UNIVERSITY OF SOUTHAMPTON

Faculty of Social, Human and Mathematical Sciences

- Mathematical Sciences -

Noise Processing by Networks

by

Styliani Kontogeorgaki

Thesis for the degree of Master of Philosophy

October 2016

Declaration of Authorship

I, Styliani Kontogeorgaki, declare that the thesis entitled *"Noise Processing by Networks"* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:

  S. Kontogeorgaki, R. Sanchez-Garcia, R. Ewing, K. Zygalakis, and B. D. MacArthur, "Noise-processing by signaling networks." biorxiv:10.1101/075366, Submitted in *Sci. Rep.*, 2016

Signed: .....

Date: .....

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

Faculty of Social, Human and Mathematical Sciences

Mathematical Sciences

Master of Philosophy

NOISE PROCESSING BY NETWORKS

by Styliani Kontogeorgaki

Cell behaviour is determined by complex molecular regulatory networks. Signalling networks are particularly important since they are responsible for the robust transmission of noisy environmental information to the cell's nucleus. However, although important signalling pathways have been well studied, and the manner of noise propagation through regulatory networks has been discussed, the relationship between network architecture and the cell's ability to process environmental noise is not well understood. To approach this problem in this thesis we derive a mathematical formula relating a network's structure to its noise processing ability. We find that noise processing is highly affected by the networks complexity, and in particular by the number and length of the weighted paths from the noisy input(s) to the output. In order to explore the utility of this mathematical expression, we apply it to the regulatory network for pluripotency in mouse embryonic stem (ES) cells and assess the effects of network topology on the propagation of noise through this system. We conclude by using the underlying theory to explain the interaction patterns in the ES cell's transcriptional circuit.

Acknowledgements

I would like to thank my supervisors Dr Ben Macarthur and Dr Konstantinos Zygalakis who were the ones that shared and challenged ideas, made me aware of my weaknesses and suggested me ways to be improved. Thank you for the patient guidance, the constant motivation and encouragement to accomplish this research. Thank you for the integrity and honour to contribute into a respected academic paper. I greatly appreciate the support and influence received from Ben, as well as the opportunity to explore this part of the systems biology world. In particular, I would like to mention my appreciation for Kostas's continued feedback and suggestions, despite him having moved to a new position in another university. I would also like to thank my supervisor Dr Rob Ewing for our biology discussions and also Dr Ruben Sanchez-Garcia for their contribution in our paper.

I would like to thank the Institute for Life Sciences (IfLS) of Southampton University for my research funding and for giving me the opportunity to attend to interesting workshops and conferences. I would like to thank the Mathematical Sciences of the University of Southampton which allowed me to go through my research and meet all these interesting people.

I would like to thank my group people: Rosanna, Patrick, Sonya and Enrico for our discussions, seminars, co-attendance in conferences, and all their useful comments.

I would also like to thank my maths friends that made my time in Southampton these last three years amazing. I would like to thank *señorita* Marta for being a real friend and adopted me in the GR group. Thank you for everything, you are the best. Thank you Greg for always being there to give feedback, and in combination with Yafet for all the discussions about life and science, and for the endless times of fun. Thank you Vanessa for all the feedback and the infinite baking for birthday and seminar cakes. Thank you Emma, Christian and Ingrid.

I would like to express my gratitude to Thomas for standing by me all this time, reading endlessly all the biology and mathematics, giving feedback whenever needed, cooked for me, taking care of me, supporting me in hard times and sharing all the happiness; thank you for all the inspiration.

I would like to thank my family; mum, dad and my sister. Although they are far away, they encourage and support me as much as anyone can and much more.

Without all of you this thesis would not be possible.

This thesis is dedicated to Νάσια

Contents

List of Figures

List of Tables

Nomenclature

$\boldsymbol{A}$	adjacency matrix
$\boldsymbol{K}(t)$	covariance matrix
$\boldsymbol{W}(t)$	Wiener process
$\boldsymbol{\Sigma}$	noise coefficient matrix
$\mathcal{G}$	graph
$w(P)$	weight of walk P between two nodes
$\boldsymbol{\nu_j}$	stoichiometric vector
$\boldsymbol{n}(t)$	state vector
$p^*(n)$	stationary distribution
CH	glycogen synthase kinase 3 inhibitor (TGF-β)
Lif	cytokine Leukemia inhibitory factor
MAP Kinase	mitogen-activated protein kinase
Oct4	octamer DNA binding transcription factor 4
PD	mitogen-activated protein kinase (Mek)
Stat3	signal transducer and activator of transcription 3
TGF-β	transforming growth factor beta
Wnt	wingless-integrated
BM	Brownian motion
CFPE	chemical Fokker-Planck equation
CLE	chemical Langevin equation
CME	chemical master equation

