aligned} \mathbb{E}[e_i^2] &= \frac{\rho}{\gamma\varphi} + \frac{\rho}{c_M + \varphi} \\ \mathbb{E}[e_i e_j] &= \frac{\rho}{c_M + \varphi}. \end{aligned} \tag{2.12}$$

Note that here, $\mathbb{E}[e_i^2]$ is a simple shift of $\mathbb{E}[e_i e_j]$ by $\frac{\rho}{\gamma\varphi}$. We verify this by using exact stochastic simulations [79] of a population of estimators with size N communicating across a fully-connected networks. As depicted in Fig. 2.4, by increasing number of neighbors n, the MSE of estimation does not change.

In order to investigate how the coupling terms (coupling strength α and number of neighbors n) are canceled out, we write Eqs. 2.11 in the matrix form i.e.

$$\begin{aligned} &\begin{bmatrix} -(\varphi + c_M) - n\alpha & n\alpha \\ \alpha & -(\varphi + c_M) - \alpha \end{bmatrix} \begin{bmatrix} \mathbb{E}[e_i^2] \\ \mathbb{E}[e_i e_j] \end{bmatrix} \\ &\quad = \begin{bmatrix} -\frac{\rho}{\gamma\varphi}(\varphi(1+\gamma) + c_M) - n\alpha\frac{\rho}{\gamma\varphi} \\ -\rho + \frac{\alpha}{\gamma\varphi}\rho \end{bmatrix}. \end{aligned} \tag{2.13}$$

This can be rearranged as

$$\left(\begin{bmatrix} -(\varphi + c_M) & 0 \\ 0 & -(\varphi + c_M) \end{bmatrix} + \begin{bmatrix} n\alpha \\ -\alpha \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix}\right) \begin{bmatrix} \mathbb{E}[e_i^2] \\ \mathbb{E}[e_i e_j] \end{bmatrix} = \begin{bmatrix} -\frac{\rho}{\gamma\varphi}(\varphi(1+\gamma) + c_M) \\ -\rho \end{bmatrix} + \begin{bmatrix} n\alpha \\ -\alpha \end{bmatrix} \frac{-\rho}{\gamma\varphi}. \tag{2.14}$$

As is evident in this form, the terms including the coupling strength and the number of neighbors are separated into the second term on either side. Thus, they cancel out if their coefficients on both sides are equal. This requires that $\mathbb{E}[e_i^2] - \mathbb{E}[e_i e_j] = \rho/\gamma\varphi$ which is indeed true in our case according to 2.12. This relation can also be written in terms of second order-moment and covariance of estimators as following:

$$\mathbb{E}\left[(M^{(i)})^2\right] - \mathbb{E}[M^{(i)}M^{(j)}] = \frac{\rho}{\gamma\varphi} = \frac{1}{\gamma^2}\mathbb{E}[M^{(i)}]. \tag{2.15}$$

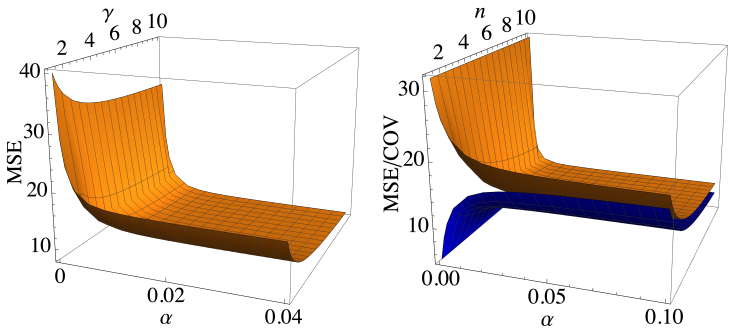
It should be noted that we here solely considered fully-connected network since it is analytically tractable. However, any arbitrary topology can be investigated by numerically solving the set of equations derived in App. A.2 and the general results we presented here are valid, independent of the topology.

In summary, our analysis here shows that communication in a population of Poissonian estimators allows them to reach the steady-state MSE more quickly. This could play an important role during cell fate determination, where cells have to make decisions upon external cues within a limited period of time. However, the communication here does not affect the steady-state MSE. This result is generally because of the fact that the MSE is fundamentally bounded by the intrinsic Poissonian fluctuations, which cannot be overcome by diffusive transport (i.e. spatial averaging). This behavior has been observed in studies of gene expression [80, 81]. Nevertheless, this is new to the context of dynamical signal sensing.

Communication effect on sensing accuracy in super-Poissonian estimators

Having established that the communications in a population of Poissonian estimators solely enhance the rate at which the MSE approaches its steady state value and leaves the the final accuracy unchanged, we now focus on the super-Poissonian estimators.

We first consider the simple case of fully-connected network again since the it is analytically tractable. We solve Eqs. A.9, A.13, A.15, A.16 and A.17 assuming that all cell are identical in terms of topology being connected to all other cells in the population. We then compare the effect of the coupling strength α to that of the enhancement factor γ. Increasing either of these parameters enhances the estimation and decreases the MSE as one can see in Fig. 2.5a. Note that although the MSE



(A) (B)

FIGURE 2.5: **The effect of coupling on the estimation in super-Poisson estimators. (A)** MSE vs. coupling strength α and cell's enhancement factor γ in the fully-connected network of cells with following set of parameters: $\rho = 0.1$, $\varphi = c_M = 0.01$, $n = 10$ and $\sigma^2 = 0.01$. **(B)** MSE (orange surface) and covariance of errors between two cells (blue surface) vs. number of neighbors n and coupling strength α in the fully-connected network of super-Possonian estimators with $\gamma = 6$.

values here are obtained for a specific set of parameter, the qualitative behavior of the results doesn't change over a broad range of relevant parameters.

One can see in Fig. 2.5a that even small values of the coupling strength α can decrease the MSE much faster than by increasing the enhancement factor of the estimation γ. For example, note that even when $\gamma \to \infty$ while $\alpha = 0$, the MSE is larger than when $\gamma = 1$ and $\alpha = 0.002$. These results indicate that cell-to-cell communication is more beneficial for enhancing the estimation compared to expending energy for producing more copies of the signal though the same noisy reaction channel.

Fig. 2.5b shows the MSE of fully-connected super-Poissonian estimators as well as the covariance (COV) of the errors in two cells $\mathbb{E}\left[(e^{(i)})e^{(j)})\right]$ for $i \neq j$. As one can see here, both coupling strength α and number of neighbors n enhance the estimation, but for the specific set of parameters used here, the coupling strength is more effective. Additionally, at a given coupling strength, the COV reduces as the number of neighbors increases which is due to the fact that each cell's estimation would be affected by more cells and therefore each cell will have a smaller effect.

Before further studying the effect of coupling through complex topologies, we analyze the effect of coupling between neighbors of a cell (also known as local clustering) on its MSE. To this end, we compare our results from fully-connected network with another analytically solvable case: A mean field case where a hub cell is connected to n neighbors which are not connected to each other. The results obtained at the steady state in App. A.2.1 show that the MSE of the hub cell is equal to the MSE of any of the cells in a fully connected network of the same size. Thus, to our surprise, the local clustering of a cell does not affect its MSE of estimation, while it decreases the neighbors' MSEs. In order to test this finding in more complicated cases, we study all intermediate steps of transition from a sparse star-shape network

of size $N = 5$ to a dense fully-connected population with the same size. We then calculate the MSE of the cells as well as the COVs. Fig. 2.6 depicts these steps and shows that the MSE of the cell at the center (number 0) is not affected by adding links among its neighbors. This indicates the fact that MSE does not depend on the local clustering of the system is not limited to the mean field cases.

Sensing accuracy in spatial and complex networks

Although the MSE of estimation in a cell at steady-state is not affected by the the connection between its neighbors, its evolution is affected by the COV of its error and the error of the neighbors. The COV between a pair of cells in turn depends on the COVs of each of the two cells with their neighbors (see Eq. A.13 and A.14). Therefore, important effects can arise from next nearest neighbors. This can be of significant importance in a more realistic and spatially extended case where cells interact through networks that are neither fully-connected nor star-shaped but somewhere in between and embedded in real space. We thus study a biologically important scenario in which cells are epithelial-like, with six neighbors (except for those at the boundary) and placed in a two-dimensional regular lattice. We start by constructing a perfect 2D hexagonal lattice with open boundaries. Then, we solve the set of equations derived in appendix A.1 numerically, to find the time evolution of the MSE estimations. As shown in Fig. 2.7a and 2.7b, the MSE as well as the COV depend on the position of the cells within the network. Moreover, the covariance of errors (which can be interpreted as information shared between cells) increases when estimations are worse (i.e. the MSE increases). This implies that cells with worse estimation tend to get more information through their available links. In order to further investigate this observation, we use three ensembles of random networks: Random Spatial (RS), Scale-Free (SF) and Small-World (SW) networks.

We start with the RS ensemble, capturing interactions in tissues and similar systems. In order to construct a RS network without boundary effects, we first make a regular hexagonal lattice with periodic boundaries, and then, cut links with a constant probability of 0.05. We generate 30 realizations with fixed $N = 100$ and then calculate the steady-state value of MSE's. Fig. 2.7d shows one example of these networks while Fig. 2.7e shows its unwrapped version in which some links are cut for a better visualization. In order to further investigate the effect of communication in a broader range of node degrees s, we next generate 30 realizations of Scale-Free (SF) networks with size $N = 100$ and mean degree 6 using the Barabási-Albert model [82]. Finally, since information spreading in complex networks is inherently connected to the small world property of the network [83], we generate 30 realizations of Small-World (SW) networks with $N = 100$ and rewiring probability of 0.05 by employing the Watts-Strogatz model [84].

Fig. 2.8 summarizes the results for these three ensembles of random networks coupling super-Poissonian estimators with the following set of parameters: $\rho = 0.1$,

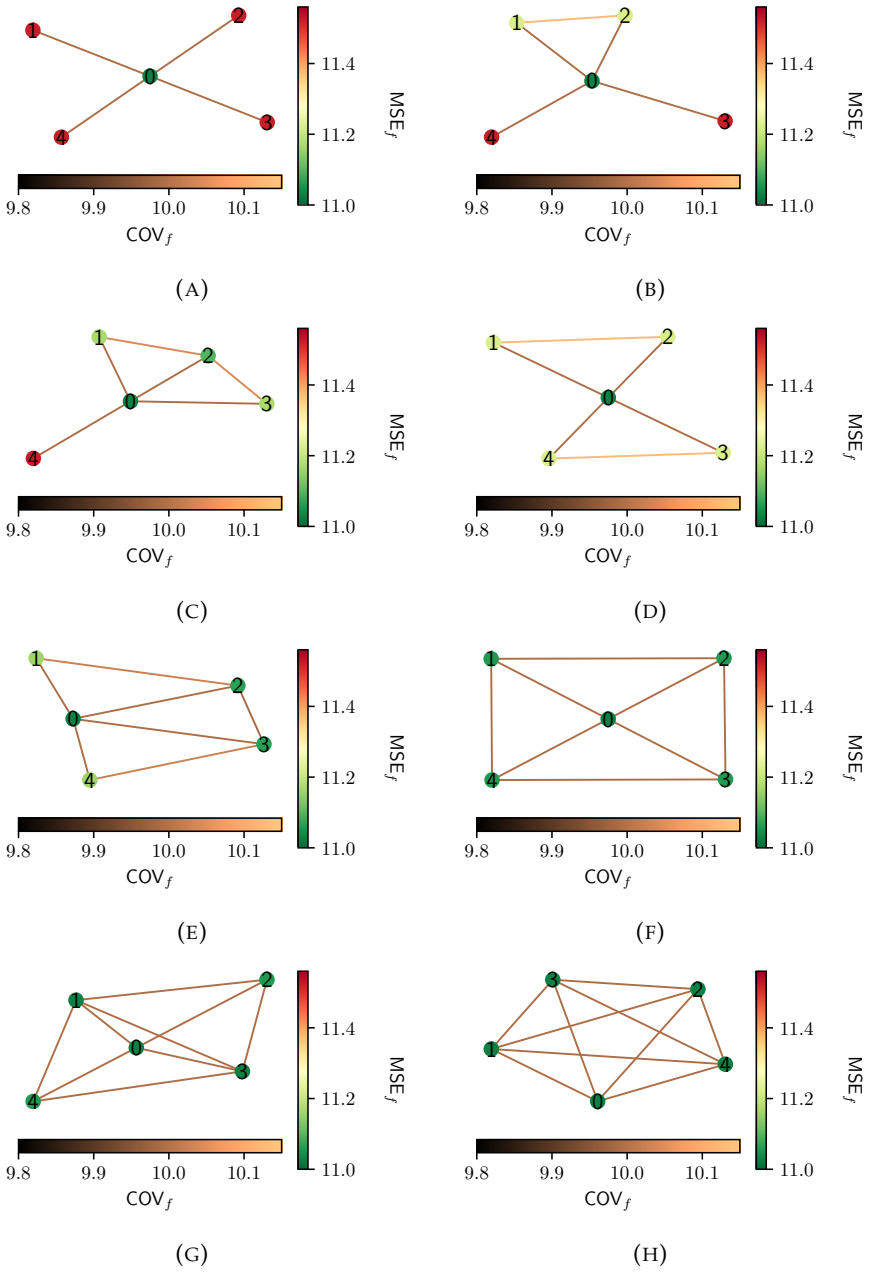


FIGURE 2.6: **The effect of local clustering on a cell's MSE of estimation.** The MSE of cells connected through a sparse stare-shaped network **(A)**, a dense fully-connected network **(H)** and all intermediate steps **(B)–(G)**

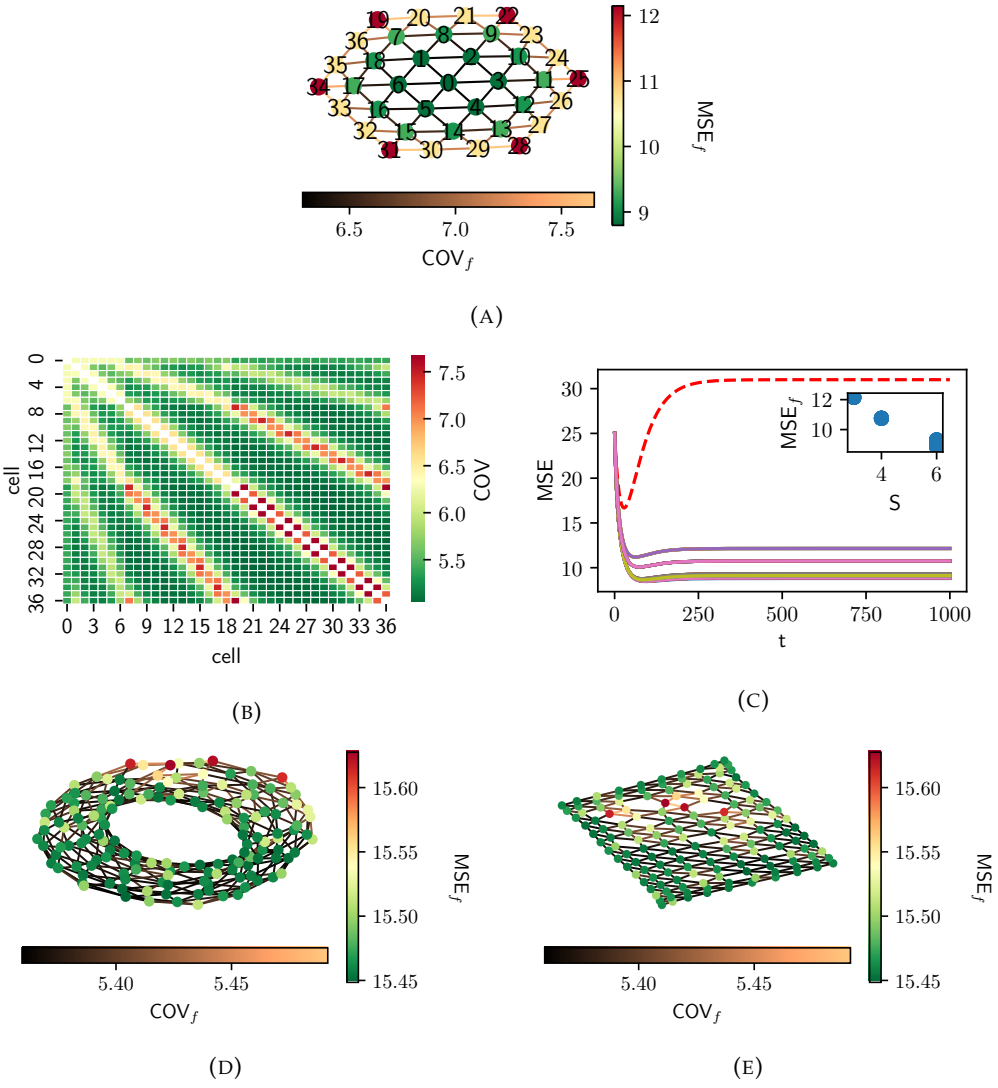


FIGURE 2.7: **Sensing accuracy in spatial networks.(A)** Sensing accuracy in a perfect hexagonal lattice at stationary state **(B)** Covariance between cells in the network at stationary state **(C)** Dynamics of MSE of cells in the network in comparison with a single cell (red dashed line) starting from the same initial conditions **(D)** The MSE and covariance of errors of cells in a typical RS network with real representation. **(E)** The same network in (D) which is only unwrapped for the sake of visibility

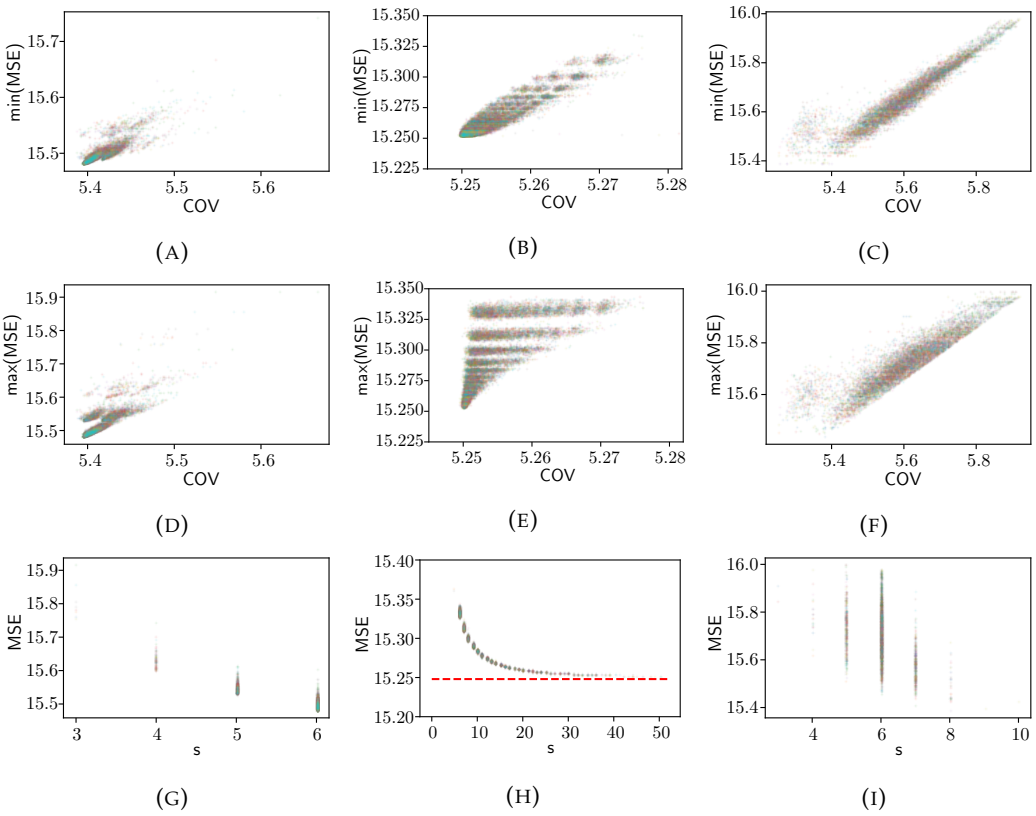
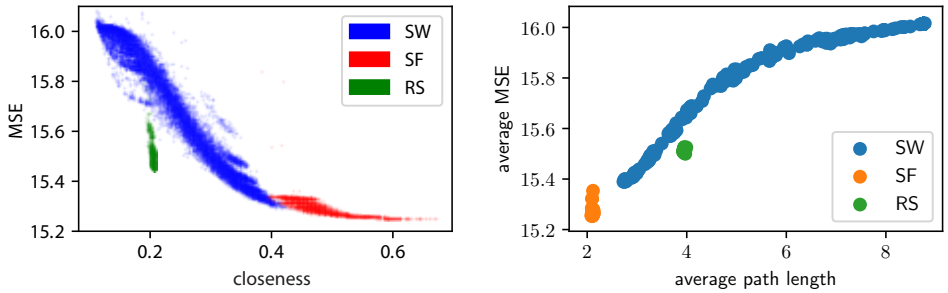


FIGURE 2.8: **Sensing accuracy in random network.** The results of 30 realizations of random networks with size 100 and topology of RS (the first column), SF (with mean degree 6 in the second column) and SW (with rewiring probability of 0.05 and mean degree 6 in the last column). Each color represents a specific realizations. Scatter plot of COV of a link vs. the minimum of MSE over the two ends of that link is shown in the first row. **(A)**, **(B)** and **(C)** depict this for RS, SF and SW topologies, respectively. Scatter plot of COV of a link vs. the maximum of MSE over the two ends of that link in RS networks **(D)**, SF networks **(E)** and SW networks **(F)**. Scatter plot of MSE vs. degree of the nodes s for RS networks **(G)**, SF networks **(H)** and **(I)** . The red dashed line in (H) shows the MSE of a cell in a fully connected network with size 100 and the same set of parameters used in SF network. This implies that in SF networks, the hubs have almost the same quality of estimation as in a fully connected networks with the same size, although their degrees do not exceed 50.



(A) (B)

FIGURE 2.9: **The role of local and global topological features on the MSE of estimation(A)** Scatter plot of MSE of each cell vs. its closeness centrality for all realizations of three different topologies presented in Fig. 2.8. **(B)** Average MSE vs. average path length for different network topologies: blue circles correspond to SW networks with rewiring probability ranging from 0.0001 to 0.5, orange circles correspond to SF network and green circles show the result of RS networks. In all of these simulations, the mean degree is equal to 6.

$\phi = 0.01$, $c_M = 0.01$, $\gamma = 1$, $\alpha = 0.06$ and $\sigma^2 = 0.01$. The relations between covariance of errors of two neighboring cells and the smaller MSE of estimation of those two cells are shown in Fig. 2.8a for RS, in Fig. 2.8b for SF and in Fig. 2.8c for SW topology. Similarly, Figs. 2.8d, 2.8e and 2.8f show the relation between covariance of errors of pairs of neighboring cells and the higher MSE of estimation of those two cells for RS, SF and SW topologies, respectively. As on can see here, there is a strong correlation between COV and the MSE of estimation of the cells. This topology-independent correlation verifies our previous hypothesis about the relation between MSE and covariance in the hexagonal lattice. This suggests that cells relying more on a few links generally achieve a worse estimation accuracy than cells which can average over more individual links. Moreover, we check the relation between MSE of estimation in a cell and its degree s. Figs. 2.8g, 2.8h and 2.8i show this relation for RS, SF and SW topologies, respectively. The negative correlation shows that indeed by increasing the degree of a cell the quality of its estimation increases. Surprisingly however, in SF networks, as shown in Fig. 2.8h, the MSE of the hubs approach a value (shown by dashed line) which corresponds to the MSE of a cell in a fully-connected network with the same size (i.e. 100). One should note that in these realizations, the hubs' degrees do not exceed 50 and yet the MSE of their estimation is very close to that of cells with 100 neighbors. This may well be because of the high closeness centrality (defined as the average of distance from the given node to all other nodes [85]) of the hubs of SF networks. We therefore investigate the dependence of MSE of each cell on its closeness centrality for all different topologies. As depicted in Fig. 2.9a, the scatter plot of all nodes from all random networks generated with the three topologies shows an anti-correlation, indicating that closeness centrality is a key property with critical effect on the quality of estimation in a cell, independent of the network topology.

Accuracy of estimation at the population level

So far we have only focused on the effect of coupling on MSE of estimation at the individual cell level, but now, we investigate the role of coupling on the overall MSE at the population level. Recall that the MSE of estimation on a node depends not only on the number of its first neighbors, but also on its next neighbors and beyond. However, the effect of more distant neighbors is weaker on a given cell. This can be seen in Fig. 2.7a where the effect of boundary cells gets weaker as one moves towards the center. As a consequence, the *average path length* of a network (the average number of steps along the shortest paths for all possible pairs of nodes [85]) is a great candidate for the read-out of the effectivness of communications in a network. In other words, one would expect that at a given population size, the cells will be on average more affected by the other cells if the network of interactions has a lower average path length. We hypothesize that populations interacting through a network with lower average path length will have a better average estimation. To test this hypothesis, we generate SW networks with different rewiring probabilities which result in different average path lengths. Then, we compare their average MSEs. Fig. 2.9b depicting the average MSEs of the SW networks with 15 different rewiring probabilities (from 0.0001 to 0.5) and 10 realizations for each rewiring probability shows that indeed the average path length is a crucial parameter. This figure also depicts average MSE of SF and RS networks collapsing on the same curve as the SW networks.

2.4 Conclusion and outlook

We employed the continuous-time Markov chain formalism to develop a framework for studying a collective signal sensing based on an approximation of the optimal Bayesian estimation. Our framework is general and capable of incorporating various noise types: intrinsic noise, cell-to-cell variability and spatial fluctuations of neighborhood. We demonstrated this by a few case studies. We showed that in identical cell populations where only intrinsic noise is present (i.e. Poissonian estimators), estimation at the steady state does not benefit from communication between cells. Here, communication solely reduces the rate at which cells reach unbiased estimation and their steady state value of MSE. This transient enhancement is important when cells need to collect information and make decision in a limited time such as cell fate determination in developing tissues. We furthermore considered extrinsic noise or in other words unidentical cell populations. In this setup, communication indeed reduces MSE of estimation at steady-state. However, to our surprise, communication between a cell's neighbors and their MSEs do not affect its quality of estimation. By studying estimation in random networks, we show that at the individual cell level, the MSE is inversely related to the closeness centrality and the average MSE of a population is related to the average path length of the interaction network.