DNA deoxyribonucleic acid

ES cell embryonic stem cell

mRNA messenger ribonucleic acid

ODE ordinary differential equation

PPI protein-protein interaction

RNA ribonucleic acid

RRE reaction rate equation

SDE stochastic differential equation

SSA stochastic simulation algorithm

TF transcription factor

Chapter 1

Introduction

Recent research developments in cell biology increasingly show the utility of networks for better understanding and exploring the cell's organisation and function [9–11]. Regulatory signalling pathways and their underlying networks are particularly important for their role in transmitting noisy environmental information to the cell nucleus [6, 12–15]. However, a malfunction in signal transduction is commonly associated with cancer, as well as with deficiencies in embryonic development [16–20]. Although specific molecular mechanisms that are involved in noise signalling, along with the dynamics of these functions, are well defined and studied, the role by which signalling networks receive environmental noisy signals remains ambiguous [21–26]. To address this problem we focus on exploring the relation between these networks' architecture and their effect on environmental noise processing; the graph representation of such networks can shed light on their latent properties and allows the assessment of the important pathways. To gain insight into the way signals propagate across such organizations, we will use tools from graph theory and stochastic differential equations (SDEs). In this thesis, we consider and model 'noise' as the inherent stochasticity of the system, as well as the external environmental noisy signals, which could be thought of as bearing essential information or just random environmental fluctuations. First we discuss biochemical systems considering the former type of noise, which has been studied at length [27–29]. Later we focus on common signalling motifs to explore the relationship between network structure and noise transduction from a noisy input(s).

1.1 Biological Networks

Biological systems can be represented by networks in order to reveal and understand the underlying biological processes, as well as to interpret interactions or states of biomolecules that are provided in biological data [30, 31]. The analysis of biological networks elucidate the behaviour of such complex systems. Therefore, biological systems with a large copy number of proteins, genes, molecular species or signals can be

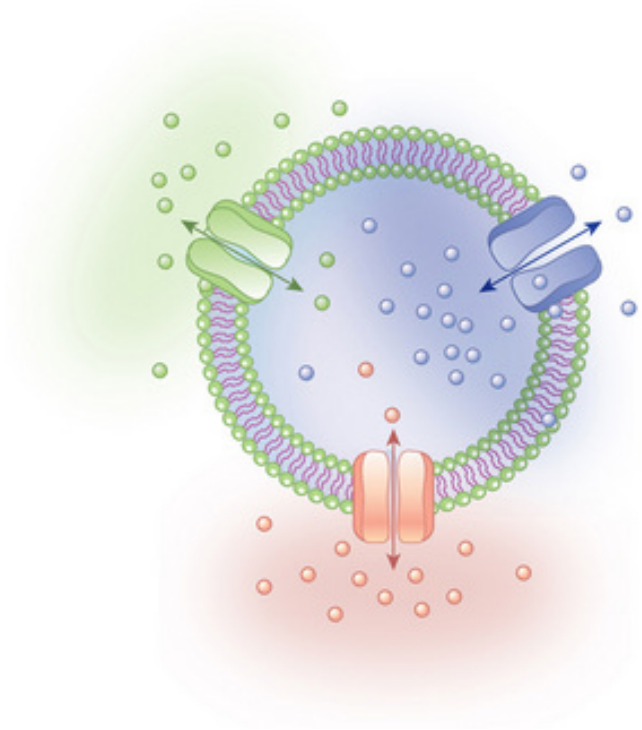


Figure 1.1: The image shows how proteins (different colours) pass the membrane of the cell into the cytoplasm and move into the inner cell, transferring external signals. Image source [2].

better understood, controlled and predicted. Important properties such as topology, expression patterns, interactions, or signal transduction can be revealed [30, 32]. Important functions in the inner cell are described via a variety of complex molecular networks [9, 12]. These functions can be metabolic, DNA synthesis or repair, protein to protein interactions, or information processing. They are implemented by transcription regulatory networks or signal transmission networks among proteins, transcription factors, genes and other regulatory molecules [3, 11, 30, 33–36]. Here we are discussing some important biological networks such as protein to protein interactions, transcription regulatory networks, and we are focusing on interactions among molecules in signalling networks.