Although we here solely focused on the estimation of an external bio-chemical signal, our framework is general and applicable to other similar problems where estimation of a dynamic external physical quantity (other than concentration) for a cell population is needed. For example, when a cell sheet is bounded by a curved surface, the cells' membrane acquire a similar curvature. This membrane curvature which is a readout for the surface curvature is sensible through activity of bent macro-molecules such as BAR domains [86]. Therefore, our work presented in this chapter is an important step towards understanding and modeling information processing tasks performed by a large coupled and, spatially distributed system with complex network of interactions relevant in biology, distributed swarm computing across a complex network, etc.

Chapter 3

Context-dependent information processing in living cells

In order to survive, cells need to collect information regarding their dynamic environment. Then this collected information, along with the information regarding the internal states are processed via the cell's regulatory network. Given that numerous information processing tasks are required for a typical cell's life, reusing the circuits are of significant importance. In other words, information processing circuits must be flexible to perform multiple necessary tasks. In this chapter, we introduce a few *dynamically switchable logic gates* that can perform different logical functions depending on the basin of attraction that the system resides in. We study the robustness of these gates against various noise types, and characterize the trade-off between reliability and their multifunctionality. Much of our results and discussions in this chapter are adopted from our pre-print in arXive (see Ref. [87]).

3.1 Introduction

Various types of information processing tasks are necessary in biology at all levels of complexity and across many length- and time-scales [21]. For example, pack animals collectively make high-level social decisions requiring integration of vast amounts of information about other individuals in the pack [88, 89].Also, predators must make rapid, complicated decisions during their hunt for food while potential prey must navigate similar complexities to avoid being predated upon [90, 91]. At the level of individual high level multicellular organisms, numerous information is processed in the central nervous system consisting of billions of neurons communicating via electrophysiological pulses [92]. At the individual cell level, however, information typically takes the form of concentration of bio-chemical species and is processed in non-trivial and non-linear ways via cell's enormous and complex regulatory network [29, 93, 94].

Depending on the context (e.g. different environmental conditions or stages of life

cycle), cells perform a variety of distinct functions [95]. Much effort has been devoted to studying the relation between functions and the regulatory mechanisms assuming each circuit carries out one information processing task. However, recent studies have shown that the whole regulatory network can change the effective topology to incorporate multifunctionality [96].

The observed multifunctionality in living organisms have driven some attempts to explain it in terms of multistable dynamical systems [97]. For example, Jiménez et al. computationally survey a broad range of bi-functional circuits and show that in many cases two qualitatively distinct functions cannot be mapped to the sub-circuits [98]. Payne and Wagner with a computational model exhaustively study many configuration of three-gene circuits and show that they can perform enormous diversity of patterns some of which exhibit multifunctionality [99]. Perez et al. combine a bistable motif with an oscillator and show that the circuit can have different dynamics depending on the control parameters and initial conditions [100]. Moreover, many experiments have confirmed existence of multistability in regulatory networks [101, 102]. However, fundamental questions about the multifunctionality in regulatory network remain unanswered. For example, what are the advantages and disadvantages of using multistable circuits instead of many separate circuits in the presence of intrinsic and extrinsic noise?

Flexible information processing and multifunctional circuits are well-established and relatively common concepts in other contexts. In the realm of silicon-based information processing, for example, *field-programmable gate arrays*, in which a layer of memory bits sets the connection between logic gates at the circuit layer, are widely used due to their flexibility and reusability. Similarly, in the context of neural information processing, context-dependency is utilized to achieve function switching at time-scales much smaller than what is needed for plastic changes [103, 104]. In biology, however, context-dependancy is more than just a luxury to achieve fast function switching. Here, as mentioned earlier, information has the from of copy number or concentration of bio-molecules. Rerouting this type of signal to the desired circuit is very challenging if possible making context-dependency or multifunctionality of the circuits of critical. Besides, since living cells generally have several functions, being able to perform many of them with a given subsystem is crucial. Especially, when there are limited resources. Nevertheless, the trade-off between multifunctionality and robustness in biology is well studied. Therefore, in this chapter we propose a framework for studying context-dependent information processing which allows different operations (e.g. logical functions) by the same unit on a given input.

Studying information processing in a given setup benefits from investigating properties of its building blocks, i.e. *logic gates*. For example, discussions about quantum information processing are formulated around its constitutes: quantum logic gates [105]. Similarly, in biology, many studies have focused on designing an investigating logic gates implementable by living systems [106, 107, 108].

While traditional logic gates make useful toy models for traditional and static information processing devices, they are unable to constitute a multifunctional circuit. For studying such circuits, dynamically switchable logic gates are required that can dynamically switch between different functions when needed. This concept was first introduced in the context of neural information processing in Ref. [109]. It should be noted that some of the currently available biological gate designs have the capability of multifunctional operations, e.g. by increasing the output's threshold [110], but a systematic study of such multifunctionality, their costs and benefits is missing. Also, the switching in the current designs are dictated externally and directly to the output by changing its inhibitor concentration. This is equivalent to altering the circuit's structure, and it is not a result of the underlying dynamics of the system. However, we here introduce some examples of dynamically switchable logic gates based on dynamics of regulatory network of cells. We then show their applicability by constructing a binary adder/subtractor, and discuss the advantages and disadvantages of multifunctionality based on these examples.

3.2 Dynamics of gene regulatory networks

Before discussing the designs of dynamically switchable logic gates, let us review the dynamics of their components. Here, we assume that our circuits are composed of biochemical species used by cells to process information and make decisions. In principle, these species can be gene products or any other regulatory components of a cell. The time evolution of concentration of such molecules are typically described by a simple and nonlinear dynamical equation with two terms corresponding to the production and degradation processes. To be more specific, we here assume that the dynamics of concentration x at time t is governed by

$$\dot{x}(t) = F\left(s(t)\right) - \gamma x(t), \tag{3.1}$$

where γ is the degradation rate constant, s is the sum of all incoming regulatory signals (activation, repression, and self-regulation loops) to this element, and $F\left(s\right)$ is the regulatory function that describes the production rate as a function of the input regulatory signal s. Note that activation signals have positive contribution while repressive ones have negative contribution to the total incoming signal s. Throughout this study, we use a phenomenological and common sigmoidal function as

$$F\left(s\right) = \frac{1}{1 + e^{-\beta(s-\alpha)}} \ , \tag{3.2}$$

in which β controls the sharpness (i.e. inverse of the fuzziness), and α controls the location of the switch (threshold of sigmoidal function) or in other words, the value of the input signal after which the production occurs [98]. Note that the threshold α can also acquire negative values meaning the species will be produced even if it is

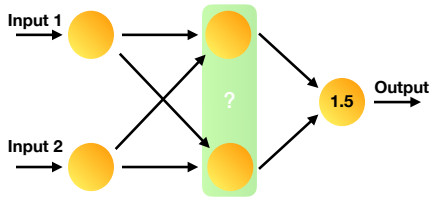


FIGURE 3.1: **The generic dynamically switchable logic gate.** In this structure, there are two inputs and one output as well as an intermediate layer shown in green box. Each element of the circuit is represented by an orange circle and the number written on it shows its threshold of activation α. Normal arrows (i.e. $\rightarrow$) show activation (positive) signal.

inhibited with a strength smaller than this negative value.

It should be noted that we use this specific type of dynamics because of its simplicity. However, all the discussions and results presented here are valid when assuming a different dynamics such as Hill function, as long as the production term has switch-like behavior. We demonstrate this generality in in App. B.1 by reproducing part of our results using a Hill function.

3.3 Dynamically switchable logic gates

Let us begin by considering a simple and generic configuration of a dynamically switchable logic gate, as depicted in Fig. 3.1, based to the proposed circuit in Ref. [109] in the context of neural networks and their corresponding dynamics. In our case, as mentioned earlier each element of the circuit shown by an orange circle in the diagrams represents a biochemical species. Moreover, the number on each element shows its threshold of activation α. In this configuration, an intermediate layer (highlighted by the green box) receives the signal from two inputs (upstream elements), and sends a signal to the output (downstream element) according to its state variables (i.e., concentration of the elements in this layer). In our setup, positive signal (activation) from the two components in the intermediate layer is necessary for activating the output since its threshold α is equal to 1.5. It should also be noted that in this diagram and the following ones, the links have unit weight unless it stated otherwise. Negative signals (inhibition) will represented by bar-headed lines (i.e. $\dashv$) as opposed to positive signals that are represented by arrows (i.e. $\rightarrow$.)

For the sake simplicity, we consider the state of each element in the left layer (inputs to the intermediate layer) to be constant over time, independent from each other and acquiring either zero or one corresponding to OFF and ON state, respectively. Therefore, we here only focus on the dynamics and fluctuations of the intermediate layer.

3.3.1 The ANDOR gate: structure and dynamics

The simplest case of a dynamically switchable logic gate is when the circuit can perform "AND" and "OR" functions hence called ANDOR. According to the truth tables of these two gates, the output should be OFF regardless of the gate type when there is no input. Similarly, for both gate types, the output should be ON when both inputs are ON. The only difference between the output of these two gates is when one of the inputs is ON and the other is OFF (intermediate signal level). In this case, an OR gate results in ON output while AND gate's output is OFF. Therefore, replacing the green box in Fig. 3.1 with a circuit which has bistability at intermediate input level enables the gate to perform OR and AND functions. This can be achieved by replacing the green box of Fig. 3.1 with a bistable motif which has two element inhibiting each other and one of them has a self-promotion link as shown in Fig. 3.2a. Note that here, s corresponds to the sum of the two inputs to the intermediate layer (i.e. $s \equiv$ Input 1 + Input 2) which can only be zero, one or two and is assumed to be static. The concentration of the two elements in the intermediate layer x_1 and x_2 can however acquire any positive real values, and their time-evolution is governed by

$$\dot{x}_1 = \frac{\lambda_0}{1 + e^{-\beta(s - x_2 + \omega(x_1 - \alpha_1))}} - \lambda_1 x_1 \tag{3.3}$$

$$\dot{x}_2 = \frac{\lambda_0}{1 + e^{-\beta(s - x_1 - \alpha_2)}} - \lambda_2 x_2. \tag{3.4}$$

Here, λ_1 and λ_2 are the degradation rate constants for x_1 and x_2, respectively. λ_0 is the production rate of both genes when they receive a very strong input signal. Also, α_1 and α_2 are the activation thresholds for x_1 and x_2 which in our case are equal to 0.5 and 0.3, respectively. Finally, ω controls the slope of the separatrix of the two basins of attraction which equals to 1 unless otherwise is mentioned.

Fig. 3.2b shows the phase portrait of the system with no input. In this type of plots, streamlines represent the flow in the phase space and and their color shows the strength of the flow field. Fig. 3.2b also shows the nullclines $\dot{x}_1 = 0$ (in blue) and $\dot{x}_2 = 0$ (in orange) that cross each other three times producing three fixed points. Two of them are stable fixed points (marked by filled black circles) while the other one is a saddle point (marked by a hollow circle). As one can see here, with $s = 0$ two stable fixed points shown by filled black circles are at $(1, 0)$ and $(0, 1)$ both corresponding to one of the elements being ON and the other one being OFF. Approaching either of these fixed points results in OFF output.

At high input level $s = 2$, as depicted in Fig. 3.2d, there's only one crossing resulting in one stable fixed point at $(1, 1)$ which corresponds to availability of both x_1 and x_2 that can together turn the output ON. Finally, at the intermediate input level (i.e. $s = 1$), the dynamics shown in Fig. 3.2c has two stable fixed point. The stable fixed point at $(1, 1)$ again corresponds to the situation where both elements are produced and available for the downstream genes while the other stable fixed

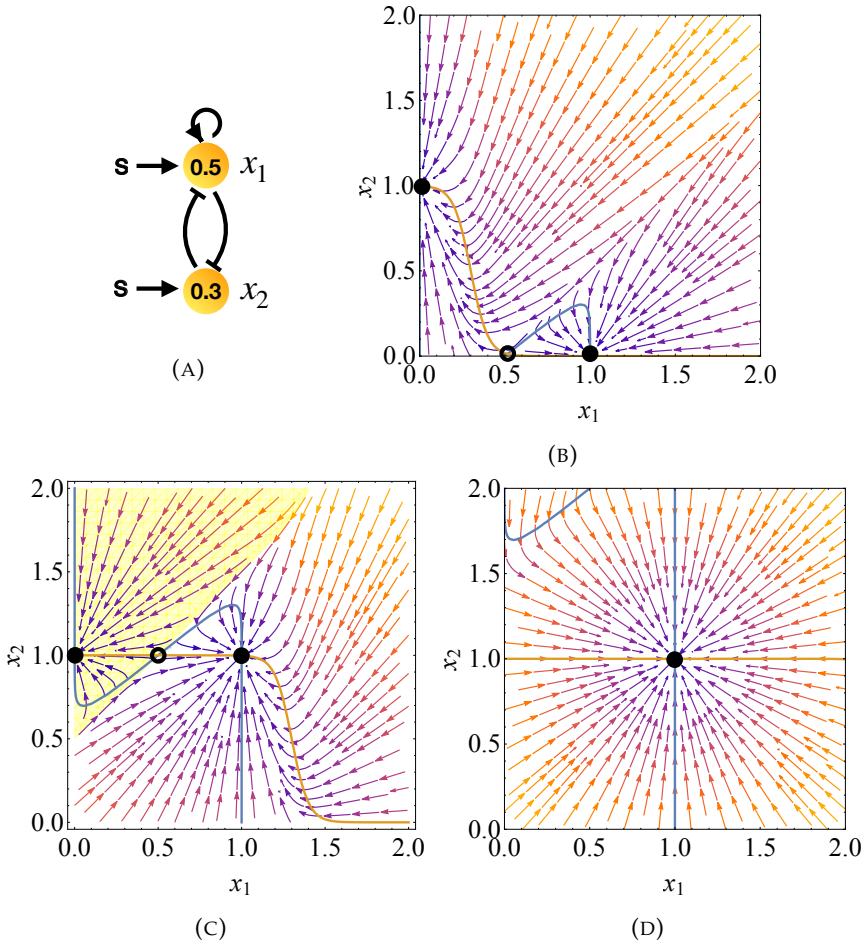


FIGURE 3.2: **The ANDOR gate. (A)** Topology of the ANDOR gate. Here, $s \equiv$ Input 1 + Input 2, and the bar-headed lines (i.e. $\dashv$) represent repression (negative) signal. **(B),(C) ,(D)** The phase portrait of the bistable motif shown in (a) with no input signal $s = 0$, intermediate signal $s = 1$, and high input $s = 2$, respectively. The black filled and hollow circles show the stable fixed points and saddle point, respectively. Moreover, the shaded area in (B) shows the basin of attraction of the left fixed point.

point corresponds the situation in which only x_2 is expressed, and it represses the production of x_1. Therefore, replacing the green box of the generic configuration in Fig. 3.1 with this motif enables the system to perform both AND and OR functions. Whether the ANDOR gate performs AND or OR function depends on the "context" of the decision and is discussed in Sec. 3.3.1.

It should be noted that the ANDOR gate only requires five components while traditional static AND and OR gates three components each. Besides, an additional controller unit causing an even higher component cost is needed to redirect the signal to the desired gate if function switching is not possible. Therefore, using the ANDOR gate can reduce the number of necessary components, significantly. It is also worth noting that the bistability which allows this *emerges* from the interaction between the components. In other words, there is an *emergent behavior* of the system that enables it to perform multiple functions without requiring as many components compared to other systems lacking this behavior.

Logical function switching in the ANDOR gate

In the ANDOR gate, switching between the two functions is possible by simply transitioning from one basin of attraction (shaded area or the rest of the phase space in Fig. 3.2c) to the other one. This transition can happen due to an additional and external signal to the dynamics of x_2 which might be tissue specific and originate from the extracellular environment. For example, there can exist another production reaction (additional term in Eq. 3.1) dedicated for the production of species 2. By such a signal, the value of x_2 can be controlled and the system can be steered vertically to the desired basin of attraction. Using such a strategy, cells would be able to make different decisions depending on the tissue which they are a part of, and the external signal that they receive because of that. Moreover, the decision made (i.e. the approached fixed point) can also be controlled via the initial concentration of the components i.e. $x_1(t = 0)$ and $x_2(t = 0)$. This strategy is useful in cell differentiation processes where two daughter cells acquire different fates due to unequal partitioning of the original cell content during the division. For example, amount of x_2 could be divided unequally between two daughter cells, resulting in two distinct initial conditions and giving rise to two distinct decisions and cell fates. This type of asymmetric cell division leading to distinct fates is observed in the development of retina [111]. Finally, if the system has to be initialized at the origin of the phase space (i.e. zero concentrations), the final state can be controlled by adjustment of the slope of the separatrix (the line which separates two basins of attraction). In this scenario, the function (i.e. AND or OR) that the system performs can be interpreted as the "ground state" of the system since it is the function performed naturally, without requiring expending any additional external energy for steering the system.

3.3.2 Beyond ANDOR: the ANDOROFF and XORANDOROFF gates

Having multifunctionality enables a system to perform more functions without requiring as many components as its static counterparts, but it may also result in unwanted noise-induced transitions. In order to have a framework to study the tradeoff between multifunctionality and robustness, one needs to have circuits with more functions and possibly more components. Therefore, we here go beyond the simple ANDOR gate and design more novel dynamically switchable logic gates capable of performing a greater number of distinct functions. This is achievable by increasing the complexity of the intermediate layer (the green box in Fig. 3.1). For example, we consider a motif in which both genes in the toggle switch have a self-induction loop, as depicted in Fig. 3.3a. This circuit with no input (i.e. $s = 0$) has three stable fixed points as shown in Fig. 3.3b none of which send strong enough signal to the output to turn it ON. For the intermediate signal level (i.e. $s = 1$) shown in Fig. 3.3c, there are three stable fixed points. Two of them are identical to the ones in Fig. 3.2 while the third stable fixed point corresponds to a state in which x_1 dominates and inhibits the expression of x_2. Finally, when $s = 2$ (i.e. both inputs are ON), one of the fixed points in which output is OFF will be preserved (as shown in Fig. 3.3d) if the self promotion loop of x_2 has a relatively high strength, for example twice as strong as the other promoting links. The basin of attraction of this fixed point (shaded area in Fig. 3.3d) corresponds to the set of initial conditions from which the system cannot reach the central fixed point (neither with one input nor with two). Thus, this area represents the situations where the system acts as an OFF switch. This means that adding a new self loop and producing a new fixed point enables the system to switch among three different functions: AND, OR and OFF. We therefore name this gate "ANDOROFF".

Although increasing the number of fixed points can in principle enable the system to perform more functions, it decreases the size of the basin of attraction of each fixed point. Accordingly, the reliability of each decision, defined as the robustness of each fixed point against uncertainty in the initial conditions, decreases. One way to overcome this limitation is by increasing the dimension of the phase-space by increasing the number of regulatory components at the intermediate layer. This will be discussed in detail later in Sec. 3.5.1. Our proposed "XORANDOROFF" gate depicted in Fig. 3.4a has three genes in the intermediate layer which, with a proper set of parameters, can have four different fixed points at $((0,0,0)$, $(1,0,0)$, $(0,1,1)$ and $(1,1,1))$. However, for each input signal ($s = 0, 1$ or 2), only two of them occur, and with the appropriate set of parameters, this circuit can switch between four logical operations: XOR (exclusive OR), AND, OR and OFF. In Fig. 3.4b, consider a system initially located at the blue region. For no input the system approaches the fixed point at which no gene is expressed $(0,0,0)$ and the output is OFF. When one of the inputs is ON, the system, starting from this region, approaches the fixed point at which only x_2 and x_3 are expressed $(0,1,1)$ and therefore, the signal for

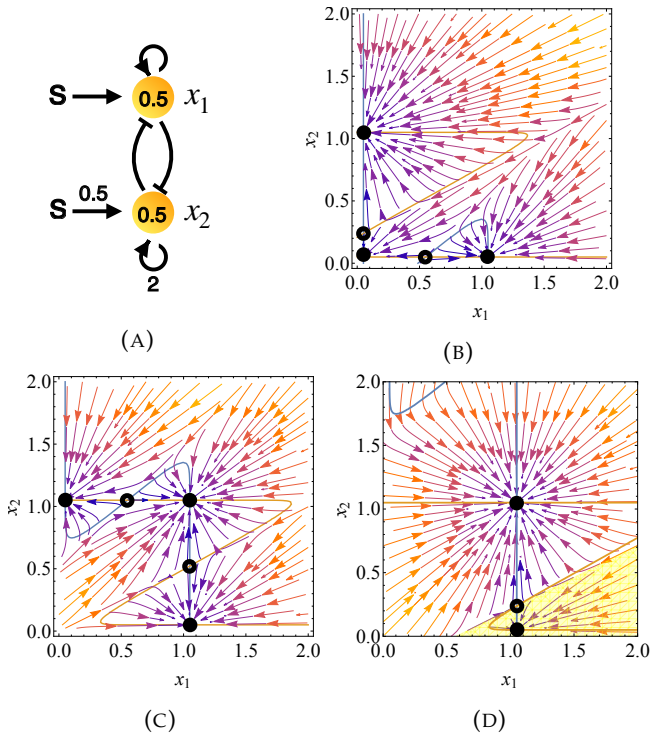
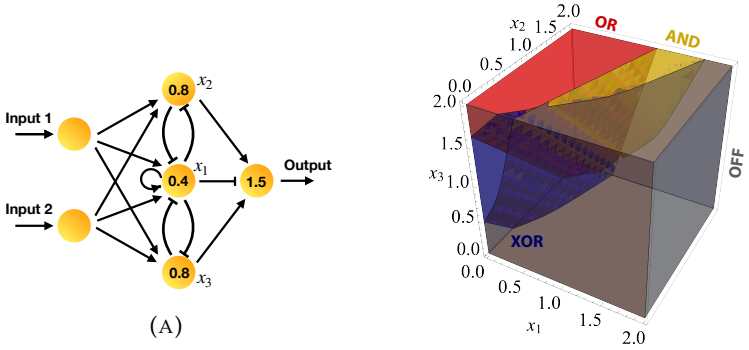


FIGURE 3.3: **The ANDOROFF gate. (A)** The topology of the ANDOROFF gate.**(B),(C),(D)** The phase portrait of the bistable motif shown in (A) with $s = 0$, $s = 1$ and $s = 2$, respectively. The black filled and hollow circles show the stable fixed points and saddle point, respectively. Moreover, the shaded area in (D) shows the basin of attraction of the lower fixed point.



(A) (B)

FIGURE 3.4: **The XORANDOROFF gate. (A)** A multistable circuit which can dynamically switch among XOR, AND, OR and OFF logical functions. **(B)** The regions of phase space corresponding to each function. The colors yellow, red, gray, and blue correspond to the function AND, OR, OFF, and XOR, respectively.

output is strong enough to turn it ON. However, when both inputs are ON all three genes (x_1, x_2 and x_3) will be expressed and expression of x_1 inhibits the expression of output. Therefore, if the system is initially located in the blue region it performs XOR function. Similarly, locating the system in other regions enables it to act as other gates and to perform the corresponding functions.

3.4 Switchable binary adder/subtractor: an example

After introducing our novel dynamically switchable logic gates, let us demonstrate the applicability and power of this framework for reducing the number of required elements. We design a circuit that is capable of performing binary addition as well as binary subtraction [112]. The traditional static circuits for either of these functions are depicted in Fig. 3.5a and 3.5b.

As one can see in Fig. 3.5a, an adder circuit requires XOR and AND gates while a binary subtractor shown in Fig. 3.5b needs XOR, NOT and AND. Being able to perform these two functions (i.e. addition and subtraction) in a traditional static framework, requires having these circuits in the system and rerouting signal (e.g. using a controller circuits) to the desired one when needed. However, by utilizing context-dependency and reusing of the components, one can perform these computations using only three logic gates: AND, XOR and NOT1+AND. In the previous section, we already demonstrated that the XORANDOROFF gate having three nodes in the intermediate layer is able to perform AND, XOR and more. By breaking the symmetry of the inputs to the intermediate layer and adjusting the other parameters, one can modify this circuit into the one shown in Fig. 3.5c. This circuit is able to perform all three logic functions required for binary addition and subtraction. This circuit is also able to perform the OR operation, act as an OFF switch or just simply reflect input 2 as its output. The basins of attraction of these functions are shown

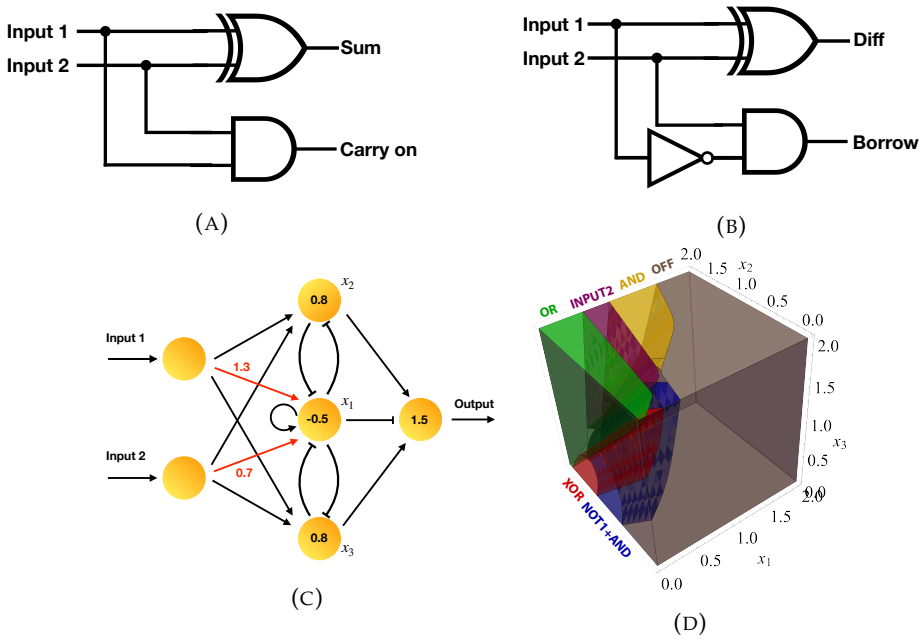


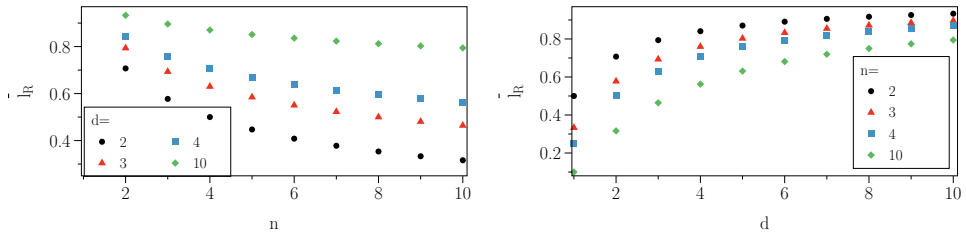
FIGURE 3.5: **An Example: Adder/subtractor circuits**. Traditional binary **(a)** adder and **(b)** subtractor circuits based on static components. **(c)** Dynamically switchable binary adder/subtractor circuit. **(d)** Initiating the system from each region shown here, enables it to perform a distinct function. Starting from green results in OR function, from red results in XOR, from blue performs NOT1+AND, from yellow does AND and if starting from grey the output is always OFF. Finally, starting from the purple area, the output simply reflects Input2.

in Fig. 3.5d. Similar to the ANDOR gate discussed in Sec. 3.3.1, switching among different functions in this circuit can be achieved by various means.