1.1.1 Protein-Protein Interaction Networks

Proteins are biomolecules that consist of a long sequence of twenty different types of amino acids. The position of each amino acid in the sequence determines the protein's unique three-dimensional structure, specifying their activity [31, 36]. The role of proteins is constitutive not only because they are elementary units of the cell, but also for the reason that they perform an abundance of functions within the cell. Proteins can act as enzymes and incite the activity of chemical reactions involved in metabolic

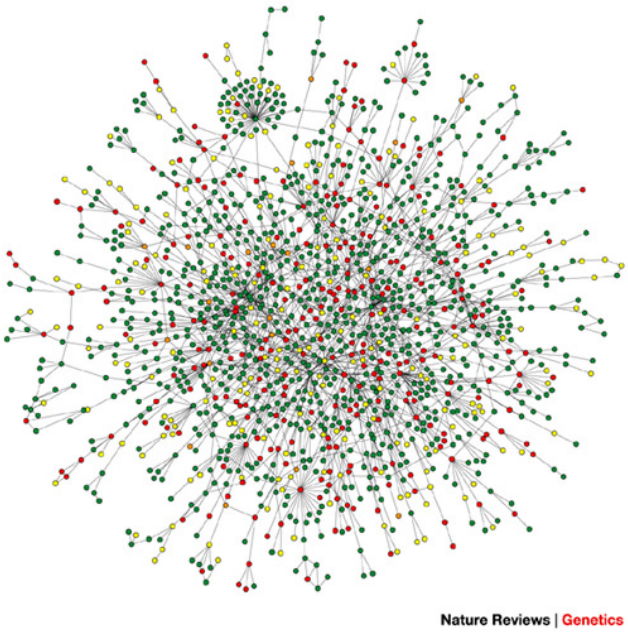


Figure 1.2: Protein-protein interaction network in yeast. The structure of highly connected proteins is three times more important than the proteins that are not highly connected [3].

processes. Other proteins either are responsible for the cell's structure and organization, or implement transcriptional regulation. Alternatively, proteins can deliver signals among cells or from the membrane to the cell's nucleus [2, 12, 30, 31, 36, 37]. Fig. 1.1 illustrates different proteins that transmit signals from the cell's membrane into the inner cell. Proteins interact with other proteins in order to form protein complexes. The collection of all protein-protein interactions (PPI) forms a protein-protein interaction network [30, 31, 36]. In order to identify physical PPI, many experimental techniques have been developed [3, 38–41]. The structure of PPI networks can play an important role [3, 42]. For example, in yeast the robustness of the system against mutations not only is affected by the structure of PPIs, but also on how strongly proteins are connected. The more highly the proteins are linked, the more fundamental is their role in the network; in highly connected proteins (see Fig. 1.2), their structure is three times more important than the proteins that are not highly connected [3].

1.1.2 Transcription Regulatory Networks

The way a cell responds to internal or external stimuli can be described by gene regulatory networks and signalling pathways [31, 43]. Cells contain the genetic information in their nucleus as sequences of nucleic acids. All the cells have the same deoxyribonucleic acid (DNA), which in its turn encodes genes as required to function [11, 36, 43]. The way the cell decides how to function depends on the regulation of the genetic

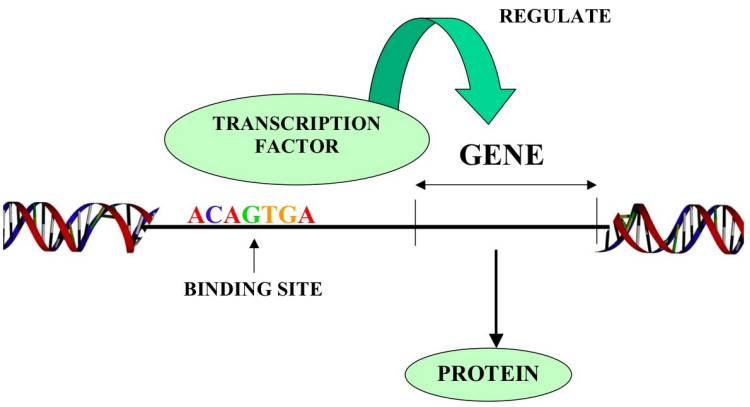


Figure 1.3: Transcription process. Image is taken from [4].

material in the DNA. This genetic information is transcribed through the process of gene expression [2, 36].

The first level of gene expression is the transcription, which is the bridge from DNA to ribonucleic acid (RNA). The enzyme RNA polymerase along with the transcription factors (TFs) execute the transcription process. TFs are molecules characterized by the property of binding on a segment of DNA, in order to produce a copy of it into mRNA. This copy is formed by the RNA polymerase, which first separates the DNA helix and then develops a new RNA strand. During the transcription process, if the DNA unwinds, the mRNA copies the genetic information in the nucleus of the cell. Fig. 1.3 describes how a transcription factor binds on a sequence of DNA and regulates the gene's decision (activates a gene) for a production of a protein [35, 36].

If a gene is encoded by a transcription factor, the information contained in the mRNA is transcribed into a protein; this process is the translation. The whole process occurs in the ribosome, which is the machinery of the synthesis of proteins. In more detail, once mRNA has transcribed the genetic information, it travels out of the nucleus into the cytoplasm; then, mRNA decodes the information in the ribosome and creates an amino acid chain that is translated into a protein. The entire process from DNA to proteins, transcription to translation is depicted in Fig. 1.4.

1.1.3 Signalling Networks

Signals in the cell are received and processed by signalling pathways from its external local micro-environment, as well as from various parts within the cell. In mammalian organisms, cell signalling pathways intercommunicate with each other and complex networks structures arise [12, 44, 45]. These signalling networks affect the coordination of cell activities and the signal processing in the cell's environment [3, 31, 46–50]. Such networks give information about the nature of the interactant molecules, along with

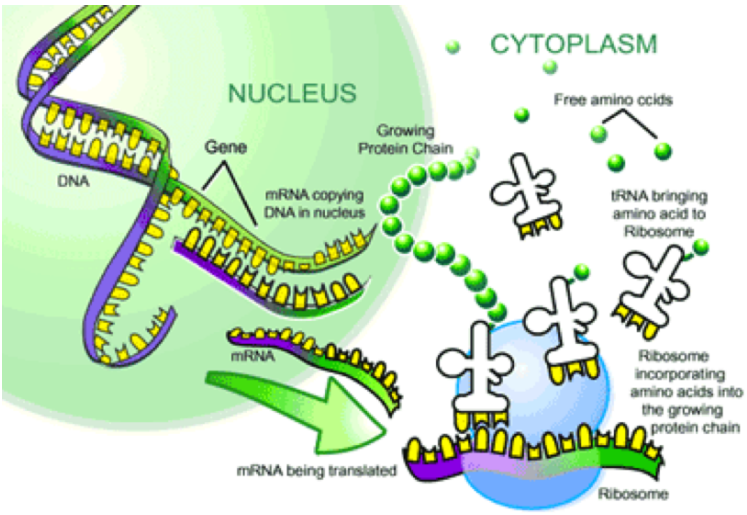


Figure 1.4: Transcription and translation process. Image is taken from [5].

their orientation, timescale and the influence on the targets of interest. Signalling networks are of particular importance in maintaining robust cellular identities, since they mediate noisy environmental information from the local cellular micro-environment to the cell nucleus [3, 46–50]. In order to perform this task effectively, they must be able to transmit complex environmental information robustly; a failure to do this has been connected to cancer initiation and progression [51–53], as well as causing deficiencies in embryonic development [12, 54, 55].

A great amount of our knowledge about signalling networks arises from the detailed analysis of their constituent signalling pathways. A variety of these pathways have been studied in great detail, and the core components and biochemical mechanisms of signal transduction in pathways such as wingless-integrated (Wnt), transforming growth factor beta (TGF-β) and mitogen-activated protein kinase (MAP Kinase) signalling are well defined [51, 52, 54, 56, 57]. Considering the example of the Wnt signalling pathway, we can describe how signals might be transmitted in the cell nucleus. The key effector of canonical Wnt signalling in the nucleus is the multitasking protein β-catenin, which plays an important role in cell-to-cell adhesion and in gene transcription. It is well known that the absence of Wnt signalling stimulation from β-catenin leads to the repression of the target genes, while the binding of it helps the signal pass into the nucleus [51, 52, 54, 58, 59]. They negatively regulate Wnt signalling by binding to Wnt receptors. Although the knowledge of specific molecular mechanisms that contribute or respond to noise in signalling and the dynamics of these functions are well defined and studied, the role by which signalling networks receive environmental noisy signals remains unclear [21, 22, 60].