Using the proposed binary adder/subtractor reduces the number of required elements and it is therefore more efficient in this sense. However, it may sacrifice the speed of computation since there's only one output from the system. This means that only either Sum or Carry (in binary adder) can be computed at a time and in the subtractor circuit, either Difference or Borrow. In order to circumvent this drawback, one may alter the circuit structure and parameters to get both outputs simultaneously, but this is beyond the scope of our study here, since we focus on maximum multifunctionality and reusibility of the elements.

3.5 Robustness of dynamically switchable logic gates

As mentioned earlier, multifunctionality of information processing circuits enables them to perform more functions than their static analogs. Unfortunately, this also makes them susceptible to errors as unwanted switching to undesired functions may occur due to various noise types. In this section, we discuss how different uncertainties can affect reliability of functions in dynamically switchable logic gates and how



(A) (B)

FIGURE 3.6: **The average resilience length of general multifunctional systems. (A)** vs. number of functions n at various dimensions d, and **(B)** vs. dimension d with different number of functions n.

one can quantitatively characterize these definitions of reliability. We start by considering uncertainty in the initial conditions of the system. Then, we discuss uncertainties in the dynamics parameters, and finally, using the ANDOR gate we discuss robustness against intrinsic noise due to low copy number of chemical species and random timing of their reactions.

3.5.1 Robustness against uncertainty in initial conditions

Let us consider a d-dimensional system (i.e. d components in the intermediate layer of of Fig. 3.1) that performs n distinct functions. Also, assume that the relevant phase space is a hypercube with sides of length L resulting in an overall hypervolume of L^d. The system is supposed to be initialized at the basin of attraction of the desired fixed point, but it may miss that due to uncertainties in initial conditions. This can, for instance, be the result of temporal fluctuations in the upstream processes or of spatial fluctuations during cell division. In any case, a larger basin of attraction results in higher resilience. In order to have a measure to compare systems with different dimensions, one can define the average resilience length as d-th root of the average size of basins of attraction i.e. $\bar{l}_R \equiv \frac{L}{\sqrt[d]{n}}$. Therefore, for a given number of functions n, higher dimension d results in higher resilience meaning that the system can better withstand uncertainties in initial conditions.

In Fig. 3.6a, one can see the average resilience length $\bar{l}_R$ vs. the number of functions n which decreases faster at lower dimensions d. This indicates that in order to have higher resilience against uncertainties in initial conditions, higher dimension is required, especially when a high number of functions is required. Additionally, Fig. 3.6b depicts the average resilience length $\bar{l}_R$ vs. dimension of the intermediate layer d with different number of functions showing how decreasing the dimension results in decrease in the resilience length particularly for greater number of functions n.

We furthermore calculate the size of the basins of attraction, numerically, for all functions of the ANDOR, ANDOROFF and XORANDOROFF gates. We employ a simple Monte Carlo method to calculate the size of the parts of the phase space from which our proposed gates perform each function. For the two-dimensional cases (ANDOR

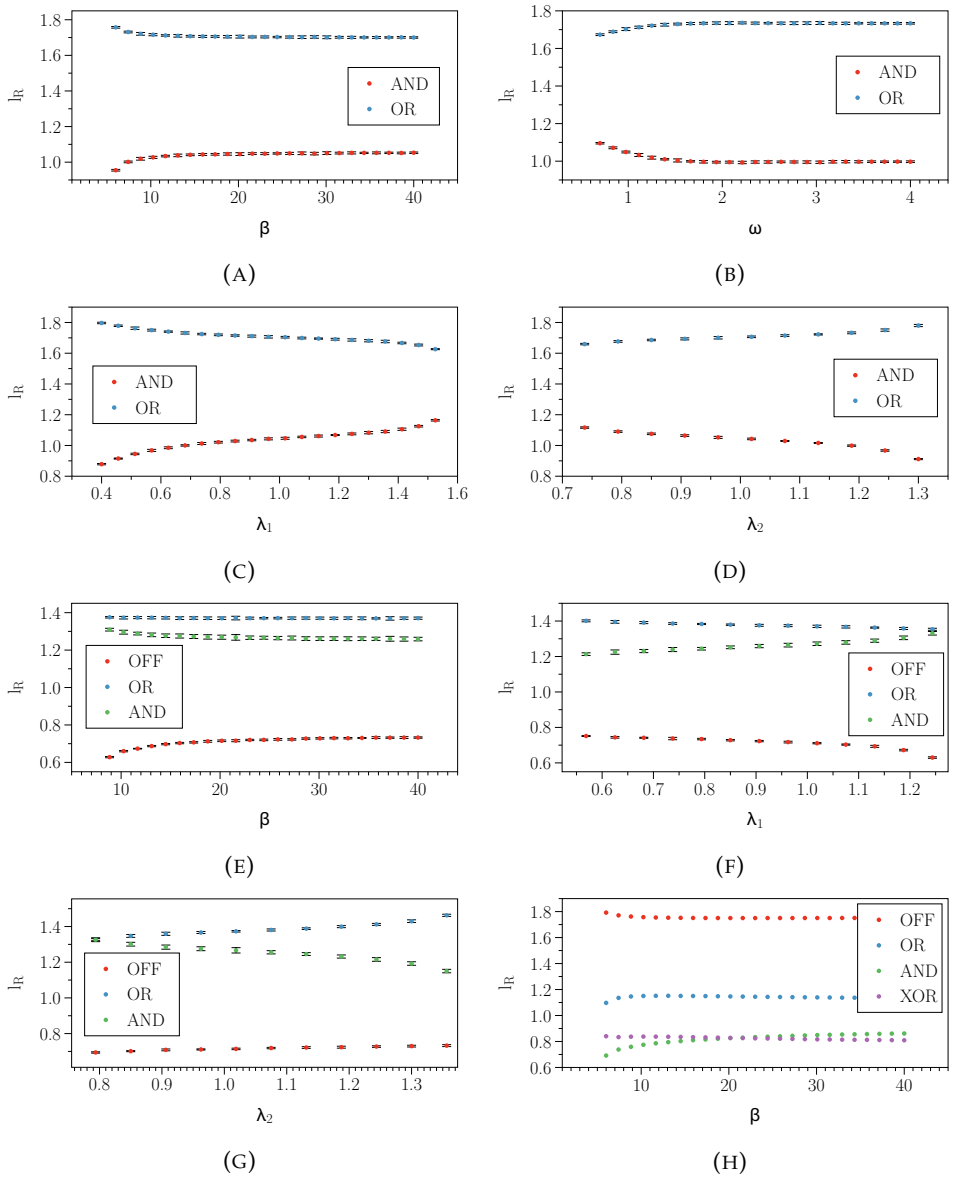
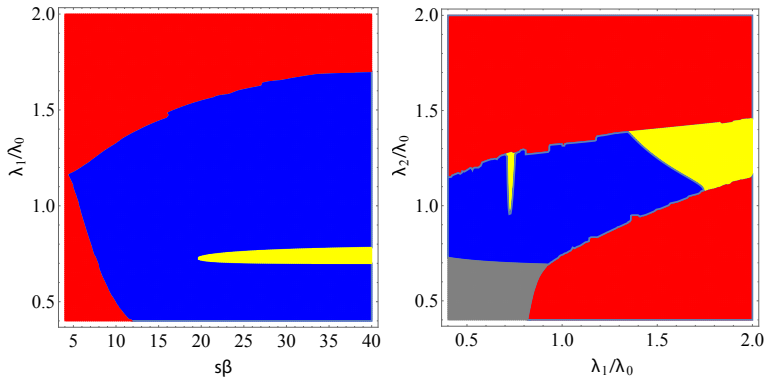


FIGURE 3.7: **The resilience length of each function vs. dynamics parameters of the three proposed dynamically switchable logic gates. (A–D)** The resilience length of the ANDOR gate functions vs.β, ω, λ_1, and λ_2 respectively. **(C–G)** The resilience length of the ANDOROFF gate's functions vs. β, λ_1, and λ_2, respectively.**(H)** The resilience length of the XORANDOROFF gate's functions vs. β.



(A) (B)

FIGURE 3.8: **The phase diagram of the ANDOR gate** described by Eqs. 1 and 2 in **(a)** $s\beta$– λ_1/λ_0 and **(b)**λ_1/λ_0– λ_2/λ_0 planes. Blue color shows the region in which system meets all requirements while red area shows where it does not have bistability. When the bistability exists, yellow shows where the sum of signals to the output does not meet its threshold, and gray shows where the output turns ON without any input (i.e. $S = 0$). Note that in both phase plots the parameter range for which the network operates as an ANDOR gate is significant.

and ANDOROFF gates) we draw 10^6 random samples and for the XORANDOFF gate 2×10^7 from the phase space with uniform distribution for every parameter value and input levels. We then run the dynamics to find the area (for 2d cases) and the volume (for the 3d case) from which each function is performed. Finally, we calculate the resilience length l_R that is presented in Fig. 3.7.

3.5.2 Robustness against uncertainty in dynamics parameters

Cell-to-cell variability in the parameters of information processing circuits can also result in errors. Also, fluctuations in the upstream processes controlling the production of the components can have similar effects. Here, like the previous section, one can consider the parameter space of the system and find the sub-volume in which the system performs all the functions properly. Then it is possible to again simply define a resilience length. However, it should be noted that since different parameters of the system have different dimensions, one needs to construct the phase plots of the system with respect to the non-dimensionalized set of parameters.

Exploring the phase diagrams of a system is also important for understanding its potential applications and practicality. Fig. 3.8a shows the different regions in the dimensionless parameter space of the $s\beta$ and $\frac{\lambda_1}{\lambda_0}$ plane when $\frac{\lambda_2}{\lambda_0}$ is set to 1. The blue area shows the parameter combinations which satisfy all conditions required for performing AND and OR functions. The yellow part shows the region in which the system at the right fixed point does not send enough signal to the output (i.e. $x_1 + x_2 < 1.5$) making the system unable to perform the AND function. Finally, the red region is where the bistability does not occur. Similarly, we also determine

the right combination of $\frac{\lambda_1}{\lambda_0}$ and $\frac{\lambda_2}{\lambda_0}$, as shown in Fig. 3.8b for $s\beta = 20$. Here, each color to the same situation as as in Fig. 3.8a,and the gray color shows the region where degradation for both x_1 and x_2 is so low that even without any inputs, their steady-state concentrations meet the output threshold turning it ON. As one can see in these figures, there is a robust range of parameters for which this circuit behaves as a context-dependent logic gate switching between AND and OR.

3.5.3 Robustness against intrinsic noise

Low copy number of species and consequently the random timing of reactions cause intrinsic noise in the dynamics of cells' regulatory network, in general. In multi-functional circuits, these temporal fluctuations allow for undesired noise-induced transitions that reduce the reliability of the decisions. In order to study these transitions, one can employ the theory of large deviations constructed by Freidlin and Wentzell [17] described as following. Consider a stochastic system whose dynamics is governed by a Langevin equation as

$$\dot{\mathbf{X}}(t) = b\left(\mathbf{X}(t)\right) + \varepsilon\sigma\left(\mathbf{X}(t)\right)\dot{\mathbf{W}}(t), \tag{3.5}$$

where $\mathbf{X}(t)$ is the n-dimensional state vector of the system at time t, $b\left(\mathbf{X}\right)$ is the drift term (i.e. deterministic part of the dynamics). The second term on the right hand side is the diffusion term which is composed of $\mathbf{W}$ an m-dimensional uncorrelated Wiener process, $\sigma\left(\mathbf{X}(t)\right)$ is the standard deviation of multiplicative noise and ε which is a small parameters determines the noise strength. For a given path φ in phase space that starts at $t = 0$ and stops at $t = T$, the Freidlin-Wentzell action can be written as:

$$S(\varphi) = \frac{1}{2}\int_0^T \sum_{i,j} a_{ij}\left(\varphi_t\right)\left(\dot{\varphi}_t^i - b^i\left(\varphi_t\right)\right)\left(\dot{\varphi}_t^j - b^j\left(\varphi_t\right)\right)\mathrm{d}t, \tag{3.6}$$

in which i and j go through dimensions of the system. Also, a_{ij} are elements of $A(x) \equiv \left(\sigma(x)\sigma^*(x)\right)^{-1}$. Then, the probability of the path φ to be taken by the system due to noise is proportional to $\exp\left(-S(\varphi)/\varepsilon^2\right)$. Note that when ε is small enough, all paths from a given point to another will have negligible probabilities compared to the one which minimizes the action in $S(\varphi)$. This path, called the Minimum Action Path (MAP), determines the trajectory in the phase space with the highest probability for a given transition at the given time period T. We employ this theory for studying characterize noise-induced transitions in our circuits.

Due to the random timing of chemical reactions in biological systems, the concentrations of species follow stochastic dynamics. In the case of well-mixed systems, as discussed earlier, one can use the chemical Langevin equation [55] to fully describe the dynamics when fluctuations are sufficiently small (i.e. the reaction volume is large). For the ANDOR gate whose deterministic dynamics is described by Eq. 3.3

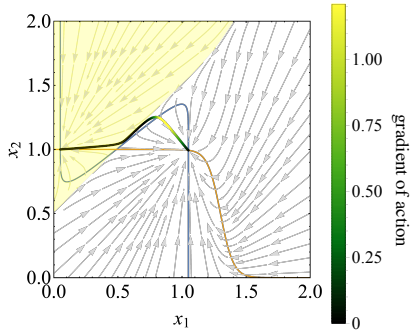


FIGURE 3.9: **The MAP for the ANDOR gate.** The phase portrait of the ANDOR gate shown in Fig. 3.2 along with the MAP which connects the central fixed point to the left one.

and 3.4, the chemical Langevin equation is as following:

$$\dot{x}_1 = F(x_1, x_2) - \lambda_1 x_1 + \Omega^{-1/2}\left(\sqrt{F(x_1, x_2)}\dot{W}_{11} + \sqrt{\lambda_1 x_1}\dot{W}_{12}\right) \tag{3.7}$$

$$\dot{x}_2 = G(x_1, x_2) - \lambda_2 x_2 + \Omega^{-1/2}\left(\sqrt{G(x_1, x_2)}\dot{W}_{21} + \sqrt{\lambda_2 x_2}\dot{W}_{22}\right). \tag{3.8}$$

Here, $F(x_1, x_2)$ and $G(x_1, x_2)$ are the production terms in Eq. 3.3 and 3.4. Moreover, Ω is the reaction volume and Ws are Wiener processes each of which arises from the corresponding reaction. Since there is no reaction coupling fluctuations of two species x_1 and x_2 (i.e. their stochastic dynamics do not share any noise term $\dot{W}$) all the non diagonal terms of the matrix A vanish and it simply reads

$$A(x_1, x_2) = \begin{bmatrix} \frac{1}{\lambda_1 x_1 + F(x_1, x_2)} & 0 \\ 0 & \frac{1}{\lambda_2 x_2 + G(x_1, x_2)} \end{bmatrix}. \tag{3.9}$$

Therefore, the Freidlin-Wentzell action for the ANDOR gate simply becomes

$$\begin{aligned} S(\varphi) =& \frac{1}{2}\int_{T_1}^{T_2} \frac{\left(\dot{\varphi}_1 - [-\lambda_1\varphi_1 + F(\varphi_1, \varphi_2)]\right)^2}{\lambda_1\varphi_1 + F(\varphi_1, \varphi_2)} \mathrm{d}t \\ &+ \frac{1}{2}\int_{T_1}^{T_2} \frac{\left(\dot{\varphi}_2 - [-\lambda_2\varphi_2 + G(\varphi_1, \varphi_2)]\right)^2}{\lambda_2\varphi_2 + G(\varphi_1, \varphi_2)} \mathrm{d}t \end{aligned} \tag{3.10}$$

in which φ_1, φ_2 are the coordinates of the path φ at any time $t \in [T_1, T_2]$.

Minimizing the action in Eq. 3.10 over the function space containing all paths which connect the point X_0 at time $t = 0$ to X_T at time $t = T$ provides the Minimum Action Path, and then, the minimum action value can be used for calculation of the rate of that transition. The MAP which connects the right fixed point at $(1, 1)$ to the left one at $(0, 1)$ for a typical set of parameters is shown in Fig. 3.9. The color on the path shows the gradient of the action at any given point. This gradient can be interpreted as the effective force exerted by the noise for causing the movement along the MAP. As one can see in this figure, the MAP goes directly towards the

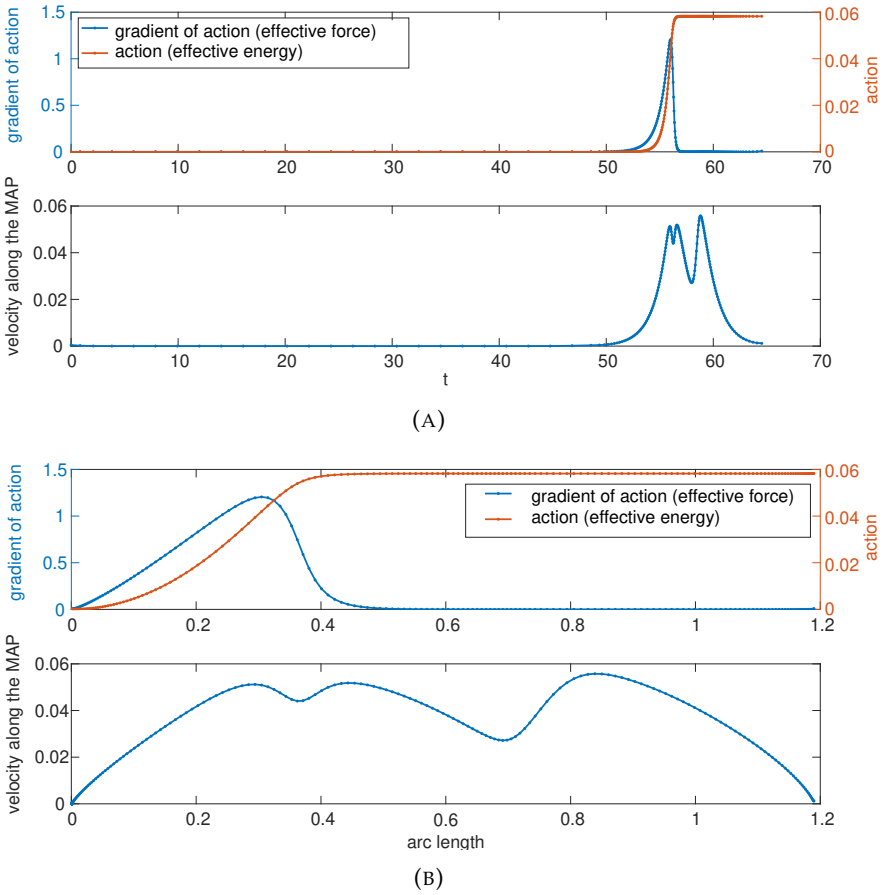


FIGURE 3.10: **Action properties along the MAP for the ANDOR gate shown in Fig. 3.9. (A)** and **(B)** shows the action, its gradient and velocity vs. time and arc length along the MAP, respectively.

separatrix in the opposite direction of stream lines. Then, it follows the streamlines getting close to the separatrix until it approaches the saddle point. Finally, it enters the other basin of attraction and follows those streamlines to the left fixed point. The observation that the MAP for the ANDOR gate crosses the separatrix close to the saddle point is consistent with the findings of other studies in different contexts [16]. It should be noted that when the system undergoes a transition from one fixed point to another, it spends most of the time at the fixed points and a small fraction of total time will be spent on the actual transition. This is evident from Fig. 3.10a which shows the action, its gradient and velocity of the transition along the MAP shown in Fig. 3.9. Therefore, in order to get an acceptable accuracy, one needs to use an adaptive minimum action path method in which the discretization of time is adaptively adjusted based on the speed (see App. B.2). The performance of the adaptive meshing is demonstrated by plotting the these quantities vs. the arc length along the MAP in Fig. 3.10b.

The rate of noise-induced transitions from the desired fixed point to another is a proxy for the reliability of the decisions. Therefore, we investigate this reliability (i.e. resilience against intrinsic noise) of the system as the parameters change. In

order to find the transition probability from one stable fixed point to another, one may intuitively expect to do this procedure for the transitions from every point in one basin of attraction to the fixed point of the other basin. However, it is shown that for small noise strength ε, all trajectories which leave a basin of attraction due to the fluctuations will visit a small neighborhood around the stable fixed point before leaving the basin [17]. Therefore the transition from one fixed point to the other represents the dominant transition trajectory and suffices for determining the resilience against intrinsic noise. We thus use this measure for studying the reliability of the decisions of the ANDOR gate.

Fig. 3.11a shows the action vs. sharpness of transition $s\beta$ for the transitions from the left fixed point to the right one (in black circles) and vice versa (in red triangles). As one can see in here, as β increases, the action for both transitions increase, but for the L$\rightarrow$R transition increases faster which results in a critical value at $s\beta^* = 14.75$. The dependence of the action $S(\varphi)$ on the slope of the sepratrix ω, degradation rate λ_1 of x_1, and that of x_2 are shown in 3.11b, 3.11c, and 3.11d, respectively. The critical values at which the action for L$\rightarrow$R transition equals R$\rightarrow$L are $\omega^* = 0.87$, $\lambda_1^*/\lambda_0 = 0.97$ and $\lambda_2^*/\lambda_0 = 1.01$. Note that in all these figures, unless a parameter is variable, we fix them as following: $s\beta = 20$ and $\lambda_1/\lambda_0 = \lambda_2/\lambda_0 = \omega = 1$. Note that although higher values of S mean higher reliability, any difference between the actions for two transitions results in a biased error. Therefore, one can set these parameters to their critical values in order to equalize the transition probability and minimize the bias. Rapid switching would be achievable in this regime. On the other hand, parameters consistent with one strongly stabilized basin at the expense of the other are attainable.

3.6 Conclusion and outlook

In this chapter we designed three dynamically switchable logic gates in which mulistability *emerges* from interaction between components, and their dynamics mimic that of gene regulatory system or any signaling pathways of living cells with switch-like production rate. These gates can perform two, three and four distinct functions based on the basin of attraction in which the system resides in or is steered to. We introduced three noise types and defined robustness against every one of them. For the uncertainty in initial condition of the system we showed a trade-off between multifunctionality and robustness which can be overcome by increasing the dimension of the system (e.g. by adding more regulatory components). For the uncertainties in dynamics parameters (extrinsic noise), we demonstrated how to construct phase diagram for the systems with respect to non-dimensionalized parameters. Finally, using the theory of large deviations, we characterized noise-induced transitions between the possible logical functions and determined the reliability of the decisions. We therefore demonstrated resilience of the proposed gates against three types of uncertainty: uncertainty in initial conditions, in dynamics parameters, and in the

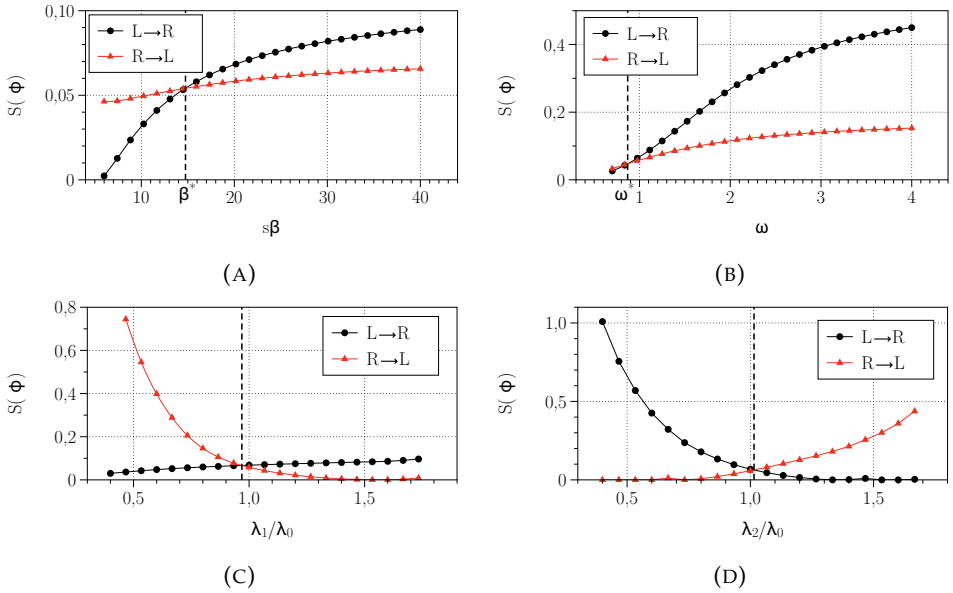


FIGURE 3.11: **The minimum action for noise-induced transition** vs. **(A)** the dimensionless sharpness of expression $s\beta$, **(B)** the slope of separatrix ω, **(C)** the degradation rate λ_1/λ_0, and **(D)** degradation rate λ_2/λ_0. Black circles show the action for the transition from the right fixed point the left one and the red triangles show the action for the opposite transitions. For each dependency a critical value at which the cost of transiting L→R and R→L is balanced can be attained and is indicated by dashed line and have the following values: $s\beta^* = 14.75$, $\omega^* = 0.87$, $\lambda_1^*/\lambda_0 = 0.97$ and $\lambda_2^*/\lambda_0 = 1.01$.

dynamics. Finally, to demonstrate generality and applicability of our framework, we provided an example in which a more complex circuit could carry out more complex computations dynamically switchable, otherwise requiring multiple logic gates with many more components.

To construct high-order information processing circuits, the existence of memory is a necessary feature. In the engineering applications, sequential logic circuits are widely used in which the output depends not only on the current input but also on the history. In the context of biological information processing, memory can be attained by addition of a toggle switch to an existing combinational (i.e. memoryless) logic gate [113]. In our setup, however, the bistability that results in switchability plays an additional role by capturing a memory of the last action without needing the addition of an extra toggle.

We believe that our work here is only the first steps towards a comprehensive understanding of how network topology, dynamics, and information processing can combine in flexible and non-trivial ways in the context of gene regulatory networks, signaling networks, and (bio-)chemical computations. Many extensions and generalizations are to be explored. Here, we have assumed that these biochemical systems are in a well-mixed reaction volume, but spatial localisation, compartmentalisation and even fluctuations are more and more appreciated as being important players

in cell biology [114, 115, 116]. These phenomena in principle have significant effects on the operation of these switchable information processing elements. Another important direction of extension of framework is considering how families of more complex calculations constructed from the ANDOR gate and others can be coupled to adaptive pressures on evolutionary timescales. This can provide a framework analog to deep learning but in the context of bio-chemical computations. Finally, of course, *in vitro* implementation of these gates is an important first step towards any real-world application of our framework.

Chapter 4

Fractal fluctuations as a tool for characterizing patterns

Development of tissues during embryonic stage typically involves cell division, cell differentiation and material secretion. How biology coordinates these processes to make tissues robustly despite the temporal and spatial fluctuations present at various scales is a crucial and interesting problem. The complexity of this problem further increases when proliferation, differentiation and material deposition occur simultaneously. Development of flat bones such as skull caps is one of such cases that are not well-studied. During these processes, non-local and non-linear feedbacks are employed to overcome noise and maintain important tissue properties. Therefore, studying such processes requires advanced non-local methods to characterize the patterns as well as generative models to explore the mechanisms responsible for the observed features. In this chapter, we employ Multi-Fractal Detrended Analysis (MFDFA) to robustly characterize ossification patterns in developing skull caps of mouse embryos based on their singularity spectra. We then produce various surrogate data sets to unveil the origin of observed multifractality. We next use Wavelet Transform Modulus Maxima Maxima Method (WTMMMM) to investigate the spatial distribution of singularities in the data. Finally, using a simple generative model we explore the mechanisms involved in ossification process that are needed to produce multifractal features. All the experimental data used in this chapter are courtesy of our collaborators in Tabler lab at MPI-CBG.

4.1 Introduction

Multicellular organisms are composed of many cells of various types, and the development of such organisms involves many processes such as cell proliferation, cell differentiation, migration and deposition of non-living organic and inorganic materials. To better control the tissue development, biology has chosen to take these actions in a sequential manner, most of the time. Occasionally, however, they occur simultaneously rendering a robust development nearly impossible without coordination across time- and length-scales especially in the presence of temporal and

spatial noise [117, 118]. A typical hallmark of such an orchestration is complex spatial patterns whose fluctuations (spatial inhomogeneities) carry crucial information about the underlying mechanisms [119, 120, 121]. The first step of studying these processes is characterizing the patterns which requires advanced cross-scale methods.

The concept of *statistical self-similarity* is proven useful for quantifying patterns formed by complex systems composed of many elements interacting through multiple scales [122]. A rough definition of self-similarity is having the same statistical behavior at various scales. This has been applied to a broad class of pattern formation systems whose details are wildly different such as diffusion limited aggregates, the Laplace equation, etc. They were initially assumed to be *monofractal* meaning that all statistics of their structures can be described with a single exponent [123, 124, 125]. However, later studies showed that many structures in nature, including the aforementioned ones, require an infinite hierarchy of exponents to be fully described, i.e. they are *multifractal* [126, 127, 128]. Understanding this multifractal complexity could unveil deeper links between structure and function across scales.

Bone is a prime example of complex pattern formation in biology as the nanoscale structure of extracellular matrix laid down by the cells determines the coarse-grained material properties of these skeletal elements [129]. Here, mineralisation is the key mechanism linking processes across many spatial scales [130]. To be more specific, carbonate-substituted hydroxyapatite crystallises on collagen fibrils and generates the bone's ultimate form and mechanical properties [131, 132, 133]. As mineralization depends on molecular interactions and geometric constraints, mineral density and morphology are great proxies for studying bone formation in wild type and mutant animals [134]. However, most studies of these patterns use mono-fractal tools relying solely on linear correlations (e.g. Fractal Dimension). Additionally, they do not take into account the development of these patterns during which the shape and form evolves. Bones of the skull vault indeed undergo anisotropic expansion of a mineralised tissue from the lateral side of the forebrain toward its apex (i.e. midline) [135]. Instead of a continuous layer, the mineral forms a fine lace-like meshwork that thickens and fills over the course of development. This gradual and hierarchical construction of bone plates is a robust evolutionary innovation that allows the skull vault to accommodate continued expansion of an underlying brain [136, 137, 138]. However, this also results in extra complexity which is why despite the importance of skull morphogenesis, how mineral pattern emerges remains largely unexplored.

In this chapter, we demonstrate the utility of two multi-fractal analyses, namely the Multi-Fractal Detrended Fluctuation Analysis (MF-DFA) and the Wavelet Transform Modulus Maxima Maxima Method (WTMMMM) to describe mineralization patterns of the developing mouse skull vault. More specifically, we study the combination of parietal and frontal bones of each hemisphere as shown in Fig. 4.1a which depicts the schematic of this tissue, its relative location, and development between

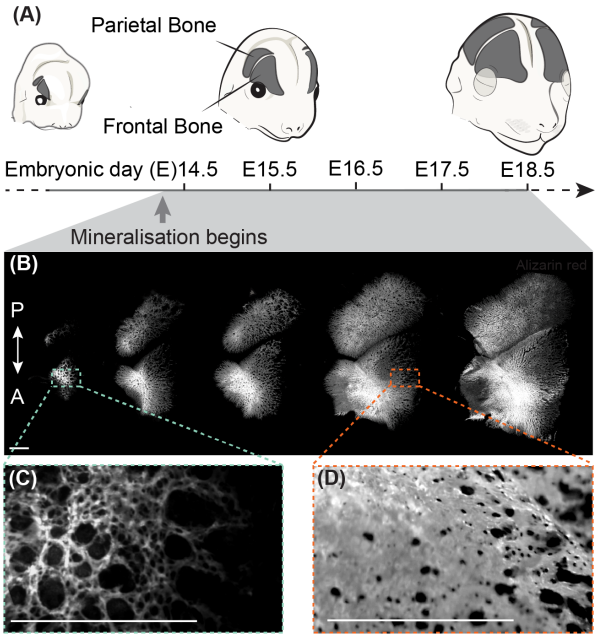


FIGURE 4.1: **Schematic of the ossification data. (A)** Schematic of bone tissues used in our study. **(B)** Typical images of ossified tissue at embryonic days 14.5,15.5, 16.5, 17.5 and 18.5. **(C)** and **(D)** magnified pieces of images from E14.5 and E17.5, respectively.

Embryonic day 14.5 (E14.5) and E18.5. Our data-set comprises images of Alizarin red-stained mineral within flat-mounted embryonic skull caps. Therefore, higher intensity in our data corresponds to higher density of minerals. We analyze the images acquired from different hemispheres separately as two independent samples of our ensemble. Fig. 4.1b, shows one typical representative from every stage included in our study. As one can see here, as the embryos develop, the ossified part of their skull grows. Also the voids (i.e. low intensity parts) get filled in. Using MF-DFA, we find that deposited mineral indeed has multifractal features i.e. a single exponent is not sufficient to fully describe the statistical self-similarity. This mulitfractality can in turn be used to characterize emergent patterns. We can also investigate its origin by employing data surrogation techniques. With a simple generative model we are able to reproduce ossification patterns observed during normal mouse development and we predict that collagen density is a key regulator of mineral pattern. Finally, we experimentally confirm this hypothesis by demonstrating the sufficiency of nanoscale collagen organisation to generate multifractal mineral patterns during development.

4.2 Multi-Fractal Detrended Fluctuation Analysis in higher dimensions

In past decades, there have been many attempts to develop tools for characterizing multifractality of stochastic time-series which can then be generalized to study

multifractal patterns in higher dimensions (e.g. 2D images). The simplest standard method for studying multifractal features is based on partition function [125, 139] and is not applicable to signals affected by trends or non-normalizable signals. Additionally, in this formalism, the negative order of moments are dominated by the voids (empty parts of the image). These problems can be solved by some variation of the box-counting method in which the signal is covered by equally sized and potentially overlapping boxes. This is then followed by studying the scaling of the statistics of the filled boxes [126]. The Wavelet Transform Modulus Maxima Method (WTMMM) can be considered an extension of box counting methods in which a proper wavelet (instead of simple boxes) is chosen depending on the properties of the system. Then detection of singularities is done by tracking the local maxima of the wavelet transform modulus throughout different scales. In this method, the chains of local maxima can also be used for constructing the partition function and studying multifractal properties [140]. However, for small data sets, it is shown that the measured multifractality by WTMMM deviates more from actual values, compared to a more recent method called Multi-Fractal Detrended Fluctuation Analysis (MF-DFA) [141]. This method was introduced by Kantelhardt et al. [142] in the context of 1D time-series, and later, was generalized by Gu and Zhou [143] for higher-order applications.

One-dimensional MF-DFA comprises a few simple steps which can be easily generalized to higher-dimensional data-sets such as conventional images or three-dimensional tomographic scans. We here focus on two-dimensional MF-DFA whose simple steps are as following:

- **Step 1.** Divide the surface under study (i.e. the intensity of the image) given by a two-dimensional array $X(i,j)$ with size $N \times M$ into $N_s \times M_s$ disjoint boxes of size $s \times s$ (see Fig. 4.2a). Here, $N_s = \left[\frac{N}{s}\right]$ and $M_S = \left[\frac{M}{s}\right]$. Each segment can be represented by $X_{v,w}(i,j) = X\left((v-1)\,s+i,(w-1)\,s+j\right)$ where $i,j = 1,2,\ldots,s$, $v = 1,2,\ldots,N_s$ and $w = 1,2,\ldots,M_s$.

- **Step 2.** Calculate the cumulative sum $Y_{v,w}$ for each segment $X_{v,w}$ via the following relation:

$$Y_{v,w}\,(i,j) = \sum_{h=1}^{i}\sum_{l=1}^{j} X_{v,w}(h,l), \tag{4.1}$$

  where $i,j = 1,2,\ldots,s$. In fig. 4.2a, one can see the cumulative sum shown by red points.

- **Step 3.** Detrend the cumulative sum surface in each segment $Y_{v,w}$ by first fitting a bivariate polynomial function $\tilde{Y}_{v,w}\,(i,j)$ and then obtaining the residuals

$$\epsilon_{v,w}\,(i,j) = Y_{v,w}\,(i,j) - \tilde{Y}_{v,w}\,(i,j)\,. \tag{4.2}$$

Then, calculate the "detrended fluctuation" at segment v, w as

$$F^2_{v,w}(s) = \frac{1}{s^2} \sum_{i=1}^{s} \sum_{j=1}^{s} \epsilon_{v,w}(i,j)^2 . \tag{4.3}$$

In fig. 4.2a, the polynomial fit to the cumulative sum is shown by the solid plane, the residuals are shown by dots above.

The procedure of choosing the order of polynomial fitted in eq. 4.2 will be discussed in App. C.1

- **Step 4.** Calculate the q-th order overall *fluctuation function*

$$F(q,s) = \left[\frac{1}{N_s M_s} \sum_{v=1}^{N_S} \sum_{w=1}^{M_S} \left[F^2_{v,w}(s) \right]^{\frac{q}{2}} \right]^{\frac{1}{q}} . \tag{4.4}$$

Fig. 4.2b, depicts $F(q,s)$ vs. s for a few values of q. As one can see here, in the selected range of s, $F(q,s)$ vs. s shows power-law behavior whose exponent depend on q.

- **Step 5.** Determine the scaling behavior of the fluctuation function $F(q,S)$ by varying s for each value of q reading

$$F(q,s) \sim s^{h(q)}, \tag{4.5}$$

where $h(q)$ is generalized Hurst exponent. This can be obtained by simply fitting a line to the $log - log$ plot of $F(q,s)$ vs. s for each value of q i.e.

$$h(q) = \frac{\log\left(F(q,s)\right)}{\log(s)}. \tag{4.6}$$

The generalized Hurst exponent for a typical ossification pattern at E15.5 is shown in fig. 4.2c.

It is worth noting that the traditional DFA method can be realized by performing steps 1–3 exactly as mentioned above and the remaining steps with $q = 2$. This way, only the Hurst exponent $H = h(q = 2)$ can be determined. Furthermore, it should be noted that MF-DFA is capable of extracting multifractal characteristics of the data subject to many forms of trends and artifacts except for periodic trends which appear in our data. However, removing these types of trends is straightforward and can be done with the help of Fourier analysis (see App. C.2 for the details). Apart from this simple step, there is no pre-processing required for our imaging data. Analyzing these images, we observe that all of them show some level of multifractality that can be used for characterizing the patterns. Before discussing the multifractal measures, we investigate the origin of multifractality in our data.

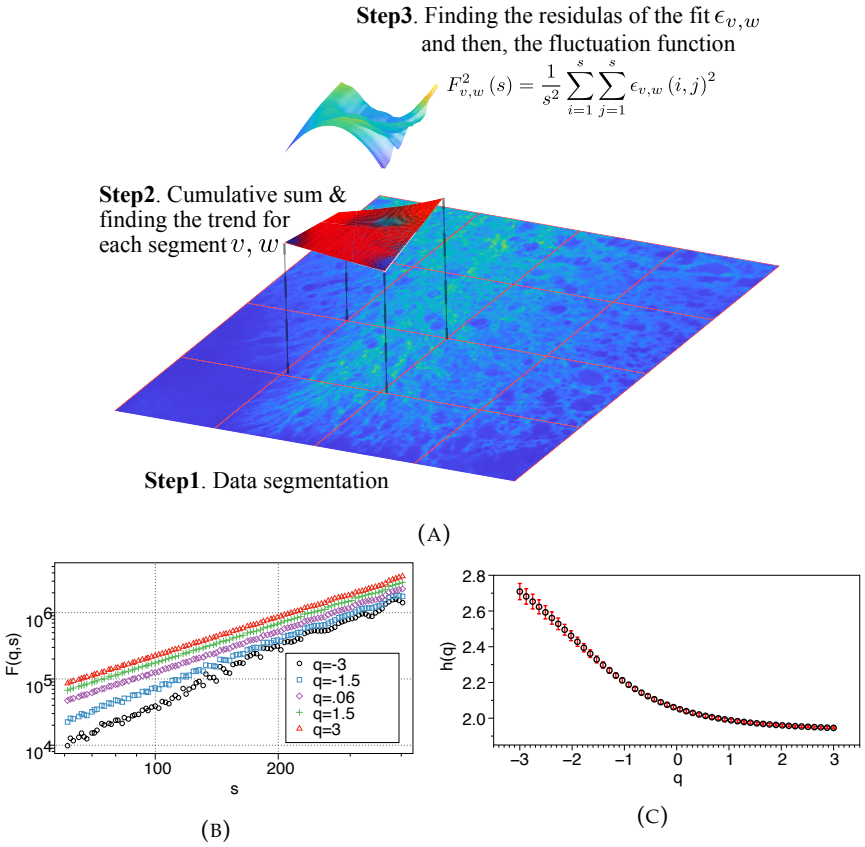
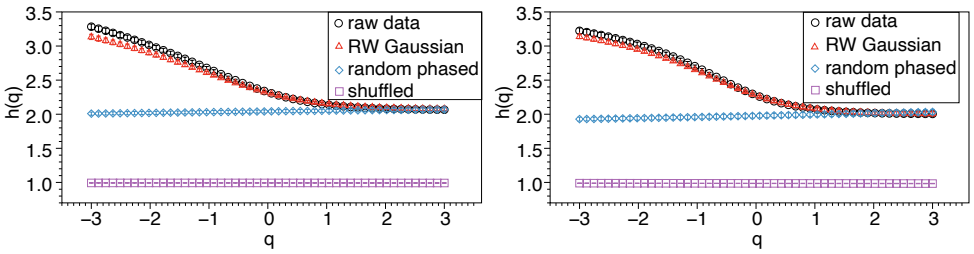


FIGURE 4.2: **The basic steps of MFDFA. (A)** the first three steps of MFDFA. **(B)** The fluctuation function of the data $F(q,s)$ vs. scale s for a few q. **(C)** The generalized Hurst exponent $H(q)$ vs. q.



(A) (B)

FIGURE 4.3: **Investigating the origin of multifractality in our data** from **(A)** E15.5 and **(B)**E16.5

4.2.1 Origin of multi-fractality in data

A data set can exhibit multifractality for various reasons including: I) non-Gaussian (e.g. fat-tailed) distributions, II) existence of linear or III) nonlinear long-range correlations. One can investigate the contribution of each of these properties to the overall multifractality using surrogates which change one or more properties while preserving the others. Here, we first use Ranked-Wise (RW) Gaussian surrogate [144], which preserves all linear and non-linear correlations but changes the distribution of the intensities to Gaussian. We also use random phased surrogate to preserve linear correlations, remove nonlinear correlations, and change the distribution to Gaussian with the help of the central limit theorem. Finally, we shuffle the pixels in the image to preserve the distributions while removing all correlations.

We analyzed two images from 15.5 and 16.5 stages, and as one can see in Fig. 4.3, the closest curve to the $h(q)$ of the original data is the RW-Gaussian surrogate. The one which differs the most is unsurprisingly the shuffled one, which only preserves the distribution. One can accordingly conclude that the non-Gaussian distribution has the least contribution, and the long-range correlations contribute most to the overall multifractality. Moreover, one can consider a random-phased surrogate and argue that the linear correlations set the offset of the generalized Hurst exponent $h(q)$ but do not contribute significantly to the multi-fractality. We therefore established that the observed multifractality originates primarily from *long-range non-linear* correlations in the data.

4.3 Singularity spectrum and multi-fractal measures of the data

Singularity points at which a signal's value or its derivatives are discontinuous carry crucial information. Various statistics of singularities are widely used for data characterization. For example, the abundance and variety of singularities of continuous neural time series can be used to infer the neuron's activity [145]. Moreover, singularity exponents can be used to segment water bodies in optical and synthetic aperture radar (SAR) satellite images [146]. Additionally, in WTMMM [140], instead

of storing and analyzing information over an entire data set, information at singularities alone suffices to construct the partition function and determine many properties of the data.

The generalized Hurst exponent $h(q)$ can be transformed into the singularity spectrum $f(\alpha)$ using the following Legendre transformation:

$$\alpha = h(q) + qh'(q) \tag{4.7}$$

$$f(\alpha) = q^2 h'(q) + D, \tag{4.8}$$

where D is the space dimension and equals to two in our case. Here, $f(\alpha)$ is the Hausdorff dimension of the set of all data points with the given singularity exponent α. In order to develop an intuition about the singularities, it is useful to consider a simple one dimensional singularity. A singularity exponent of a one-dimensional function $X(t)$ at point t_0 is the largest exponent α such that there exists a polynomial $P_n(t)$ of order $n < \alpha$ that satisfies the Hölder condition

$$|X(t) - P_n(t)| = O\left(|t - t_0|^{\alpha}\right) \tag{4.9}$$

for any point t at a small neighborhood of t_0 [147]. It should be noted that having a larger value of α at a singularity means that the function has higher-order derivatives and therefore is more regular. In other words, the singularity exponent at a point determines how regular the function at that point is [140]. Since $h(q)$ is always monotonically decreasing, $f(\alpha)$ is less than or equal to D. The closer $f(\alpha)$ gets to D, the more uniformly the singularity points with strength α are distributed throughout the space.

The singularity spectrum $f(\alpha)$ of a typical image from stage E15.5 is shown in Fig. 4.4a. This spectrum shows a wide range of singularity exponents α with varying Hausdorff dimension that again indicates multifractality. However, in a monofractal data set, there is only one type of singularities with a given exponent α that depends on the type of the fractal, and these singularities are distributed uniformly over space (i.e. with $f(\alpha) = D$). Analyzing all images from multiple developmental stages between 14.5 and 18.5 showed that different measures have two main types of time evolution: (I) constant distribution width, but changing the average in an oscillatory way, and (II) a general narrowing of an initially wide distribution. The former behavior is shown by the $\min(\alpha)$ vs. embryonic day in Fig. 4.4b. This measure characterizes how irregular the most irregular singularities of the patterns are (as discussed in the next section). The second behavior is shown in 4.4c by the width of the spectrum (indicating the variety of singularities in the patterns) vs. embryonic day (age). This behavior indicates that some aspects of the ossification pattern start with higher randomness (or subject-to-subject variability) but are modified and reach to a more robust point over the course of development. In App. C.3, more examples of these two behaviors are presented along with their implications. Note that

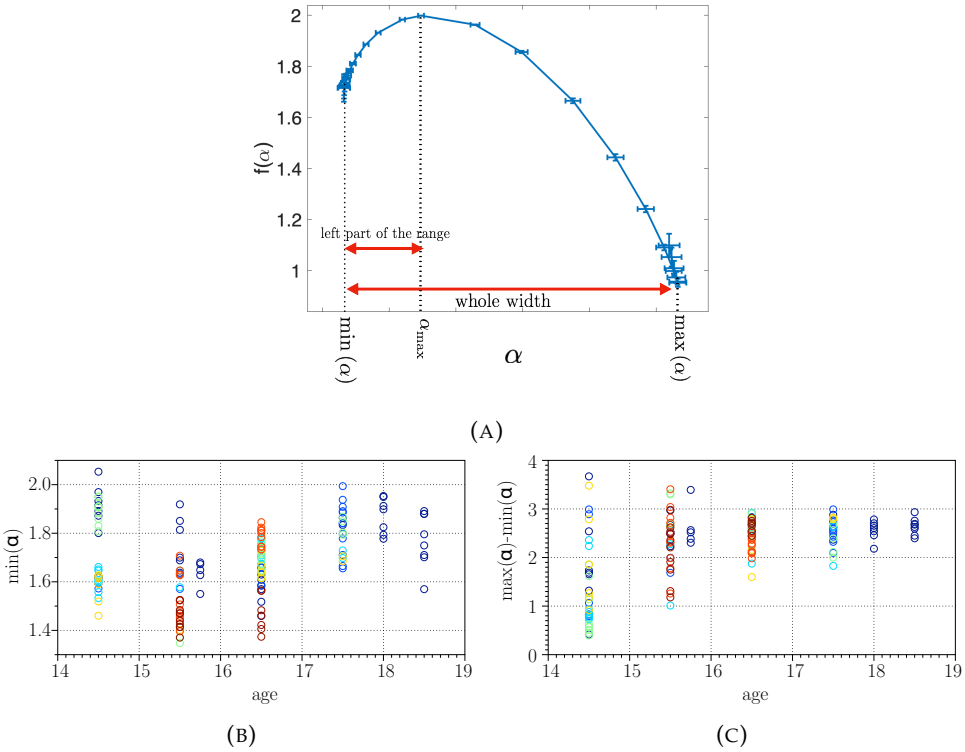


FIGURE 4.4: **The multifractal measures of the ossification data. (A)** A typical singularity spectrum for an image from 15.5D and the few measures of multifractality. **(B)** The minimum of singularity spectrum min(α) that corresponds to the most irregular singularities in the ossification patterns. This measure among some others (as discussed in App. C.3) shows an oscillatory behavior. **(C)** The width of singularity (representing the variety of singularities in the pattern) vs. age. This measure similar to a few others (as discussed in App. C.3) shows a decrease in the variance over time. Different colors in (A) and (B) represent subjects from different litters, and as one can see here, there's little intra-litter variability.

in Figs. 4.4b and 4.4c, each color at a given stage corresponds to a litter of embryos and it is easy to convince oneself that the inter-litter variability is not more than the intra-litter variability.

4.4 Singularities and WTMMMM

Having established that singularities in general carry essential information about the data and how different aspects of their spectrum have potentials for characterization of the data, we now study the spatial distribution of singularities within the images. Let us first sketch the basic ideas of singularity detection on a simple image containing a single isolated singularity as well as a Gaussian local maximum as shown in Fig. 4.5a. In this figure, the height or equivalently, the intensity $f(\mathbf{X})$ of the point

$\mathbf{X} = (x, y)$ is governed by

$$f(\mathbf{X}) = e^{-(\mathbf{X}-\mathbf{X}_1)^2/\left(2\times 64^2\right)} + |\mathbf{X} - \mathbf{X}_0|^{0.3}, \tag{4.10}$$

where $\mathbf{X}_0$ is the singularity point with exponent 0.3 and $\mathbf{X}_1$ is the maximum point (center) of the isotropic Gaussian with width of 64 pixel while the entire surface is composed of 1024×1024 pixels.