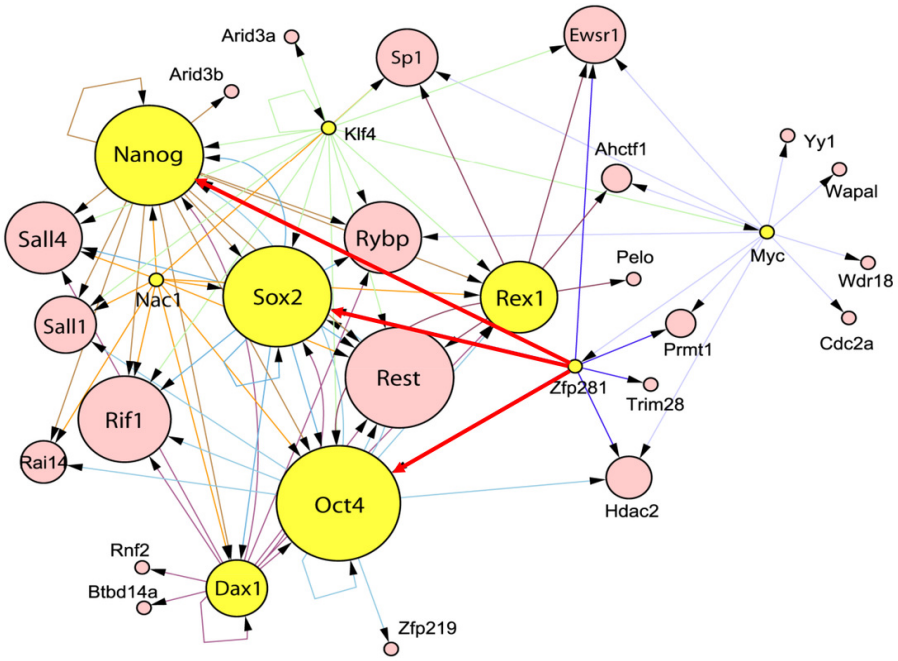


Figure 1.5: Pluripotent network of ES mouse cells showing the most important transcription factors and their interactions [6, 7]. The size of the nodes indicates the level of their connection to the other nodes. Image is taken from [7].

1.1.4 Pluripotent Network

Stem cells can be found in the pre-implantation embryo, as well as in various tissues in adults. They are distinguished for their capability in repairing tissue, self-renewal and differentiation. Embryonic stem (ES) cells are a type of mammalian stem cells which live in the inner cell mass inside the blastocyst. They have the ability to be adapted to many different functions or activities. ES cells can self-renew indefinitely *in vitro* and produce any kind of cell type and tissue [14, 36]. This property, namely pluripotency, renders ES cells distinctive [11, 34, 61–64]. The pluripotent state is characterized by a network of specific interactions between proteins and transcription factors [65, 66]. Gene expression is described by the signalling pathways of such regulatory factors.

Specific transcription factors are well known for their importance in the maintenance of the pluripotent state in ES cells. TFs such as Oct3/4, Sox2 and Nanog are the main body of the transcriptional circuit in both mouse and human embryonic stem (ES) cells [67–70]. Despite the fact that the signalling pathways in mouse and human ES cells vary, the TFs mentioned above are noticeably maintained [62, 63]. The factors Oct3/4, Sox2 and Nanog are essential for the pluripotent state and in conjunction with Myc, Tfcp2l1, Klf4, Lin28 and Esrrb are the fundamental proteins involved in the self renewal ability of ES cells [6–8, 61, 68, 70–73]. It has been established that the forced expression of combinations of these factors in somatic cells is adequate to

induce pluripotency *de novo* [61, 71, 74]. Although this core transcriptional circuit is self-sustaining when shielded from external stimulation [75], it is known that a network of signalling pathways which process extra-cellular environmental noise are also essential both to the maintenance of, and exit from, the pluripotent state [76]. Importantly, while the core transcriptional circuitry is broadly similar in mouse and human pluripotent cells [7], their dependency on external signalling is markedly different: mouse ES cells are dependent on Lif/Stat signalling [77, 78], Bmp [79], and canonical Wnt [75] to promote self-renewal, while Fgf/Erk signalling disrupts pluripotency [75, 80–82]; by contrast human ES cell self-renewal is independent of Lif [83], yet requires Activin and Fgf [84, 85] signalling and furthermore, human ES cells undergo differentiation when exposed to Bmp [85].