The simplest way of detecting a singularity is to define a measure $\mathcal{M}$ over the data at a given scale s, change the scale and study the power-law behavior of the measure at every point. Then the exponents of non-singular points have a trivial value determined by the characteristic behavior of the measure while at singularity points the exponents obtain nontrivial values carrying crucial information about the singularity. For demonstration, we use a sum measure defined as $\mathcal{M}_s(\mathbf{X}) = \sum_{(i,j)\in\Omega} g(i,j)$, where $g(i,j)$ is the grey scale intensity at pixel (i,j) and Ω is a box with size s around the point $\mathbf{X}$. Fig. 4.5b shows the sum measure over entire the image in 4.5a (except for the points close to the boundary) at three scales $S \in \{8, 10, 13\}$. In this figure, three specific points are marked: the red triangles show the singularity at $\mathbf{X}_0$, the blue circles show the center of the Gaussian maximum $\mathbf{X}_1$, and finally the purple squares show a random point within the space. Fig. 4.5c shows the sum measure of those marked points over a wider range of scales s in log-log scale. As one can see here, all three points show a power-law over this range but the exponent at the singularity is 2.3 while at the other points is 2 which is due to the scaling behavior of the sum measure. This simple procedure has many limitations. For example, in practice it requires saving many copies of the data and tracking the measure values at all points which is computationally expensive. Besides it does not provide the precise location of the singularities as all points around it will have nontrivial scaling behavior if their distance is smaller than the scales used. However, this method still captures the basic idea of singularity detection that is shared with more advanced methods such as WTMMMM.

4.4.1 The Wavelet Transform Modulus Maxima Maxima Method

The Wavelet Transform Modulus Maxima Maxima Method (WTMMMM) is based on tracking the gradient maxima of the wavelet transform moduli over a range of small scales. A detailed discussion of this and the necessary steps is provided in Ref. [140]. We here only discuss this method briefly. The first step of this procedure is choosing a proper wavelet $\boldsymbol{\phi}$ depending on the characteristics of the singularities (e.g. isotropy or the biggest singularity exponent) in the data f. The wavelet used in this method works as a lens for detection of singularities which is blind to singularities that have an exponent larger than the order of the wavelet n_ϕ. The order of a wavelet $\boldsymbol{\phi}$ is defined as the number of vanishing moments of its derivative $\boldsymbol{\psi}$ whose

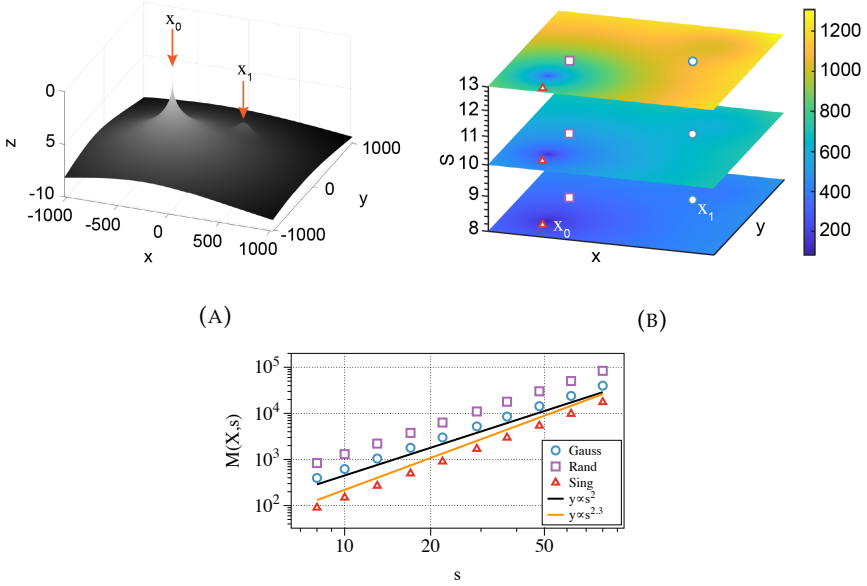


FIGURE 4.5: **A basic singularity tracking procedure. (A)** a simple image which contains an isolated singularity as well as a Gaussian local maximum described by Eq. 4.10. **(B)** The sum measure at three different scales. **(C)** The sum measure of marked points in (B) over a wider range of s in logarithmic scale. The black and orange lines here show two power-laws with exponents equal to 2 (typical scaling behavior of the measure) and 2.3 (corresponding to singularities) respectively.

components are defined as

$$\psi_1(\mathbf{x}) = \frac{\partial}{\partial x}\phi(\mathbf{x}) \tag{4.11}$$

$$\psi_2(\mathbf{x}) = \frac{\partial}{\partial y}\phi(\mathbf{x}). \tag{4.12}$$

In this study, we determine the range of singularities via MFDFA method. Therefore, we can choose the wavelet accordingly. Next, we perform the wavelet transform as following:

$$\mathbf{T}^{\boldsymbol{\phi}}[f](\mathbf{b}, s) = s^{-2} \begin{pmatrix} \frac{\partial}{\partial b_x}\left[\int \mathrm{d}^2\mathbf{x}\phi\left(a^{-1}(\mathbf{x}-\mathbf{b})\right) f(\mathbf{x})\right] \\ \frac{\partial}{\partial b_y}\left[\int \mathrm{d}^2\mathbf{x}\phi\left(a^{-1}(\mathbf{x}-\mathbf{b})\right) f(\mathbf{x})\right] \end{pmatrix}, \tag{4.13}$$

where s is the scale, $\mathbf{x}$ is the position vector, and $\mathbf{b}$ is its conjugate. We use the modulus of this transform $\mathcal{M}^{\boldsymbol{\phi}}[f]$ as a measure to find the singularities in a fashion similar to the aforementioned example. Fig. 4.6a show the components of the transform and its modulus as well as the argument for the simple example of Fig. 4.5a. Here, we used Gaussian wavelet which is a first-order wavelet ideal for detecting isotropic singularities with exponents smaller than 1.

After performing the transform, one can find the local maxima of modulus along the argument in order to determine the location of highest intensity changes, similar to the Canny's multi-scale edge detection [148]. This enables us to construct the maxima lines at each scale which is shown in Fig. 4.6b for the case of the isolated singularity of Fig. 4.5a. In this case, at each scale, there are two maxima lines: one is a closed loop around the singular point and the other one which is an open curve around the Gaussian maximum.

The next step is finding the local maxima along the maxima lines. In Fig. 4.6b, these are shown by red circles. Then one needs to chain these maxima points together throughout different scales to construct the maxima chains and the *skeleton* of the data which is the collection of all maxima chains. The blue lines in Fig. 4.6b show these chains for our simple example in which two of the lines point towards the singularity point while the other one is pointing towards the steepest side of Gaussian local maximum (the side opposite to the singularity). The last step of singularity detection via WTMMMM is determining whether a chain corresponds to a singularity or not. To do this, one can simply investigate the power-law exponent of the modulus along the maxima chains. The exponent of the power-law is dictated by the order of the wavelet if the chain corresponds to a point with more derivatives than the order of the wavelet. Obviously, this is the case when the point that the chain is pointing toward is a local maximum with infinite derivatives, but it is also the case when the order of the wavelet is less than the exponent of the singularity. Therefore, these two cases are not differentiatable via this method if the wavelet is not chosen properly. Nevertheless, if the point is a singularity with an exponent smaller than

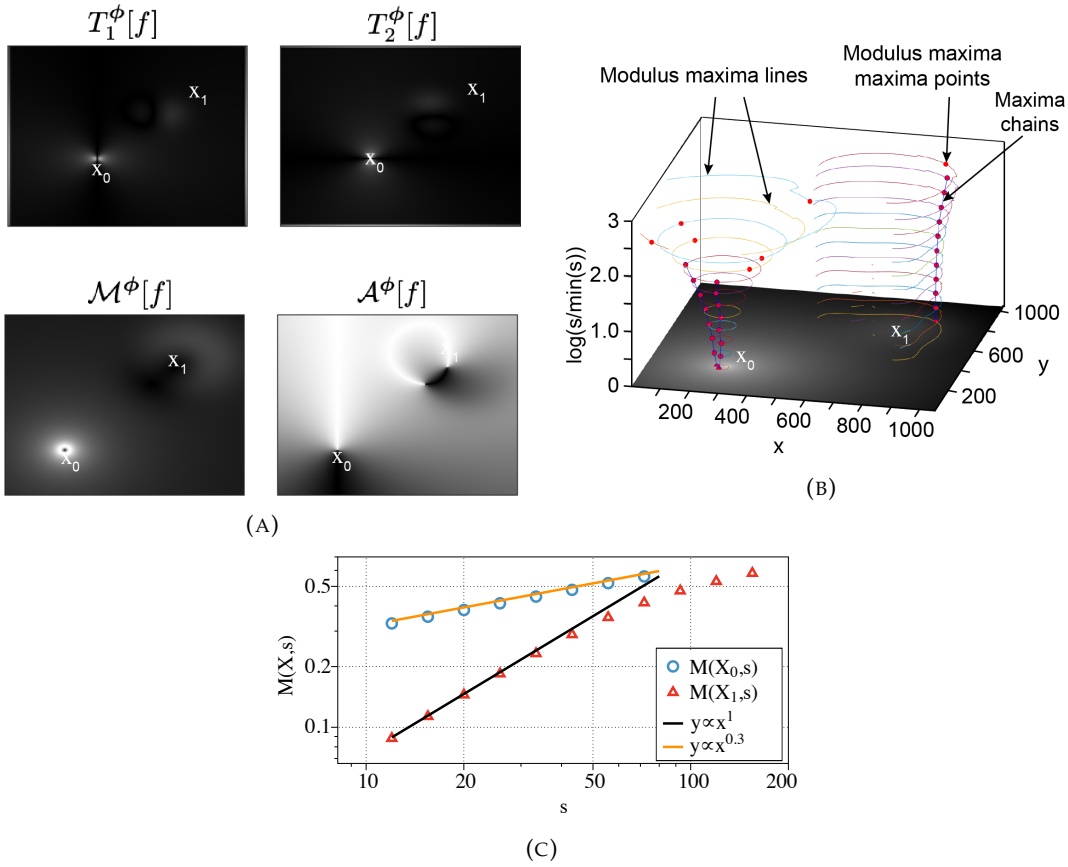
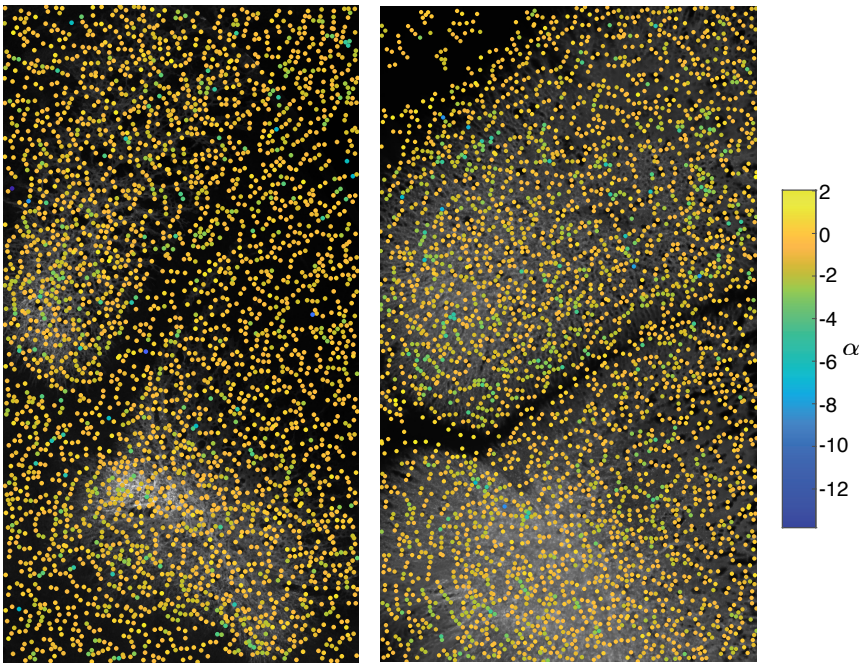


FIGURE 4.6: **Simple singularity tracking procedure via WTMMMM. (A)** The components ($T_1^{\phi}[f]$ and $T_2^{\phi}[f]$) and modulus $\mathcal{M}^{\phi}[f]$ and argument $\mathcal{A}^{\phi}[f]$ of the wavelet transform of the simple example in Fig. 4.5a using Gaussian wavelet with $s = 20$ **(B)** The Wavelet modulus maxima lines, modulus maxima maxima points (red circles) and maxima chains through different scales. **(C)** The power law behavior of the modulus $\mathcal{M}^{\phi}[f]$ along the maxima chains.The black and orange lines here show two power-laws with exponents equal to 1 (typical scaling behavior of the Gaussian wavelet) and 0.3 (corresponding to singularity exponent) respectively.



(A) (B)

FIGURE 4.7: **The result of WTMMMM on ossification data. (A)** and **(B)** The detected singularities and their exponents along with a typical data from E15.5 and E17.5, respectively.

the wavelet order, the exponent of the power-law is equal to that of the singularity. This is demonstrated by using our simple example in Fig. 4.6c where the the blue circles show the modulus along the chains that are pointing to the singularity point and follow a power-law with exponent 0.3. The red triangles however correspond to the chain of the Gaussian maxima that show a power-law with exponent 1 which is unsurprisingly equal to the order of the Gaussian wavelet. It should be noted that as a noisy data gets smoothed by wavelets with larger scale s, more information about the smaller scale fluctuations is lost. Therefore, at smaller scales there exist more chains that vanish as one goes to the larger scales.

So far we have discussed the singularity detection aspect of WTMMMM. However, this method with a few extra steps can be used to characterize multifractal features of images as discussed in App. C.4. Since this method is less accurate than MFDFA [141], we only use MFDFA for determining multifractal features and WTMMM for detecting singularities. In Fig. 4.7a and Fig. 4.7b, one can see a typical ossification pattern from E15.5 and E17.5 along with their singularities detected by a Morlet wavelet (a third-order wavelet). As one can see here, there are singularities with exponents ranging from -13 and 2. By comparing these two figures. One can convince oneself that as the ossified tissue develops, more singularities with negative exponents appear. This method is more insightful when combined with our generative model presented and discussed in the next section as it can unveil the relation between singularities and components of ossification process.

4.5 A toy model for generating ossification patterns

We investigated the origin of the observed multifractality in the data using surrogates in sec. 4.2.1, but implications of these features on the biological mechanisms involved in ossification remain yet to be uncovered. In this section, we therefore construct a simple model to generate patterns similar to those produced by ossification and analyze the results with MFDFA to find the necessary components to produce multifractal structures. We then apply WTMMMM to unveil the relation between features in the model and singularities.

We here propose a random coarse-grained model to produce multifractal patterns observed in real data. Note that our objective here is not to reproduce the entire ossification patterns with their complicated boundary since it is a result interactions with the surrounding tissues and vasculature. We solely focus on recapitulating a small section of the ossified tissue such as the magnified pieces in Fig. 4.1d which still shows multifractality when analyzed separately. Therefore, in our model, the tissue is composed of 900×1500 pixels each of which correspond to a region with area of approximately $1\mu m^2$. In each iteration, the intensity of a given pixel at (i, j) can increase by one with the probability

$$P(i,j) = P_{grad}(i,j) \times \big(P_{neigh}(i,j) + P_{coll}(i,j)\big) \tag{4.14}$$

which incorporates three basic mechanisms known to be involved in mineral deposition [149, 150]: Phosphate metabolism which acts as an effective morphogen gradient, a positive feedback from neighboring pixels, and a field of collagen fibres that act as ossification nucleators. To be more specific, P_{grad} corresponds to the effect of the phosphate metabolism, which is assumed to effectively act as a morphogen gradient. This term is proportional to the concentration of phosphate in the form of a one-dimensional exponential gradient $\exp(-i/a)$. It is multiplied into the other terms so that there is no ossification where this function vanishes. P_{neigh} takes into account the fact that once a pixel's intensity is increased, its neighboring pixels are more likely to ossify and increase their intensity in the next iterations. In order to implement this, we use the convolution of the ossification intensity $X(i, j)$ with a kernel with a given width β (i.e. $\sum_v \sum_u K(i-u, j-v, \beta)X(i, j)$). Here, we use box kernel for faster computations. Finally, P_{coll} is the contribution of collagen fibers such that presence of a fiber in a neighborhood increases probability of ossification. For this term, we again use convolution of a collagen field with a box kernel with width γ. Substituting these into Eq. 4.14, one gets the explicit form of the probability as

$$P(i,j) \propto \exp\left(-\frac{i}{a}\right)\left(\sum_v \sum_u K(i-u, j-v, \beta)X(i,j) + b\sum_v \sum_u K'(i-u, j-v, \gamma)X(i,j)\right). \tag{4.15}$$

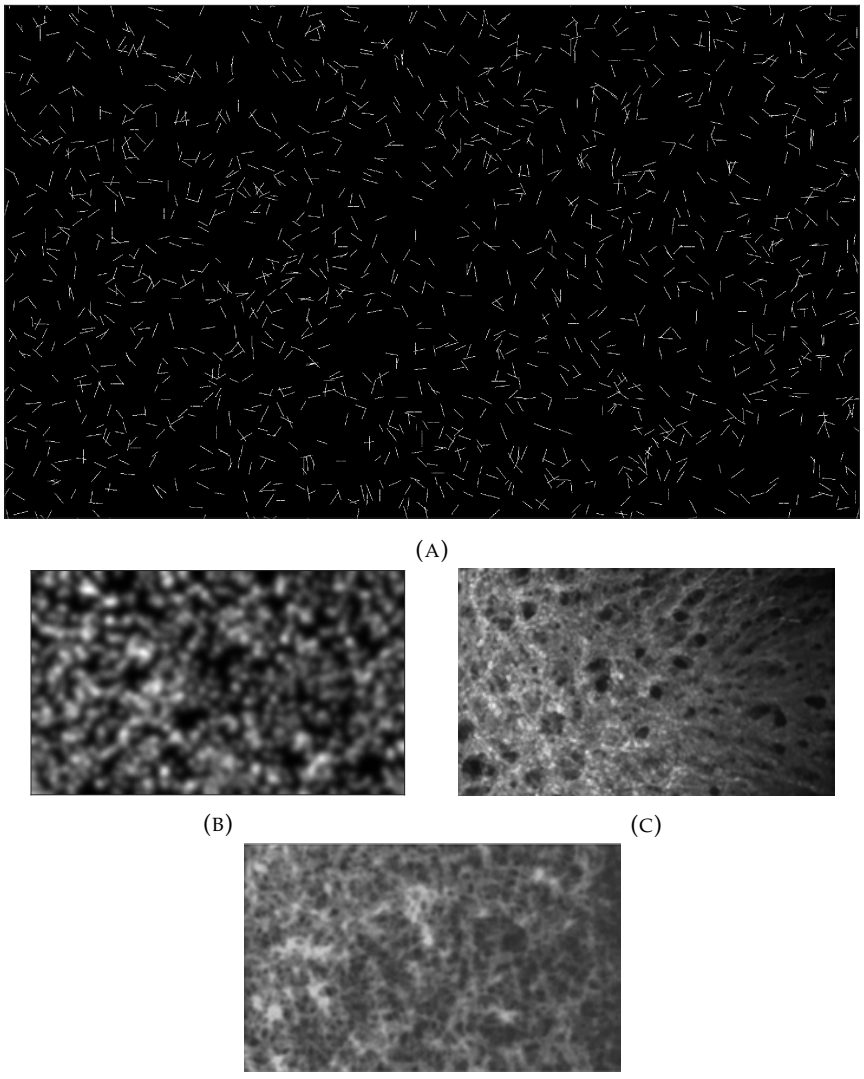
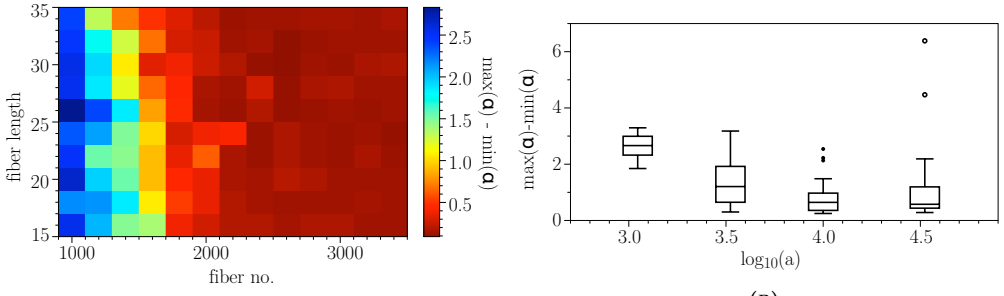


FIGURE 4.8: **Typical example of our ossification simulations. (A)** the collagen fiber template used containing 1400 fibers with length 20 distributed with uniform distribution across the field. **(B)** The result of our generative model for the fiber template shown in (A). **(C)** A small segment of ossification pattern from E15.5 with the same dimensions as (B). **(D)** Simulation results similar to (B) but with 100 pixels long fibers.

Fig. 4.8a shows a typical fiber field which contains 1400 fibers with length of 20 pixels randomly distributed in space. Running our simulation on this fiber template results in an image similar the one depicted in Fig. 4.8b if the following biologically relevant parameters are used: $a = 30000$, $\beta = 20$, $\gamma = 12$, and $b = 7.5$. Comparing this to a typical ossification pattern from E15.5 shown in Fig. 4.8c, one can argue that longer fibers are needed. In Fig. 4.8d, on can see a simulation result with very long (100 pixel long) fibers. Although, this image is more similar to the experimental data it does not exhibit any multifractality (i.e. the generalized Hurst exponent is constant). In principle, it might be possible to recapitulate both multifractality and appearance of the data by incorporating long-range interactions of fibers resulting in long persistence length. However, this type of interactions and alignment of fibers adds extra complexity which defies the purpose of our toy model. We therefore ignore the visual differences and focus on the multifractal features of our toy model.

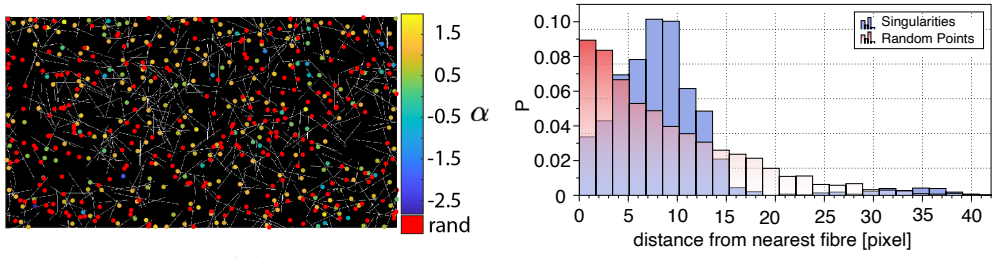
We first investigate the effect of the collagen fiber density (length and number) on the overall multifractailty as measured by $\max(\alpha) - \min(\alpha)$ which is the width of singularity spectrum. Fig. 4.9a summarizes the result of our simulations in which number of collage fibers vary from 1000 to 3400 and their lengths range from 16 to 34. Each value of the heatmap is the average over 30 realizations. As one can see here, increasing the number of fibers independent of their length results in suppression of multifractality. This is due to the fact that with very high number of fibers, the image fills with fibers and no voids (areas with low ossification) form. We also change the length scale of the phosphate gradient a to study its effect on the multifractality. Fig. 4.9b shows the width of singularity spectrum for four different values of a. As a increases, the phosphate gradient becomes shallower resulting in a more homogeneous ossification pattern with smaller multifractality. We also investigated if any effect arises from preferential orientation of collagen fibers. We assumed that fibers either acquire a predefined orientation (a multiple of $\pi/4$) with a given probability (decreasing exponentially from left to right) or a random orientation with uniform distribution over $[0, \pi)$. By varying the scale of the exponential over more than an order of magnitude, we did not observe any significant change in the multifractality.

Having studied how different components of the generative model affect multifractality, one can use WTMMMM to study the relation of singularities (whose exponents and their spectrum are used for characterizing multifractality) with fibers (whose density and number is the most important parameter affecting multifractailty). Fig. 4.8a depicts a typical fiber field used in a simulation along with the singularities detected by WTMMMM whose exponents are shown by the color of the circles. It appears that singularities are mostly located next to fibers. This can be tested by comparing the singularities to the same number of random points drawn from uniform spatial distribution shown by red circles. Fig. 4.10b shows the histograms of distance of points to their nearest fiber. As one can see here, the histogram for singularities has a peak around 9 pixels while for the random points ,



(A) (B)

FIGURE 4.9: **The role of components of our generative model on the overall multifractality** $\max(\alpha) - \min(\alpha)$**. (A)** Effect of fiber density (length and number). **(B)** The effect of phosphate gradient scale a. Each value in (A) and (B) is the overage of 30 realizations.

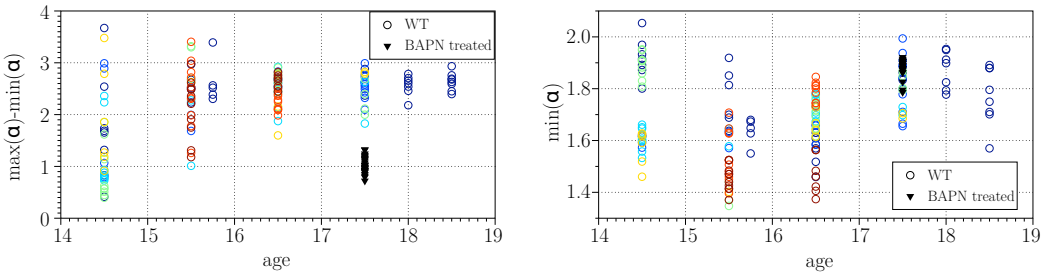


(A) (B)

FIGURE 4.10: **WTMMMM on simulation results. (A)** shows the collagen fibers along with the resultant singularities (in color coded circles) and random points with uniform distribution (in red circles). **(B)** The histogram of distances of the singularities (in blue) and random points (in red) to their the nearest fiber.

the histogram shows an exponential like decrease indicating a Poisson distribution. Therefore, one can conclude that the distance of singularities from fibers are indeed not random and they are produced as a result of fibers being nucleators of the ossification process. It should be noted that smallest scale used for WTMMMM here is 12 pixels which contributes to the separation between fibers and singularities.

Studying our simple generative model suggests that collagen fibers as nucleation centers are the most crucial components of ossification in terms of multifractality. In order to explore this implication experimentally, we studied ossification process of skull bone in Beta Aminoproprionitrile (BAPN) treated embryos. This drug is an irreversible inhibitor of the collagen crosslinker highly enriched in bone [151]. Therefore, this treatment in principle disrupts collagen fiber formation. Analyzing skull caps of BAPN treated embryos at E17.5, we found that the width of singularity spectrum $f(\alpha)$ indeed reduces significantly as shown in Fig. 4.11a. However, this perturbation does not affect $\min(\alpha)$ corresponding to the most irregular singularities of the data as shown in Fig. 4.11b.



(A) (B)

FIGURE 4.11: **Multifractal features of BAPN treated ossification patterns. (A)** The width of singularity spectrum of BAPN treated data at E17.5 (shown by black triangles) compared to the wild type subjects (shown by hollow circles). **(B)** Comparison of min(α) of BAPN treated ossification patterns (shown by black triangles) and wild type subjects (shown by hollow circles).

4.6 Conclusion and outlook

In this chapter, we demonstrated the utility of fluctuation statistics for studying biological mechanisms. We employed multi-scale analyses, in particular multifractal methods, to study ossification process in flat bone formation which involves mechanisms from various scales. We first showed that multifractal measures such as the width of singularity spectrum provides a robust measure for characterizing the patterns produced in ossification of skull caps. We then investigated the origin of multifractality in the data and showed that the observed multifractality is mostly emerging from long-range non-linear correlations. We furthermore developed a simple generative model capable of producing multifractal patterns based on mechanisms involved in ossification process. Using MFDFA and WTMMMM, we studied the relation between singularities and the components of ossification showing that collagen fibers as nucleation centers of ossification cause singular points and the density of fibers in our model is a crucial parameter for producing multifractality. This was tested by perturbations of ossification process in the embryo which resulted in suppressed multifractality.

While we have only tested the practicality of multifractal analyses in one specific tissue, we believe that such techniques are extremely valuable for addressing a more broad range of biological questions. For example, mesenchymal tissues fill space in an irregular manner. Therefore, applying tools developed for studying regular epithelial sheets on mesenchyme is not as practical. More advanced tools such as multifractal approaches are required to better study their formation and effects of different perturbation on mesenchymal tissues.

It should also be noted that all the characterization approaches used here are based on the spatial fluctuations of the data. This emphasizes the fact that fluctuations are not merely random, but they carry crucial information about the system. Therefore, filtering out these fluctuations causes loss of information.

Chapter 5

Conclusion

Living systems are subject to various types of noise originating from different sources. Even their tools for coping with these noises are inherently stochastic. Nevertheless, they manage to overcome all these uncertainties and flourish in a wide range of conditions. Therefore, studying how this robustness is achieved is an imperative aspect of biology. Also, this knowledge can be potentially helpful for diagnosis and treatment purposes. Additionally, it can guide engineers when designing delicate systems functioning under noisy conditions with inherent stochasticity.

Living organisms in order to survive, need to collect information about their surroundings and react to changes. Much effort has been devoted to understanding limits and mechanisms of sensing external signal via cells individually or collectively. However, a rigorous framework incorporating temporal and spatial fluctuations including intrinsic noise of reactions, cell-to-cell variability and neighborhood fluctuations was unfortunately missing. Therefore, we constructed a stochastic model of collective signal sensing based on continuous-time Markov chain formalism that enables us to study the quality of estimation in populations with interactions beyond simple mean-field.

Our results presented in chapter 2 show that in identical populations (without cell-to-cell variability), communication between cells, independent of interaction topology, solely reduces the time required for reaching steady state MSE while leaving the final values intact. This is relevant in scenarios where a consensus decision is needed in a limited time, but in the case of inhomogeneous cells, communication also enhances steady-state estimation quality. Our framework is applicable to any arbitrary interaction topology which we utilized to reveal the effect of interaction at the individual cell and population level. At the individual cell level, the MSE of a given cell reduces as the number of neighbors or strength of coupling increases, but to our surprise, it does not depend on the communication between the neighbors. As an element in large network, cell's quality of estimation depends on its closeness centrality. At the population level, we showed that the average MSE of estimations decreases as average shortest paths between cells decreases.

Generalizations of our model are possible by minor modifications in order to make it more realistic. For example, the interaction between cells is assumed to be instantaneous while in reality, the transport of molecules takes time. Therefore, incorporating a time delay between disappearing a sensor molecule from a cell and its appearance in the receiving one is needed. Also, in our framework all cells are exposed to the same exact environmental signal while graded signal or its spatial fluctuation defy this assumption. Finally, in many real-world cases, only a fraction of cells are exposed to the given external signals and the rest of the population's estimation is solely through the communication.

Applicability of our framework goes beyond just studying collective estimation of biochemical cues. Any physical property of the environment that can be represented by an internal concentration of chemical species can be studied by our framework. For example, cells in an epithelial sheet attached to a curved surface can sense the curvature via activity of macro-molecules with intrinsic bent such as BAR molecules.

It is easy to convince oneself that our results on collective sensing have deeper implications even outside the realm of small and non-intelligent cells. For instance, noisy communication between identical autonomous agents is not expected to enhance their perception of reality. Similarly, in a scientific community if all members have the exact same school of thought, nonideal communications may not reduce the errors and just speed up the process of reaching a consensus. On the other hand, in order to gain a better grasp of reality in a diverse community, people with more central positions have a better advantage. Finally, in order to reduce the errors in our scientific community, establishing links and reducing the average distance is indeed profitable.

Having collected the necessary information about their environment, cells need to process it along with their internal states via their regulatory machinery to perform crucial decisions. Given the enormous number of decisions and functions needed during cells' life cycle, multifunctional motifs are common in their regulatory network. Robust decisions are especially challenging here because of various types of noise that the regulatory network is subject to.

In chapter 3, we introduced a few examples of dynamically switchable logic gates that are capable of performing multiple functions depending on their *context*. They serve as a toy model to study the trade-off between multifunctionality and robustness against three types of uncertainty: cell-to-cell variability, uncertainty in initial conditions and intrinsic noise. We defined robustness against each of these noise types and showed how one can calculate them for a given system. We furthermore demonstrated the applicability of our setup by designing a simple circuit capable of performing binary addition and subtraction.

Our proposed multifunctional circuits are crucial for designing high-order computing devices based on (bio-)chemical reactions since in such systems, directing the

signal to different sub-units is particularly difficult if possible. Similar ideas exist and is utilized in the realm of silicon based information processing. Here, in so called *field-programmable arrays*, connections in the processing layer is manipulated by the control layer allowing for design flexibility and device reusability. In our circuits, however, the multifunctionality *emerges* from interactions between components and enables them to perform more number of functions without requiring as many components as the traditional static counterpart. Additionally, in our design there's no need for extra components dedicated to control the connections as switching between functions is possible via initial conditions.

There are several follow-up questions worth mentioning here. For instance, how can high-order and complex calculations be performed by combining the ANDOR gate and its cousins? Is it possible to "train" such networks to perform a given computation similar to artificial neural networks but by adjusting the initial conditions instead of adjusting strength of the connection? Furthermore, is there a way of controlling the initial conditions and consequently the performed functions via dynamic inputs? For example, if the connections between the input layer and the intermediate one is asymmetric, is it possible to steer the system towards the desired basin of attraction by a time-dependent input signal? Additionally, it is imperative to explore actual biological systems and regulatory networks to see when robustness is preferred over multifunctionality and how their trade-off compares to the idealized theoretical cases.

Finally, after processing the necessary information, a community of cells may take action in a collective manner and possibly by combining multiple processes across different scales. This type of activities results in complicated and non-trivial patterns requiring cross-scale analyses. A prime example of such processes is development of flat bones like skull caps. In chapter 4, we employed two-dimensional multifractal analyses to characterize and study spatial fluctuations in ossification patterns of developing skull caps of mouse embryos. We demonstrated the utility of these analyses by robustly characterizing the experimental data using multifractal measures such as width of the singularity spectra. We studied surrogates of the data to show that the observed multifractality originates from long-range non-linear correlations in the data. We also introduced a simple generative model capable of producing multifractal patterns with the right set of parameters. We furthermore introduced a generative model to reproduce multifractality of ossification pattern based on simple mechanisms. We then analyzed the simulation results via MFDFA and showed that the density of the collagen fiber patterns in the tissue are crucial to organize the ossification dynamics in order to produce the observed multifractality. We found further support for these predictions by analyzing the patterns in which the proper production and organization of collagen in the developing mouse embryo was perturbed. We showed that in this case multifractality is significantly reduced. Using

WTMMMM, an advanced singularity tracking method, we revealed that singularities co-localize with the collagen fibers.

There are many follow-up directions to pursue from where we ended this project. One crucial extension to our generative model is including interactions between ossification nucleators (i.e. collagen fibers). The alignment of these fibers can improve the appearance of simulation results without suppressing multifractality. Another important extension is to consider the thickness of the bone. Ossification and presence of fibers at different depths could result in a richer dynamics with potentially more control on producing multifractal features. Furthermore, the relation between ossification and underlying collagen fiber can be investigated via analyzing the electron microscopy images.

Note that multifractality is a property of fluctuations in the data and our study here shows how insightful they are for studying biological pattern formation. This information will be lost if crude filtering or some smoothing process is done to the raw data. Hopefully, this would increase appreciation of fluctuations in biology and related fields.

To sum up, this thesis was a step towards understanding the limitations arising from stochasticity, and appreciating the information that can be acquired from investigating fluctuations in the context of biology. We here studied robustness of biological processes despite various types of stochasticity present at all scales and stages from sensing to functioning. Biological systems achieve robustness by following a general scheme composed of three steps: sensing the environment, processing information and taking the action. More specifically, we considered three problems of significant importance in biology: collective sensing of dynamic environment, flexible information processing via internal machinery of cells, and stochastic pattern formation during tissue growth. We combined techniques from various fields such as stochastic processes, dynamical systems, and network science to tackle the problem of *robustness in biology: from sensing to functioning*.

Appendix A

More on collective sensing of environmental signals

A.1 Sensing dynamics in interacting cell communities

As mentioned in the main text, the dynamics of the environmental signal $Z(t)$ and its estimators $M^{(i)}(t)$ can be described by a set of stochastic differential equations which have independent unit Poisson processes counting the occurrences of the reactions, i.e.

$$Z(t) = Z(0) + \underbrace{R_b^Z(\rho t)}_{\text{birth reaction}} - \underbrace{R_d^Z\left(\varphi \int_0^t Z(s)\mathrm{d}s\right)}_{\text{death reaction}} \tag{A.1}$$

$$\begin{aligned} M^{(i)}(t) = M^{(i)}(0) &+ \underbrace{R_s^{(i)}\left(\gamma c_M \int_0^t Z(s)\mathrm{d}s\right)}_{\text{sensor reaction in cell } i} \\ &+ \underbrace{R_b^{(i)}\left(\gamma(\rho+\Delta\rho^{(i)})t\right) - R_d^{(i)}\left((\varphi+c_M)\int_0^t M^{(i)}(s)\mathrm{d}s\right)}_{\text{estimator reactions in cell } i} \\ &+ \sum_{j=1}^{N}\left[\underbrace{R_t^{i\leftarrow j}\left(\alpha_{ij}\int_0^t M^{(j)}(s)\mathrm{d}s\right)}_{\text{transport to cell } i} - \underbrace{R_t^{j\leftarrow i}\left(\alpha_{ij}\int_0^t M^{(i)}(s)\mathrm{d}s\right)}_{\text{transport from cell } i}\right]. \end{aligned} \tag{A.2}$$

Here, R_b^Z, R_d^Z, $R_b^{(i)}$, $R_d^{(i)}$, $R_s^{(i)}$ and $R_t^{j\leftarrow i}$ are the independent unit Poisson processes counting the occurrences of the respective reaction. A single isolated estimator will therefore follow a stochastic differential equation:

$$\mathrm{d}M^{(i)}(t) = \mathrm{d}R_b^{(i)}(t) - \mathrm{d}R_d^{(i)}(t) + \mathrm{d}R_s^{(i)}(t), \tag{A.3}$$

with $R_b^{(i)}(t)$, $R_d^{(i)}(t)$ and $R_s^{(i)}(t)$ as the reaction counters of the birth, death and sensing reactions, respectively. Note that these can each be decomposed into a predictable part and a zero-mean process such that $R_b^{(i)}(t) = \gamma\rho^{(i)}t + \tilde{R}_b^{(i)}(t)$ and $R_d^{(i)}(t) = (c_M+\varphi)\int_0^t M^{(i)}(s)\mathrm{d}s + \tilde{R}_d^{(i)}$. However, in the presence of cell-to-cell communication,

one needs to take into account the molecule exchange. Therefore, we define the net flux of the transport reactions as

$$\begin{aligned} \mathrm{d}R_t^{ij}(t) =& \mathrm{d}R^{i\leftarrow j}(t) - \mathrm{d}R^{j\leftarrow i}(t) \\ =& \big(\alpha_{ij}M^{(j)}(t) - \alpha_{ji}M^{(i)}(t)\big)\mathrm{d}t + \mathrm{d}\tilde{R}_t^{(ij)}(t), \end{aligned} \tag{A.4}$$

with $\mathrm{d}\tilde{R}_t^{(ij)}(t) = \mathrm{d}\tilde{R}^{i\leftarrow j}(t) + \mathrm{d}\tilde{R}^{j\leftarrow i}(t)$. By adding this transport reaction to the birth-death reactions of estimator $M^{(i)}$ Eq. A.3, it's dynamics would be

$$\mathrm{d}M^{(i)}(t) = \mathrm{d}R_b^{(i)}(t) - \mathrm{d}R_d^{(i)}(t) + \sum_j \mathrm{d}R_t^{(ij)}(t) + \mathrm{d}R_s^{(i)}(t) \tag{A.5}$$

where $i, j = 1, 2, ..., N$ are indices and N is the size of the system (i.e. the number of estimators). Note that in practice, these interactions are symmetric and $\alpha_{ij} = \alpha_{ji}$.

Now, by substituting values of $R_t^{(ij)}$, $R_b^{(i)}$ and $R_d^{(i)}$ into Eq. A.5, the evolution of the estimator in cell i becomes

$$\begin{aligned} \mathrm{d}M^{(i)}(t) =& \big[\gamma(\rho + \Delta\rho^{(i)}) - (\varphi + c_M)M^{(i)}(t) + \gamma c_M Z(t) \\ &+ \sum_j \big(\alpha_{ij}M^{(j)}(t) - \alpha_{ji}M^{(i)}(t)\big)\big]\mathrm{d}t \\ &+ \mathrm{d}\tilde{R}_b^{(i)}(t) - \mathrm{d}\tilde{R}_d^{(i)}(t) + \mathrm{d}\tilde{R}_s^{(i)}(t) + \sum_j \mathrm{d}\tilde{R}_t^{(ij)}(t), \end{aligned} \tag{A.6}$$

which is the stochastic differential equation describing the dynamics of estimator i. The third term in this equation stems from the interactions with the neighboring cells and it is commonly known as a *consensus* term [152]. Taking the expectation of Eq. A.6 gives:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[M^{(i)}(t)] =& \gamma\rho - (\varphi + c_M)\mathbb{E}[M^{(i)}(t)] + \gamma c_M \mathbb{E}[Z(t)] \\ &+ \sum_j \big(\alpha_{ij}\mathbb{E}[M^{(j)}(t)] - \alpha_{ji}\mathbb{E}[M^{(i)}(t)]\big). \end{aligned} \tag{A.7}$$

Note that the time-evolution of estimators are coupled to each other and also to the expectation of the environmental signal $\mathbb{E}[Z(t)]$, whose time evolution similarly satisfies:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[Z(t)] = \rho - \varphi\mathbb{E}[Z(t)]. \tag{A.8}$$

A.2 Sensing accuracy in interacting cell communities

To assess the bias and accuracy of each cells estimator, we define the Mean Error (ME) $\mathbb{E}\left[e_i(t)\right] = \mathbb{E}\left[Z(t) - {}^{M^{(i)}(t)}/_{\gamma}\right]$ and the Mean Squared Error (MSE) $\mathbb{E}\left[e_i^2(t)\right] = \mathbb{E}\left[\left(Z(t) - {}^{M^{(i)}(t)}/_{\gamma}\right)^2\right]$, respectively. It is also easy to show (by subtracting Eq. A.8

from A.7) that the expected sensing error of cell i (i.e. $\mathbb{E}[e_i(t)] = \mathbb{E}[Z(t) - \frac{1}{\gamma}M^{(i)}(t)]$) is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[e_i(t)] = -(\varphi + c_M)\mathbb{E}[e_i(t)] + \sum_j \big(\alpha_{ij}\mathbb{E}[e_j(t)] - \alpha_{ji}\mathbb{E}[e_i(t)]\big). \tag{A.9}$$

Since the stationary state of the mean error is always zero, the Mean Squared Error (MSE) represents a better measure for quality of the estimation. In order to calculate this, we employ the Itô formula for counting processes, which can be formulated for our special case as following: assume a counting process $X(t)$ has the form $\mathrm{d}X(t) = a\mathrm{d}N(t)$, where $\mathrm{d}N(t)$ is a counting process. Any function of this process $F(X(t))$ will then evolve in time as

$$\mathrm{d}F(X(t)) = [F(X(t) + a) - F(X(t))]\mathrm{d}N(t). \tag{A.10}$$

For convenience, we separate the transport reaction term in Eq. A.4 as $\mathrm{d}R_t^{(i)} = \mathrm{d}R_{in}^{(i)} - \mathrm{d}R_{out}^{(i)}$ in which $\mathrm{d}R_{in}^{(i)} = \sum_j \mathrm{d}R^{i\leftarrow j}(t)$ and $\mathrm{d}R_{out}^{(i)} = \sum_j \mathrm{d}R^{j\leftarrow i}(t)$. By definition we have

$$\begin{aligned}
\mathrm{d}e_i(t) =& \mathrm{d}Z(t) - \frac{1}{\gamma}\mathrm{d}M^{(i)}(t) \\
=& \mathrm{d}Z_b(t) - \mathrm{d}Z_d(t) \\
&- \frac{1}{\gamma}\big(\mathrm{d}R_b^{(i)}(t) - \mathrm{d}R_d^{(i)}(t) + \mathrm{d}R_{in}^{(i)}(t) - \mathrm{d}R_{out}^{(i)}(t) + \mathrm{d}R_s^{(i)}(t)\big)
\end{aligned} \tag{A.11}$$

using Eq. A.10 (the Itô formula) yields

$$\begin{aligned}
\mathrm{d}e_i^2(t) =& \big(1 + 2e_i(t)\big)\mathrm{d}Z_b(t) + \big(1 - 2e_i(t)\big)\mathrm{d}Z_d(t) \\
&+ \big(\frac{1}{\gamma^2} + \frac{2}{\gamma}e_i(t)\big)\big[\mathrm{d}R_d^{(i)}(t) + \mathrm{d}R_{out}^{(i)}(t)\big] \\
&+ \big(\frac{1}{\gamma^2} - \frac{2}{\gamma}e_i(t)\big)\big[\mathrm{d}R_b^{(i)}(t) + \mathrm{d}R_{in}^{(i)}(t) + \mathrm{d}R_s^{(i)}\big]
\end{aligned} \tag{A.12}$$

Substituting the values of each term and taking the average over realizations, gives the following differential equation for the MSE of the cell i:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[e_i^2(t)] =& \rho(1 + \frac{1}{\gamma}) - 2(\varphi + c_M + s_i)\mathbb{E}[e_i^2(t)] \\
&- \frac{1}{\gamma}(\varphi + c_M + s_i)\mathbb{E}[e_i(t)] - 2\mathbb{E}[\Delta\rho_i e_i(t)] \\
&+ \sum_j \alpha_{ij}\big(2\mathbb{E}[e_i(t)e_j(t)] - \mathbb{E}[e_j(t)]\big) \\
&+ \frac{1}{\gamma}\big(\varphi(1 + \gamma) + 2c_M + 2s_i\big)\mathbb{E}[Z(t)]
\end{aligned} \tag{A.13}$$

where $s_i = \sum_{j=1}^{N} \alpha_{ij}$ is degree of cell i and $\mathbb{E}[\Delta\rho_i e_i(t)]$ is the covariance of cell's uncertainty and its error of estimation. Since the evolution of MSE depends on the

covariance of the errors (i.e. $\mathbb{E}[e_i(t)e_j(t)]$), we also must find its time-evolution to obtain a closed set of equations. To this end, we start with the chain rule, i.e.,

$$\mathrm{d}\big(e_i(t)e_j(t)\big) = e_i(t)\mathrm{d}e_j(t) + \mathrm{d}e_i(t)e_j(t) + \mathrm{d}e_i(t)\mathrm{d}e_j(t). \tag{A.14}$$

The third term in Eq. A.14 is always zero unless $\mathrm{d}e_i$ and $\mathrm{d}e_j$ jump simultaneously. This happens only if: (1) an estimation molecule is exchanged between cell i and j, or (2) a birth or death reaction occurs for the Z species. Using this assumption, substitution of Eq. A.11 into Eq. A.14, and finally taking the expectation, results in

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[e_i(t)e_j(t)] =& \rho - (2\varphi + 2c_M + s_i + s_j)\mathbb{E}[e_i(t)e_j(t)] \\ &+ \sum_k \big(\alpha_{jk}\mathbb{E}[e_i(t)e_k(t)] + \alpha_{ik}\mathbb{E}[e_j(t)e_k(t)]\big) \\ &+ \alpha_{ij}\big(\mathbb{E}[e_j(t)] + \mathbb{E}[e_i(t)]\big) + \big(\varphi - \frac{2\alpha_{ij}}{\gamma}\big)\mathbb{E}[Z(t)] \\ &- \mathbb{E}[\Delta\rho_i e_j(t)] - \mathbb{E}[\Delta\rho_i e_j(t)]\end{aligned} \tag{A.15}$$

Similarly, one can easily find the time evolution of $\mathbb{E}[\Delta\rho_i e_i(t)]$ and $\mathbb{E}[\Delta\rho_i e_j(t)]$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\Delta\rho_i e_i(t)] = -(\varphi + c_M + s_i)\mathbb{E}[\Delta\rho_i e_i(t)] + \sum_j \alpha_{ij}\mathbb{E}[\Delta\rho_i e_j(t)] - \mathbb{E}[\Delta\rho_i^2] \tag{A.16}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\Delta\rho_i e_j(t)] = -(\varphi + c_M + s_j)\mathbb{E}[\Delta\rho_i e_j(t)] + \sum_k \alpha_{jk}\mathbb{E}[\Delta\rho_i e_k(t)] \tag{A.17}$$

Solving Eqs. A.13, A.15,A.16 and A.17 jointly with Eq. A.9 gives the time evolution of the MSE (and all other variables) for any given initial conditions and set of parameters.

A.2.1 Sensing accuracy in star-shaped network

In this section, we study another simple yet informative case in which only a hub is connected to the rest of nodes and there are no other connections. Such networks which are known as star-shaped networks, represent one type of mean-field configuration. In this case, the stationary state version of equations A.13 and A.15 will

be

$$0 = -(\varphi + c_M + n\alpha)\mathbb{E}[e_c^2] + n\alpha\mathbb{E}[e_c e_m] - \mathbb{E}[\Delta\rho_c e_c] + \frac{1}{\gamma}(\varphi(1+\gamma) + c_M + n\alpha)\frac{\rho}{\varphi} \tag{A.18}$$

$$0 = -(\varphi + c_M + \alpha)\mathbb{E}[e_m^2] + \alpha\mathbb{E}[e_c e_m] - \mathbb{E}[\Delta\rho_m e_m] + \frac{1}{\gamma}(\varphi(1+\gamma) + c_M + \alpha)\frac{\rho}{\varphi} \tag{A.19}$$

$$0 = 2\rho - (2\varphi + 2c_M + (n+1)\alpha)\mathbb{E}[e_c e_m] + \alpha\mathbb{E}[e_c^2] + (n-1)\alpha\mathbb{E}[e_m e_{m'}] + \alpha\mathbb{E}[e_m^2] - \frac{2\alpha}{\gamma}\frac{\rho}{\varphi} - \mathbb{E}[\Delta\rho_c e_m] - \mathbb{E}[\Delta\rho_m e_c] \tag{A.20}$$

$$0 = -(\varphi + c_M + \alpha)\mathbb{E}[e_m e_{m'}] + \alpha\mathbb{E}[e_m e_c] + \rho - \mathbb{E}[\Delta\rho_m e_{m'}] \tag{A.21}$$

where index c indicates the central cell (hub) while m and m' indicate two distinctive marginal cells (on the periphery) which are only connected to the central one. In the case of Poissonian estimators, $\Delta\rho_i$ will be zero for every i and the set of equations will be closed. With this assumption, solving these equations gives us the same MSE independent of coupling similar to the case of fully connected or any other topology. However, in a more realistic case, $\Delta\rho_i$ will not be zero due to the presence of super-Poissonian statistics. In this case one should also consider $[\Delta\rho_i e_j]$ from Eqs. A.17 and A.16 which obey the following equations for different indices

$$0 = -(\varphi + c_M + n\alpha)\mathbb{E}[\Delta\rho_c e_c] + n\alpha\mathbb{E}[\Delta\rho_c e_m] - \mathbb{E}[\Delta\rho^2] \tag{A.22}$$

$$0 = -(\varphi + c_M + \alpha)\mathbb{E}[\Delta\rho_c e_m] + \alpha\mathbb{E}[\Delta\rho_c e_m] \tag{A.23}$$

$$0 = -(\varphi + c_M + n\alpha)\mathbb{E}[\Delta\rho_m e_c] + (n-1)\alpha\mathbb{E}[\Delta\rho_m e_{m'}] + \alpha\mathbb{E}[\Delta\rho_m e_m] \tag{A.24}$$

$$0 = -(\varphi + c_M + \alpha)\mathbb{E}[\Delta\rho_m e_m] + \alpha\mathbb{E}[\Delta\rho_m e_c] - \mathbb{E}[\Delta\rho^2] \tag{A.25}$$

$$0 = -(\varphi + c_M + \alpha)\mathbb{E}[\Delta\rho_m e_{m'}] + \alpha\mathbb{E}[\Delta\rho_m e_c]. \tag{A.26}$$

Appendix B

More on context-dependent information processing

B.1 The ANDOR gate with Hill dynamics

As mentioned in Chap. 3, we a specific type of dynamics for our examples which is shown in Eq. 3.1. However, one can in principle use any other type of regulatory function (e.g. the Hill function) as long as it features a switch-like sigmodal behavior and dynamically switching should still be achievable. In order to demonstrate this, we also constructed another version of the ANDOR gate based on the Hill function regulatory dynamics instead of the one shown in Eqs. 3.3 and 3.4:

$$\dot{x}_1 = \frac{\left(x_1/0.3\right)^n + \left(s/0.8\right)^n}{1 + \left(x_1/0.3\right)^n + \left(s/0.8\right)^n + \left(x_2/0.5\right)^n} - x_1, \tag{B.1}$$

$$\dot{x}_2 = \frac{\left(s/0.35\right)^n}{1 + \left(s/0.35\right)^n + \left(x_1/0.5\right)^n} - x_2. \tag{B.2}$$

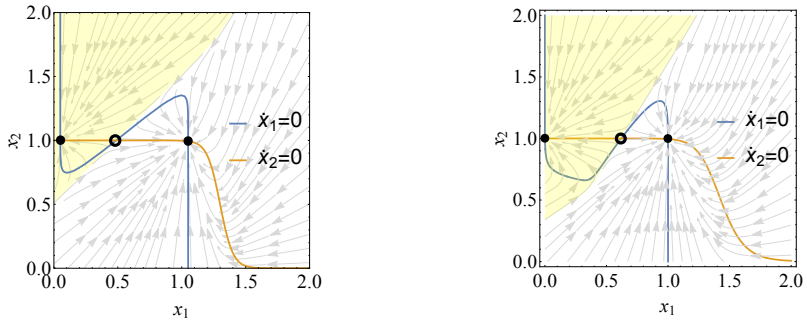
Assuming this dynamics for the intermediate layer of the ANDOR gate with $n = 15$ results in a phase portrait that is qualitativey similar to that of Eqs. 3.3 and 3.4. In Fig. B.1, one can see these two plots side by side for the intermediate input signal level $s = 1$.

B.2 Adaptive Minimum Action Path

As discussed in 3.5.3, one can study noise-induced transitions by the theory of large deviations constructed by Freidlin and Wentzell [17]. Recall that For a given stochastic system whose stochastic time evolution is given by a Langevin equation as following:

$$\dot{\mathbf{X}}(t) = b\left(\mathbf{X}(t)\right) + \varepsilon\sigma\left(\mathbf{X}(t)\right)\dot{\mathbf{W}}(t), \tag{B.3}$$

where $\mathbf{X}(t)$ is the n-dimensional state vector at time t, $b\left(\mathbf{X}\right)$ is the drift term and the diffusion. This diffusion term is composed of an m-dimensional Wiener process $\mathbf{W}$ and $\sigma\left(\mathbf{X}(t)\right)$ which is the standard deviation noise and ε which is a small parameters



(A) (B)

FIGURE B.1: **Different underlying regulatory dynamics display the same bistability.** The phase portraits of the bistable motif in Fig. 3.2 **(A)** with the phenomenological dynamics and in **(B)** with Hill function dynamics. In both of these plots, the solid black circles show the fixed points, the hollow ones show the saddle points and the shaded area show the basins of attraction of the left fixed points.

determining the noise strength. In this system, the probability of a path φ in the phase space to be taken by the system due to the intrinsic noise is proportional to $\exp\left(-S(\varphi)/\varepsilon^2\right)$, where the action $S(\varphi)$ is defined as:

$$S(\varphi) = \frac{1}{2}\int_{T1}^{T_2} \sum_{i,j} a_{ij}\left(\varphi_t\right)\left(\dot{\varphi}_t^i - b^i\left(\varphi_t\right)\right)\left(\dot{\varphi}_t^j - b^j\left(\varphi_t\right)\right) \mathrm{d}t. \tag{B.4}$$

Here, i and j count dimensions of the system and a_{ij} are elements of $A(x) \equiv \left(\sigma(\mathbf{X})\sigma^*(\mathbf{X})\right)^{-1}$.

When ε is small, all paths connecting a given pair of points have negligible probabilities compared to the one which minimizes the action in Eq. 3.6. This path is called the Minimum Action Path (MAP) which determines the path with the highest probability for a given transition at the given time interval.

In order to numerically find the MAP for a given system, one needs to discretize the system. Suppose that the path φ takes place over the time interval $[T_1, T_2]$. One can simply divide this interval into k sub-intervals as:

$$T_1 = t_0 < t_1 < \cdots < t_k = T_2. \tag{B.5}$$

Then, the path φ can be approximated by the set of its values Φ_n at these time points t_n for $n = 0, \ldots, k$. Also, the action $S(\varphi)$ can be approximated by

$$S(\Phi_0, \ldots, \Phi_k) = \frac{1}{2}\sum_{n=1}^{k} \Delta t_n \sum_{i,j} a_{ij}\left(\Phi_{n-1/2}\right)\left(\frac{2\Phi^i_{n-1/2}}{\Delta t_n} - b^i\left(\Phi_{n-1/2}\right)\right) \left(2\frac{\Phi^j_{n-1/2}}{\Delta t_n} - b^j\left(\Phi_{n-1/2}\right)\right), \tag{B.6}$$

in which $\Phi_{n-1/2} = (\Phi_n + \Phi_{n-1})/2$ and $\Delta t_n = t_n - t_{n-1}$. If the time steps Δt_n are

assumed to be constant and independent of the rate of change in systems position in the phase space $\dot{\varphi}$, then the approximation of the path will not have acceptable accuracy in most practical cases. This is due to the fact that when a path is connecting two metastable points, the majority of the time is spent at the end points. Therefore, most of the time points are dedicated to the semi-static part of the path. This problem can be circumvented via adaptive discretization of time. In other words, Δt_n should change depending on $\dot{\varphi}$ so that it can be resolved with a good accuracy. An adaptive Minimum action path was introduced in Ref. [153] to address this problem. In the following, we briefly discuss this method that we used in our study to find the MAP to the ANDOR gate.

The key concept of obtaining the adaptive MAP is using a monitor function according to which the re-meshing process of the time is performed. A good candidate would be the rate of change $\dot{\varphi}$. However, this value is zero in many cases resulting in singularities and a more practical choice would be

$$w(t) = \sqrt{1 + C\dot{\varphi}} \tag{B.7}$$

in which C is a large constant. Then, the quality of meshing can be measured by

$$Q = \frac{\max_n \left(w_n \Delta t_n\right)}{\min_n \left(w_n \Delta t_n\right)} \tag{B.8}$$

where w_n is the value of monitor function at time step t_n.

Having the monitor function and the quality of meshing, it is straightforward to find the adaptive MAP. One needs to start with a uniform meshing, find the MAP for the current meshing, and check the quality of meshing. If the quality is above a threshold, then calculate the monitor function and new meshing such that $\Delta t_n \propto 1/w_n$. Next, by interpolating, one can find the values of the calculated MAP at the newly meshed time points and use those as initial condition for the optimization procedure of the action and find a new MAP. Then, by repeating the procedure of remeshing and optimization until a good quality of meshing is achieved, one obtains the adaptive MAP.

In our study we set the constant C in Eq. B.7 to 10000, and the threshold of the Q is 3. The performance of re-meshing procedure is clear in Fig. 3.10b where discretized points cover the trajectory with little separations.

Appendix C

More on Fractal fluctuations

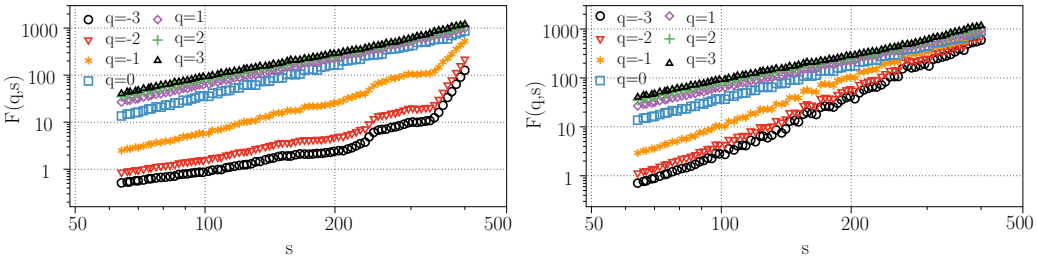
C.1 Choosing the parameter for MF-DFA

The parameters that should be used for MF-DFA vary from one system to another depending on their properties and therefore they should be determined separately for each dataset. In this study, we used the following procedure for a small subset of samples (from E15.5) to determine the parameters and then used these parameters to analyze all images.

In order to determine the range of scales s over which the analysis is done, one needs to start from the widest meaningful range. This range should start from the order of the fitted polynomial, so that the number of parameters does not exceed the number of points. Moreover, the range of s should stop at $s_{\max}$ that results in a number of segments being large enough so that the averaging process in eq. 4.4 remains meaningful. Once this maximum range is determined, one should find the sub-range in which the fluctuation function $F(q, s)$ shows power law scaling for all values of q. This range can be then used for the rest of the analysis.

In eq. 4.4, q is the order of the moment of the fluctuations. The range of this parameter should be chosen according to the minimum number of segments. There should be high number of segments for each value of q. Large values of q correspond to statistics of the segments with large fluctuations and negative values correspond to the segments with small fluctuations. Therefore, maximum and minimum q's should be chosen in a way that their corresponding fluctuation functions behave similar to others, i.e. show power law behavior over the chosen range of s.

The order of the chosen polynomial trend depends on the nature of the data and any intrinsic trends contained therein. As described in [145], we determined this order by starting from the simplest polynomial $\tilde{Y} = ai + bj + c$ and increasing the order in i and j until the surface $Y_{v,w}(i, j)$ is over-fitted and the generalized Hurst exponent is no longer a monotonic function of q. Here, it is especially loosing its accuracy in small q's corresponding to small fluctuation segments.



(A) (B)

FIGURE C.1: **The effect of removing periodic trends on measuring multifractaility. (a)** $F(q,s)$ vs. s for a typical and raw image at 15.5 stage. **(b)** $F(q,s)$ vs. s of the same image after removing 2 of the lowest frequencies in all directions.

C.2 Periodic trends and frequency analysis of the data

Although MFDFA is capable of determining multifractal features in the presence of many forms of trends in data, it fails when applied to the data sets with periodic trends. However, it is shown that such trends only cause an increase in the overall fluctuation function at scales near the wavelength of the trend which in turn results in inaccurate determination of generalized Hurst exponent [154]. One can circumvent this problem by filtering out two lowest frequency (i.e. long wavelength) components of the Fourier transform. This can be done by simply Fourier transforming the data, replacing low frequency terms by zero one by one, then taking the inverse Fourier transform, and analyzing the result until the crossover is removed.

In Fig. C.1a, the overall fluctuation function $F(q,s)$ is shown for a typical ossification at E15.5 with $q \in \{-3,2,1,0,1,2,3\}$. As one can see in this plot, there is a sudden increase in $F(q,s)$ at the large values of s. This crossover is more pronounced for smaller q's but exists in all of them which makes determination of $h(q)$ inaccurate. By removing a two lowest frequency components of the data, however, one gets proper power laws within the same range of s. In Fig. C.1b, the same plot as in C.1a is shown but after removing two lowest frequency components in all directions. It is worth noting that removing more frequency components results in flattening the generalized Hurst exponent i.e. suppressing the multifractality. In Fig. C.2, one can see the generalized Hurst exponent $h(q)$ of the data whose results are shown in Fig. C.1 with various number of low frequencies ($m \in \{0,2,5,8,11,14,17\}$) removed. As one can see here, the overall multi-fractality which can be defined as the change in $h(q)$ decreases as we increase m. Therefore, we only use $m = 2$ for analyzing all images in order to prevent over-filtering and getting the most accurate multi-fractal measures. A similar behavior has been observed in another context where low-pass filtering (removing the highest frequencies) also results in flattening of $h(q)$ when analyzing time-series recorded from V1 cortex [145].

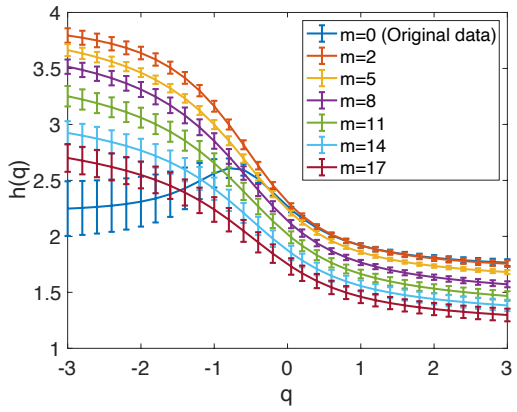


FIGURE C.2: **The effect of removing periodic trends on the multifractaility.** Generalized Hurst exponent of a typical ossification pattern after removing m lowest frequency.

C.3 Different multifractal measures of the data

In chapter 4, we used MFDFA to characterize the ossification patterns in the developing skull caps of mouse embryos. There are many different measures one can define based on the singularity spectrum shown in Fig. 4.4a. These measures are investigated for all images and the results are summarized in Fig. C.3. As one can see in this here, there are two main behaviors that the measures show over time: (I) starting from a wide distribution and getting narrower, and (II) having the width of distribution almost constant, but changing the average in a oscillatory way. The entire width of the spectrum $\max(\alpha) - \min(\alpha)$ in fig. C.3b and the maximum singularity exponent $\max(\alpha)$ in Fig. C.3d show the first behavior. They have wider range at E14.5 and their range gets narrower as the embryo develops. The average of these measures also increase over time. $\max(\alpha) - \min(\alpha)$ indicates the variety of singularities in the data and increasing its average over time means more diversity of singularities form over the course of development. Similarly, increase in average of $\max(\alpha)$ implies addition of more regular singularities.

The second type of behavior can be seen in the location of the maximum of the spectrum $\alpha_{\max}$ shown in Fig. C.3c, the minimum of the singularity exponents $\min(\alpha)$ in Fig. C.3a, left part of the range of the singularity exponents (i.e. $\alpha_{\max} - \min(\alpha)$) in Fig. C.3e. In order to interpret this type of behavior, it is useful to recall that $f(\alpha_{\max}) = 2$ meaning that singularities with exponent $\alpha_{\max}$ are distributed everywhere with Hausdorff dimension of 2. Therefore, may speculate that these correspond to non-informative and uniform noise in the data. In that case, the other oscillatory measures might also correspond to trivial and non-informative aspects of singularities.

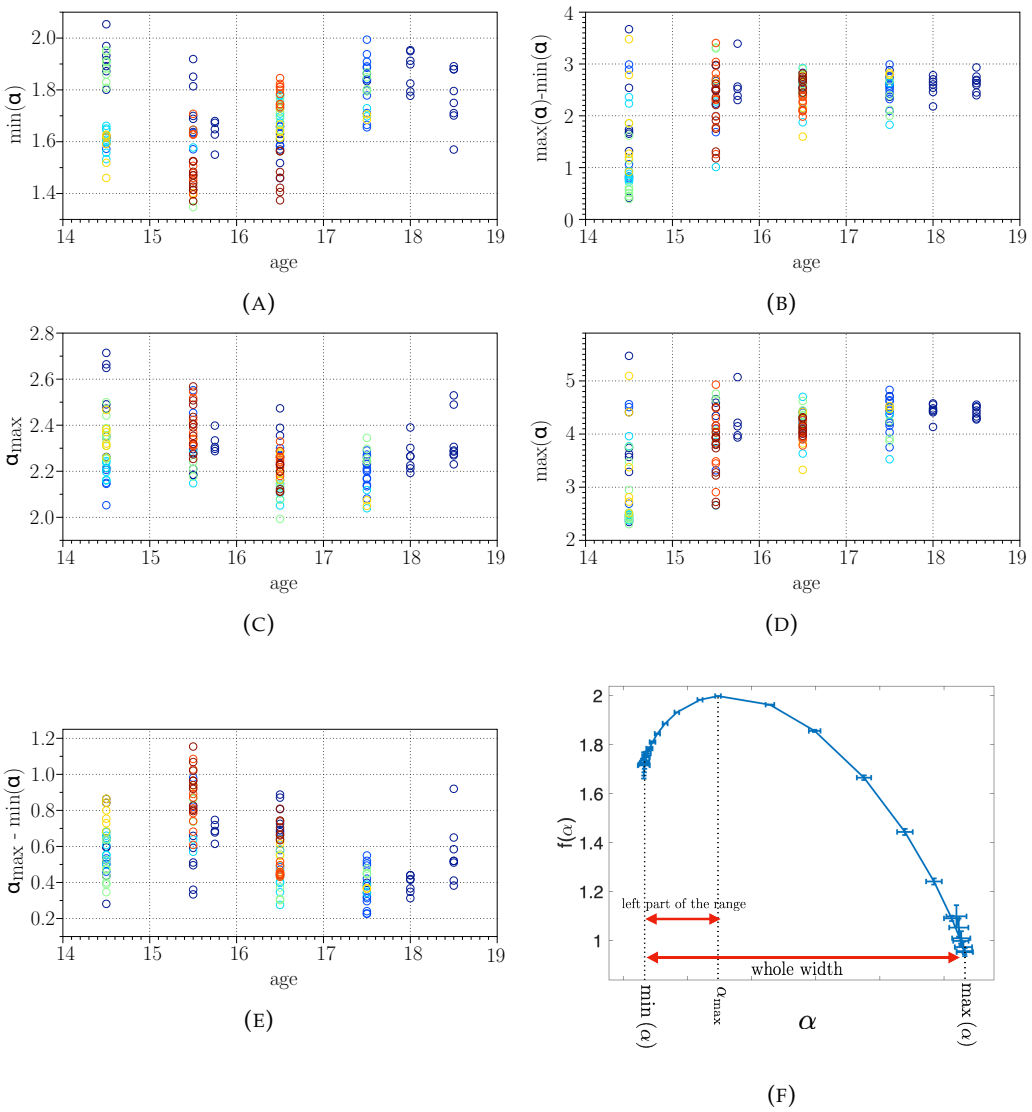


FIGURE C.3: **Multifractal measures over time.** **(A)** $\min(\alpha)$ **(B)** $\max(\alpha) - \min(\alpha)$, in **(C)** $\alpha_{\max}$ **(D)** $\max(\alpha)$ **(E)** $\alpha_{\max} - \min(\alpha)$ vs. embryonic day (age). **(F)** A typical singularity spectrum and the definition of different measures.

C.4 Characterizing multifractal features using WTMMMM

In Chap. 4, we used MFDFA for characterizing multifractal features of the data and WTMMMM for finding location of singularities. However, this method is also capable of determining all multifractal measures as we discuss in this section. Having constructed the skeleton of the data as shown in Sec. 4.4.1, one can calculate the partition function defined as

$$\mathcal{Z}(q,s) = \sum_{\mathcal{L}\in\mathcal{L}(s)} \left(\sup_{(\mathbf{x},s')\in\mathcal{L},s'\leq s} \mathcal{M}_{\psi}[f]\,(\mathbf{x},s') \right)^{q}, \tag{C.1}$$

where $\mathcal{L}(s)$ is the set of all maxima chains that exist at scale s and persist at scales $s' < s$. This partition function then can be used to determine the multi-fractal features of the data. One can define the exponent $\tau(q)$ as

$$\tau(q) := \frac{\log\left(\mathcal{Z}\left(q,s\right)\right)}{\log(s)}. \tag{C.2}$$

Here, $\tau(q)$ is linear if the data is mono-fractal, and nonlinear if it is multi-fractal. Alternatively, one can use a Legendre transform to obtain the singularity spectrum as following:

$$\alpha = \tau'(q), \tag{C.3}$$

$$f(\alpha) = \alpha q - \tau(q), \tag{C.4}$$

where results in a delta function for a mono-fractal signal and in a wide spectrum for a multi-fractal signal.

In order to show the next final steps of the WTMMMM on a data set with known fractal features, we apply it to a two-dimensional Brownian motion field with $H = 0.6$ and size 1024×1024. Note that this is a monofractal field with constant generalized exponent i.e. $h(q) = H = 0.6$. Therefore, according to Eq. 4.7, there's only one type of sibgularities in the data with exponent $\alpha = H = 0.6$ and the Hausdof dimension $f(\alpha = 0.6) = 2$. Accordingly, we can here again use the Gaussian wavelet since its order is larger than the largest singularity exponent in the field. Fig. C.4a shows the skeleton that are pointing toward the singularity points as $s \to 0^{+}$. Fig. C.4b depicts the partition function in logarithmic scale showing power law at small scales. Finally, Fig. C.4c and Fig. C.4d show the comparison between the results of MFDFA, WTMMMM and the theoretical values. The exponent can be written in terms of the generalized Hurst exponent $h(q)$ as $\tau(q) = qh(q) - D$ which for our case reads $\tau(q) = 0.6q - 2$ which is shown by black line in Fig. C.4c. As one can see here, for both measures, MFDFA gives more accurate results. This is again consistent with other studies such as Ref. [141]. However, MFDFA does not provide any information about the position of the singularities while in WTMMMM, one can simply

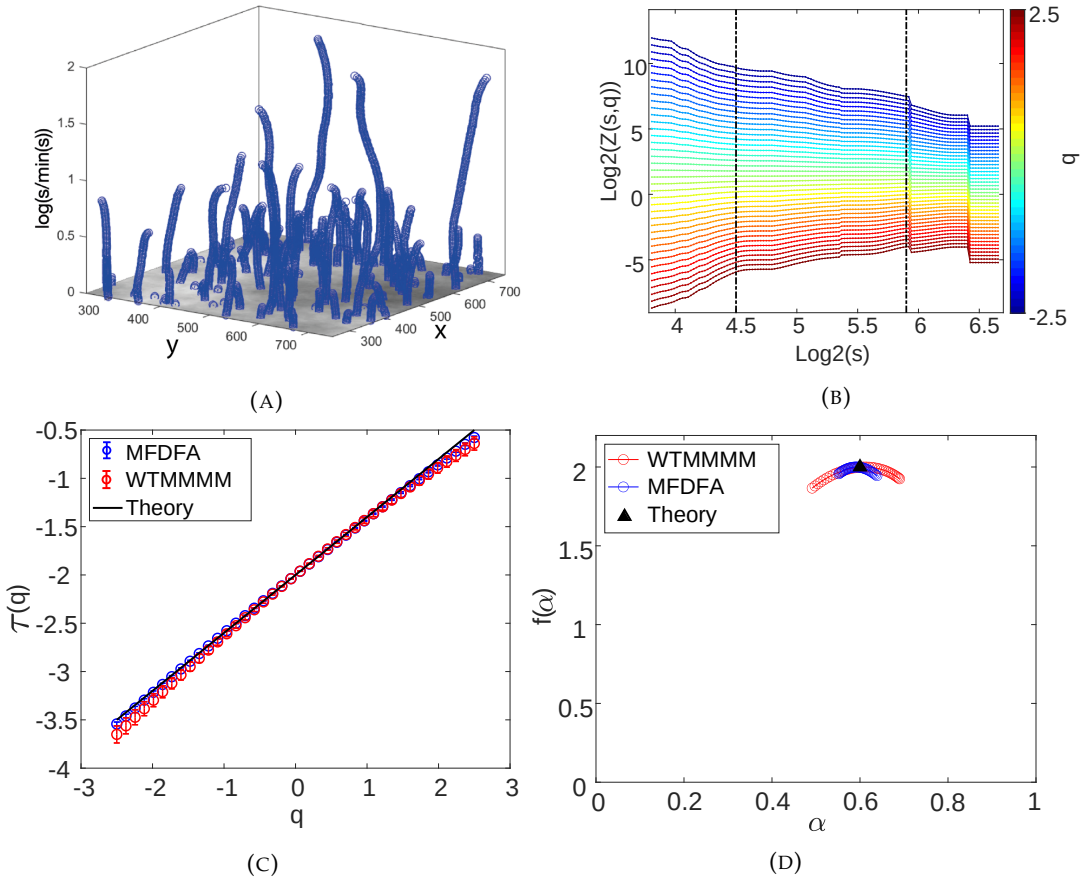


FIGURE C.4: **Application of WTMMMM on Fractional Brownian motion field with** $H = 0.6$ **(A)** A section of the skeleton of the data. **(B)** The partition function $\mathcal{Z}(q, s)$ in log-log scale. Here, the dashed black lines show the range in which the power laws are fitted to the curves. **(C)** Comparison of the exponent $\tau(q)$ obtained via WTMMMM, MFDFA and its theory value. **(D)** Comparison of the singularity spectrum obtained via WTMMMM, MFDFA and its theory value. Both analyses have good agreement with the theoretical line, but MFDFA performs slightly better.

extrapolate the maxima chains to $s = 0$ and find the approximate location of the singularities along with their exponents (by determining the exponent of the power law of the modulus $\mathcal{M}^{\Phi}[f]$ vs. s). It should be noted that some maxima chains exist because of local maxima in the data, not singular points. In these cases, the exponents of the power-laws are determined by the order of the chosen wavelet and, it is easy to distinguish them.

Bibliography

[1] Lord Rayleigh. "XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 10.60 (1880), pp. 73–78.

[2] Karl Pearson. "The problem of the random walk". In: *Nature* 72.1865 (1905), pp. 294–294.

[3] Marian V Smoluchowski. "Zur kinetischen Theorie der Brownschen Molekular Bewegung und der Suspensionen". In: *Ann. d. Phys.* 21 (1906), p. 756.

[4] Andrey A. Markov. *Wahrscheinlichkeitsrechnung*. BG Teubner, 1912.

[5] Subrahmanyan Chandrasekhar. "Stochastic problems in physics and astronomy". In: *Reviews of modern physics* 15.1 (1943), p. 1.

[6] Albert Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der physik* 4 (1905).

[7] Albert Einstein. "Zur theorie der brownschen bewegung". In: *Annalen der physik* 324.2 (1906), pp. 371–381.

[8] Adriaan D Fokker. "Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld". In: *Annalen der physik* 348.5 (1914), pp. 810–820.

[9] Max Planck. "Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie". In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* (1917).

[10] Paul Langevin. "Sur la théorie du mouvement brownien". In: *Compt. Rendus* 146 (1908), pp. 530–533.

[11] Don S Lemons and Anthony Gythiel. "Paul langevin's 1908 paper "on the theory of brownian motion"["sur la théorie du mouvement brownien," cr acad. sci.(paris) 146, 530–533 (1908)]". In: *American Journal of Physics* 65.11 (1997), pp. 1079–1081.

[12] Hendrik A Kramers. "Brownian motion in a field of force and the diffusion model of chemical reactions". In: *Physica* 7.4 (1940), pp. 284–304.

[13] W Pauli. *Festschri/t zum 6o. Geburtstage A. Sommer/dds*. 1928.

[14] Leon Van Hove. "The approach to equilibrium in quantum statistics: A perturbation treatment to general order". In: *Physica* 23.1-5 (1957), pp. 441–480.

[15] Donald A McQuarrie. "Stochastic approach to chemical kinetics". In: *Journal of applied probability* 4.3 (1967), pp. 413–478.

[16] Alexander N Pisarchik and Ulrike Feudel. "Control of multistability". In: *Physics Reports* 540.4 (2014), pp. 167–218.

[17] Mark Iosifovich Freidlin and Alexander D Wentzell. "Random perturbations". In: *Random perturbations of dynamical systems*. Springer, 2012.

[18] Changsong Zhou and Jürgen Kurths. "Noise-induced phase synchronization and synchronization transitions in chaotic oscillators". In: *Physical review letters* 88.23 (2002), p. 230602.

[19] Henrik Boije et al. "The independent probabilistic firing of transcription factors: a paradigm for clonal variability in the zebrafish retina". In: *Developmental cell* 34.5 (2015), pp. 532–543.

[20] Masoud Mohseni et al. *Quantum effects in biology*. Cambridge University Press, 2014.

[21] Gašper Tkačik and William Bialek. "Information processing in living systems". In: *Annual Review of Condensed Matter Physics* 7 (2016), pp. 89–117.

[22] Coco van Boxtel et al. "Taking chances and making mistakes: non-genetic phenotypic heterogeneity and its consequences for surviving in dynamic environments". In: *Journal of The Royal Society Interface* 14.132 (2017), p. 20170141.

[23] Thomas O Perry. "Dormancy of Trees in Winter: Photoperiod is only one of the variables which interact to control leaf fall and other dormancy phenomena." In: *Science* 171.3966 (1971), pp. 29–36.

[24] Constance A Harrington, Peter J Gould, and J Bradley St Clair. "Modeling the effects of winter environment on dormancy release of Douglas-fir". In: *Forest Ecology and Management* 259.4 (2010), pp. 798–808.

[25] Erez Dekel and Uri Alon. "Optimality and evolutionary tuning of the expression level of a protein". In: *Nature* 436.7050 (2005), pp. 588–592.

[26] Guillaume Lambert and Edo Kussell. "Memory and fitness optimization of bacteria under fluctuating environments". In: *PLoS genetics* 10.9 (2014), e1004556.

[27] Aaron M New et al. "Different levels of catabolite repression optimize growth in stable and variable environments". In: *PLoS biology* 12.1 (2014), e1001764.

[28] Howard C Berg and Edward M Purcell. "Physics of chemoreception". In: *Biophysical journal* 20.2 (1977), pp. 193–219.

[29] Thierry Mora. "Physical limit to concentration sensing amid spurious ligands". In: *Physical review letters* 115.3 (2015), p. 038102.

[30] Christoph Zechner et al. "Molecular circuits for dynamic noise filtering". In: *Proceedings of the National Academy of Sciences* 113.17 (2016), pp. 4729–4734.

[31] Tetsuya J Kobayashi. "Implementation of dynamic Bayesian decision making by intracellular kinetics". In: *Physical review letters* 104.22 (2010), p. 228104.

[32] Sean Fancher and Andrew Mugler. "Fundamental limits to collective concentration sensing in cell populations". In: *Physical review letters* 118.7 (2017), p. 078101.

[33] Robert Gilbert. *Physical biology of the cell, by Rob Phillips, Jane Kondev and Julie Theriot*. 2009.

[34] Benno Müller-Hill. "The lac operon". In: *The lac Operon*. de Gruyter, 2011.

[35] Gordon Churchward, Hans Bremer, and R Young. "Transcription in bacteria at different DNA concentrations". In: *Journal of bacteriology* 150.2 (1982), pp. 572–581.

[36] Seth A Darst, Elizabeth W Kubalek, and Roger D Kornberg. "Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography". In: *Nature* 340.6236 (1989), pp. 730–732.

[37] Keith Robison, Abigail Manson McGuire, and George M Church. "A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete Escherichia coli K-12 genome". In: *Journal of molecular biology* 284.2 (1998), pp. 241–254.

[38] Christoph Zechner et al. "Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings". In: *Nature methods* 11.2 (2014), pp. 197–202.

[39] Oleg Yarchuk et al. "Interdependence of translation, transcription and mRNA degradation in the lacZ gene". In: *Journal of molecular biology* 226.3 (1992), pp. 581–596.

[40] John L Spudich and Daniel E Koshland. "Non-genetic individuality: chance in the single cell". In: *Nature* 262.5568 (1976), pp. 467–471.

[41] Harley H McAdams and Adam Arkin. "Stochastic mechanisms in gene expression". In: *Proceedings of the National Academy of Sciences* 94.3 (1997), pp. 814–819.

[42] Michael B Elowitz et al. "Protein mobility in the cytoplasm of Escherichia coli". In: *Journal of bacteriology* 181.1 (1999), pp. 197–203.

[43] Dann Huh and Johan Paulsson. "Non-genetic heterogeneity from stochastic partitioning at cell division". In: *Nature genetics* 43.2 (2011), pp. 95–100.

[44] Ines SC Baptista and Andre S Ribeiro. "Stochastic models coupling gene expression and partitioning in cell division in Escherichia coli". In: *BioSystems* 193 (2020), p. 104154.

[45] Michael B Elowitz et al. "Stochastic gene expression in a single cell". In: *Science* 297.5584 (2002), pp. 1183–1186.

[46] Adam Klosin et al. "Phase separation provides a mechanism to reduce noise in cells". In: *Science* 367.6476 (2020), pp. 464–468.

[47] Shu-Wen Teng et al. "Measurement of the copy number of the master quorum-sensing regulator of a bacterial cell". In: *Biophysical journal* 98.9 (2010), pp. 2024–2031.

[48] Thibaud Taillefumier and Ned S Wingreen. "Optimal census by quorum sensing". In: *PLoS computational biology* 11.5 (2015), e1004238.

[49] Tetsuya Tabata and Yuki Takei. "Morphogens, their identification and regulation". In: (2004).

[50] Tobias Bollenbach et al. "Precision of the Dpp gradient". In: (2008).

[51] Raymond Cheong et al. "Information transduction capacity of noisy biochemical signaling networks". In: *science* 334.6054 (2011), pp. 354–358.

[52] Wing-Cheong Lo. "Morphogen gradient with expansion-repression mechanism: Steady-state and robustness studies". In: *Discrete & Continuous Dynamical Systems-B* 19.3 (2014), p. 775.

[53] Danny Ben-Zvi and Naama Barkai. "Scaling of morphogen gradients by an expansion-repression integral feedback control". In: *Proceedings of the National Academy of Sciences* 107.15 (2010), pp. 6924–6929.

[54] Daniel T Gillespie. "Exact stochastic simulation of coupled chemical reactions". In: *The journal of physical chemistry* 81.25 (1977), pp. 2340–2361.

[55] Daniel T Gillespie. "The chemical Langevin equation". In: *The Journal of Chemical Physics* 113.1 (2000), pp. 297–306.

[56] Lorenzo Duso and Christoph Zechner. "Stochastic reaction networks in dynamic compartment populations". In: *Proceedings of the National Academy of Sciences* 117.37 (2020), pp. 22674–22683.

[57] Johan Elf, Andreas Doncic, and Mans Ehrenberg. "Mesoscopic reaction-diffusion in intracellular signaling". In: *Fluctuations and noise in biological, biophysical, and biomedical systems*. Vol. 5110. SPIE. 2003, pp. 114–124.

[58] Robin J Wilson. *Introduction to Graph Theory*. Longman, 2010.

[59] Albert-László Barabási and Mãrton Pósfai. *Network Science*. Cambridge University Press, 2016.

[60] Jianxi Gao, Baruch Barzel, and Albert-László Barabási. "Universal resilience patterns in complex networks". In: *Nature* 530.7590 (2016), pp. 307–312.

[61] Christoph Kirst, Marc Timme, and Demian Battaglia. "Dynamic information routing in complex networks". In: *Nature communications* 7.1 (2016), pp. 1–9.

[62] Holger Kantz and Thomas Schreiber. *Nonlinear time series analysis*. Vol. 7. Cambridge university press, 2004.

[63] Jens Prusseit and Klaus Lehnertz. "Stochastic qualifiers of epileptic brain dynamics". In: *Physical review letters* 98.13 (2007), p. 138103.

[64] Fatemeh Ghasemi et al. "Analysis of non-stationary data for heart-rate fluctuations in terms of drift and diffusion coefficients". In: *Journal of biological physics* 32.2 (2006), pp. 117–128.

[65] Rudolf Friedrich et al. "Approaching complexity by stochastic methods: From biological systems to turbulence". In: *Physics Reports* 506.5 (2011), pp. 87–162.

[66] M Reza Rahimi Tabar. *Analysis and data-based reconstruction of complex nonlinear dynamical systems*. Vol. 730. Springer, 2019.

[67] Pierre A Haas et al. "The noisy basis of morphogenesis: Mechanisms and mechanics of cell sheet folding inferred from developmental variability". In: *PLoS biology* 16.7 (2018), e2005536.

[68] Brian Munsky, Brooke Trinh, and Mustafa Khammash. "Listening to the noise: random fluctuations reveal gene network parameters". In: *Molecular systems biology* 5.1 (2009), p. 318.

[69] Mohammadreza Bahadorian, Christoph Zechner, and Carl D Modes. "Gift of gab: Probing the limits of dynamic concentration-sensing across a network of communicating cells". In: *Physical Review Research* 2.2 (2020), p. 023403.

[70] Wouter-Jan Rappel and Herbert Levine. "Receptor noise and directional sensing in eukaryotic chemotaxis". In: *Physical review letters* 100.22 (2008), p. 228101.

[71] Bo Hu et al. "Physical limits on cellular sensing of spatial gradients". In: *Physical review letters* 105.4 (2010), p. 048104.

[72] Thierry Mora and Ned S Wingreen. "Limits of sensing temporal concentration changes by single cells". In: *Physical review letters* 104.24 (2010), p. 248101.

[73] Cassio G Lopes and Ali H Sayed. "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis". In: *IEEE Transactions on Signal Processing* 56.7 (2008), pp. 3122–3136.

[74] Sheng-Yuan Tu and Ali H Sayed. "Mobile adaptive networks". In: *IEEE Journal of Selected Topics in Signal Processing* 5.4 (2011), pp. 649–664.

[75] Ying Liu et al. "Distributed estimation over complex networks". In: *Information Sciences* 197 (2012), pp. 91–104.

[76] Steven M Kay. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

[77] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier, 1992.

[78] David F Anderson and Thomas G Kurtz. "Continuous time Markov chain models for chemical reaction networks". In: *Design and analysis of biomolecular circuits*. Springer, 2011, pp. 3–42.

[79] Daniel T Gillespie. "Stochastic simulation of chemical kinetics". In: *Annu. Rev. Phys. Chem.* 58 (2007), pp. 35–55.

[80] Thorsten Erdmann, Martin Howard, and Pieter Rein Ten Wolde. "Role of spatial averaging in the precision of gene expression patterns". In: *Physical review letters* 103.25 (2009), p. 258101.

[81] Thomas R Sokolowski and Gašper Tkačik. "Optimizing information flow in small genetic networks. IV. Spatial coupling". In: *Physical Review E* 91.6 (2015), p. 062710.

[82] Albert-László Barabási and Réka Albert. "Emergence of scaling in random networks". In: *science* 286.5439 (1999), pp. 509–512.

[83] Linyuan Lü, Duan-Bing Chen, and Tao Zhou. "The small world yields the most effective information spreading". In: *New Journal of Physics* 13.12 (2011), p. 123005.

[84] Duncan J Watts and Steven H Strogatz. "Collective dynamics of 'small-world'networks". In: *nature* 393.6684 (1998), p. 440.

[85] Réka Albert and Albert-László Barabási. "Statistical mechanics of complex networks". In: *Reviews of modern physics* 74.1 (2002), p. 47.

[86] Bruno Antonny. "Mechanisms of membrane curvature sensing". In: *Annual review of biochemistry* 80 (2011), pp. 101–123.

[87] Mohammadreza Bahadorian and Carl D Modes. "ANDOR and beyond: dynamically switchable logic gates as modules for flexible information processing in biochemical regulatory networks". In: *arXiv preprint arXiv:2108.00950* (2021).

[88] Jeffrey M Black. "Preflight signalling in swans: a mechanism for group cohesion and flock formation". In: *Ethology* 79.2 (1988), pp. 143–157.

[89] Julia Trillmich, Claudia Fichtel, and Peter Kappeler. "Coordination of group movements in wild Verreaux's sifakas (Propithecus verreauxi)". In: *Behaviour* 141.9 (2004), pp. 1103–1120.

[90] Nigel R Franks et al. "Speed versus accuracy in collective decision making". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270.1532 (2003), pp. 2457–2463.

[91] Erik T Frank and K Eduard Linsenmair. "Individual versus collective decision making: optimal foraging in the group-hunting termite specialist Megaponera analis". In: *Animal Behaviour* 130 (2017), pp. 27–35.

[92] Eric R Kandel et al. *Principles of neural science*. Vol. 4. McGraw-hill New York, 2000.

[93] Gašper Tkačik and Aleksandra M Walczak. "Information transmission in genetic regulatory networks: a review". In: *Journal of Physics: Condensed Matter* 23.15 (2011), p. 153102.

[94] Gašper Tkačik, Curtis G Callan Jr, and William Bialek. "Information capacity of genetic regulatory elements". In: *Physical Review E* 78.1 (2008), p. 011910.

[95] Eric H Davidson. *The regulatory genome: gene regulatory networks in development and evolution*. Elsevier, 2010.

[96] Nicholas M Luscombe et al. "Genomic analysis of regulatory network dynamics reveals large topological changes". In: *Nature* 431.7006 (2004), p. 308.

[97] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC Press, 2018.

[98] Alba Jiménez et al. "A spectrum of modularity in multi-functional gene circuits". In: *Molecular systems biology* 13.4 (2017).

[99] Joshua L Payne and Andreas Wagner. "Function does not follow form in gene regulatory circuits". In: *Scientific reports* 5 (2015), p. 13015.

[100] Ruben Perez-Carrasco et al. "Combining a toggle switch and a repressilator within the ac-dc circuit generates distinct dynamical behaviors". In: *Cell systems* 6.4 (2018), pp. 521–530.

[101] Timothy S Gardner, Charles R Cantor, and James J Collins. "Construction of a genetic toggle switch in Escherichia coli". In: *Nature* 403.6767 (2000), p. 339.

[102] Francisco JP Lopes et al. "Spatial bistability generates hunchback expression sharpness in the Drosophila embryo". In: *PLoS computational biology* 4.9 (2008), e1000184.

[103] Roy E Crist, Wu Li, and Charles D Gilbert. "Learning to see: experience and attention in primary visual cortex". In: *Nature neuroscience* 4.5 (2001), pp. 519–525.

[104] Demian Battaglia et al. "Dynamic effective connectivity of inter-areal brain circuits". In: *PLoS computational biology* 8.3 (2012), e1002438.

[105] Adriano Barenco et al. "Conditional quantum dynamics and logic gates". In: *Physical Review Letters* 74.20 (1995), p. 4083.

[106] Satoshi Matsuura et al. "Synthetic RNA-based logic computation in mammalian cells". In: *Nature communications* 9.1 (2018), p. 4847.

[107] Jeff Hasty, David McMillen, and James J Collins. "Engineered gene circuits". In: *Nature* 420.6912 (2002), p. 224.

[108] Evgeny Katz and Vladimir Privman. "Enzyme-based logic systems for information processing". In: *Chemical Society Reviews* 39.5 (2010), pp. 1835–1857.

[109] Christoph Kirst, Carl D Modes, and Marcelo O Magnasco. "Shifting attention to dynamics: Self-reconfiguration of neural networks". In: *Current Opinion in Systems Biology* 3 (2017), pp. 132–140.

[110] Lulu Qian and Erik Winfree. "Scaling up digital circuit computation with DNA strand displacement cascades". In: *Science* 332.6034 (2011), pp. 1196–1201.

[111] Elisa Nerli, Mauricio Rocha-Martins, and Caren Norden. "Asymmetric neurogenic commitment of retinal progenitors involves Notch through the endocytic pathway". In: *Elife* 9 (2020), e60462.

[112] Subir Kumar Sarkar, Asish Kumar De, and Souvik Sarkar. *Foundation of Digital Electronics and Logic Design*. CRC Press, 2014.

[113] Chunbo Lou et al. "Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch". In: *Molecular systems biology* 6.1 (2010), p. 350.

[114] Salman F Banani et al. "Biomolecular condensates: organizers of cellular biochemistry". In: *Nature reviews Molecular cell biology* 18.5 (2017), pp. 285–298.

[115] David T Gonzales, Christoph Zechner, and TY Dora Tang. "Building synthetic multicellular systems using bottom-up approaches". In: *Current Opinion in Systems Biology* 24 (2020), pp. 56–63.

[116] Ali Kinkhabwala and Philippe IH Bastiaens. "Spatial aspects of intracellular information processing". In: *Current opinion in genetics & development* 20.1 (2010), pp. 31–40.

[117] Gary S Stein and Jane B Lian. "Molecular mechanisms mediating proliferation/differentiation interrelationships during progressive development of the osteoblast phenotype". In: *Endocrine reviews* 14.4 (1993), pp. 424–442.

[118] Nigel C Harrison, Ruth Diez del Corral, and Bakhtier Vasiev. "Coordination of cell differentiation and migration in mathematical models of caudal embryonic axis extension". In: *PloS one* 6.7 (2011), e22700.

[119] Levi Wood, Roger Kamm, and Harry Asada. "Stochastic modeling and identification of emergent behaviors of an Endothelial Cell population in angiogenic pattern formation". In: *The International Journal of Robotics Research* 30.6 (2011), pp. 659–677.

[120] Madeline Galbraith, Federico Bocci, and José N Onuchic. "Stochastic fluctuations promote ordered pattern formation of cells in the Notch-Delta signaling pathway". In: *arXiv preprint arXiv:2202.00763* (2022).

[121] Dmitry Bratsun. "Protein pattern formation induced by the joint effect of noise and delay in a multi-cellular system". In: *Mathematical Modelling of Natural Phenomena* 17 (2022), p. 16.

[122] Benoit Mandelbrot. "How long is the coast of Britain? Statistical self-similarity and fractional dimension". In: *science* 156.3775 (1967), pp. 636–638.

[123] Tamás Vicsek. *Fractal Growth Phenomena*. 2nd. WORLD SCIENTIFIC, 1992.

[124] Harry Eugene Stanley and Nicole Ostrowsky. *Random fluctuations and pattern growth: experiments and models*. Vol. 157. Springer Science & Business Media, 1988.

[125] Jens Feder. *Fractals*. Springer Science & Business Media, 1988.

[126] T Vicsek, F Family, and P Meakin. "Multifractal geometry of diffusion-limited aggregates". In: *EPL (Europhysics Letters)* 12.3 (1990), p. 217.

[127] Benoit B Mandelbrot and Carl JG Evertsz. "The potential distribution around growing fractal clusters". In: *Nature* 348.6297 (1990), pp. 143–145.

[128] Carl JG Evertsz and Benoit B Mandelbrot. "Multifractal measures". In: *Chaos and Fractals: New Frontiers in Science, Springer-Verlag, New York* 1092 (1992), pp. 921–953.

[129] Natalie Reznikov et al. "Fractal-like hierarchical organization of bone begins at the nanoscale". In: *Science* 360.6388 (2018), eaao2189.

[130] Elizabeth A Zimmermann and Robert O Ritchie. "Bone as a structural material". In: *Advanced healthcare materials* 4.9 (2015), pp. 1287–1304.

[131] Yuping Li and Conrado Aparicio. "Discerning the subfibrillar structure of mineralized collagen fibrils: a model for the ultrastructure of bone". In: *PloS one* 8.9 (2013), e76782.

[132] Baptiste Depalle et al. "Large deformation mechanisms, plasticity, and failure of an individual collagen fibril with different mineral content". In: *Journal of Bone and Mineral Research* 31.2 (2016), pp. 380–390.

[133] Petar Milovanovic, Marija Djuric, and Zlatko Rakocevic. "Age-dependence of power spectral density and fractal dimension of bone mineralized matrix in atomic force microscope topography images: potential correlates of bone tissue age and bone fragility in female femoral neck trabeculae". In: *Journal of anatomy* 221.5 (2012), pp. 427–433.

[134] Claude-Laurent Benhamou et al. "Fractal analysis of radiographic trabecular bone texture and bone mineral density: two complementary parameters

related to osteoporotic fractures". In: *Journal of bone and mineral research* 16.4 (2001), pp. 697–704.

[135] Mamoru Ishii et al. "The development of the calvarial bones and sutures and the pathophysiology of craniosynostosis". In: *Current topics in developmental biology* 115 (2015), pp. 131–156.

[136] Heather E White, Anjali Goswami, and Abigail S Tucker. "The intertwined evolution and development of sutures and cranial morphology". In: *Frontiers in Cell and Developmental Biology* 9 (2021), p. 690.

[137] Joan T Richtsmeier and Kevin Flaherty. "Hand in glove: brain and skull in development and dysmorphogenesis". In: *Acta neuropathologica* 125.4 (2013), pp. 469–489.

[138] Joan T Richtsmeier et al. "Phenotypic integration of neurocranium and brain". In: *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 306.4 (2006), pp. 360–378.

[139] Albert-László Barabási and Tamás Vicsek. "Multifractality of self-affine fractals". In: *Physical Review A* 44.4 (1991), p. 2730.

[140] A Arneodo, NGRS Decoster, and SG 2 Roux. "A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and anisotropic random rough surfaces". In: *The European Physical Journal B-Condensed Matter and Complex Systems* 15.3 (2000), pp. 567–600.

[141] Paweł Oświęcimka, Jarosław Kwapień, and Stanisław Drożdż. "Wavelet versus detrended fluctuation analysis of multifractal structures". In: *Physical Review E* 74.1 (2006), p. 016103.

[142] Jan W Kantelhardt et al. "Multifractal detrended fluctuation analysis of nonstationary time series". In: *Physica A: Statistical Mechanics and its Applications* 316.1-4 (2002), pp. 87–114.

[143] Gao-Feng Gu and Wei-Xing Zhou. "Detrended fluctuation analysis for fractals and multifractals in higher dimensions". In: *Physical Review E* 74.6 (2006), p. 061104.

[144] Ata Madanchi et al. "Strong short-term non-linearity of solar irradiance fluctuations". In: *Solar Energy* 144 (2017), pp. 1–9.

[145] Zahra Fayyaz et al. "Multifractal detrended fluctuation analysis of continuous neural time series in primate visual cortex". In: *Journal of neuroscience methods* 312 (2019), pp. 84–92.

[146] Victor Manuel San Martin and Alejandra Figliola. "Application of Multifractal Analysis to Segmentation of Water Bodies in Optical and Synthetic Aperture Radar Satellite Images". In: *arXiv preprint arXiv:1604.02488* (2016).

[147] Jean-François Muzy, Emmanuel Bacry, and Alain Arneodo. "Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method". In: *Physical review E* 47.2 (1993), p. 875.

[148] John Canny. "A computational approach to edge detection". In: *IEEE Transactions on pattern analysis and machine intelligence* 6 (1986), pp. 679–698.

[149] Yan Liu et al. "Hierarchical and non-hierarchical mineralisation of collagen". In: *Biomaterials* 32.5 (2011), pp. 1291–1300.

[150] William J Landis et al. "Mineralization of collagen may occur on fibril surfaces: evidence from conventional and high-voltage electron microscopy and three-dimensional imaging". In: *Journal of structural biology* 117.1 (1996), pp. 24–35.

[151] Kenneth R Wilmarth and John R Froines. "In vitro and in vivo inhibition of lysyl oxidase byaminopropionitriles". In: *Journal of Toxicology and Environmental Health, Part A Current Issues* 37.3 (1992), pp. 411–423.

[152] Wei Ren and Randal W Beard. "Consensus seeking in multiagent systems under dynamically changing interaction topologies". In: *IEEE Transactions on automatic control* 50.5 (2005), pp. 655–661.

[153] Xiang Zhou, Weiqing Ren, and Weinan E. "Adaptive minimum action method for the study of rare events". In: *The Journal of chemical physics* 128.10 (2008), p. 104111.

[154] M Sadegh Movahed et al. "Multifractal detrended fluctuation analysis of sunspot time series". In: *Journal of Statistical Mechanics: Theory and Experiment* 2006.02 (2006), P02003.