In order to identify interactions between proteins and transcription factors *in vivo*, various experimental methods have been developed, such as chromatin immunoprecipitation with DNA micro-array (Chip-on-chip) [7, 65, 66, 68, 86, 87]. The aim of the Chip-on-chip technique is to identify the DNA binding sites of the transcription factors, and determine the protein to DNA interactions [61, 88].

1.2 Noise and Stochasticity

Cells live in complex biochemical environments, which are inherently noisy due to chemical reactions among constitutive proteins and other regulatory molecules. Noise is the variation in the expression of these molecules that interact, along with the cell to cell variability. The way an individual protein or gene is expressed can vary because of the probabilistic nature of the cell; this can lead to the distinction of states or properties of even identical cells in the same environment [89–92]. A fundamental question to consider is '*What is the origin of this noise?*'

In biology the definition of noise is distinguished by the inherent stochasticity of the system, namely *intrinsic* noise, and by the effect of an external contribution from the environment, the *extrinsic* noise. Intrinsic noise in the cell is characterized by variability in gene expression on a molecular level, by fluctuations in signal transmission in regulatory networks, usually in transcription and translation, or by specific factors and mechanisms in the cell [36, 93–95]. Noise in gene expression is derived from the way that the molecules interact and we will describe it later in Chapter 2 by modelling biochemical reactions. As the whole process is stochastic, we can describe it mathematically as probabilistic. The natural stochasticity (intrinsic noise) can arise for multiple reasons: the small copy numbers of molecules in the cell, the different states of the cell, or genetic mutations [91, 93, 96]. However, the extrinsic noise arises from the random fluctuations in the environment of the cell (due to the noise in the expression ability of the cell), from the signalling pathways, as well as from differences in their concentration [94, 97, 98]. In large biochemical systems with many copies of

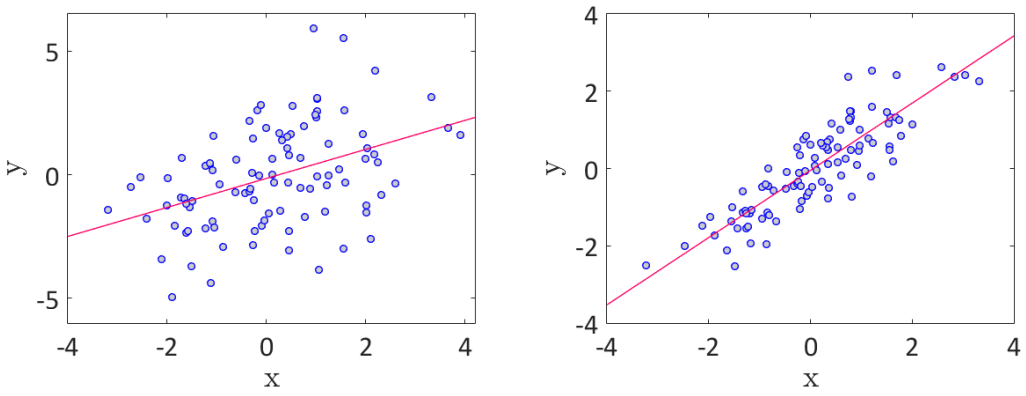


Figure 1.6: The figures show the covariance of 2 random vectors $\boldsymbol{x}, \boldsymbol{y}$ taken from a multivariate normal distribution with mean zero and covariance matrix $\boldsymbol{K} = [2, 1; 1, 4]$ (left) and $\boldsymbol{K} = [1.2, 1; 1, 1.2]$ (right). The red line shows linear regression.

molecules, the processing of the random fluctuations can be affected by the topology of the network [28, 99]. Nevertheless, noise can play a beneficial role in the cell in many levels [89]; in the coordination of gene expression for large copies of genes [100, 101], in probabilistic division and differentiation into identical cells (*i.e.* bi-stability that causes transition between cell fates) [90, 102, 103], as well as in evolution of the cell [104, 105].

One way to measure noise is to calculate the variance (for one-dimensional systems) of the molecular species of interest, which express how much the fluctuations are spread out from their expected value [94, 96, 106]. In multi-variant systems, similar information is given by the covariance matrix $\boldsymbol{K}$ [106, 107]:

$$\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{y}) = \langle (\boldsymbol{x} - \langle \boldsymbol{x} \rangle), (\boldsymbol{y} - \langle \boldsymbol{y} \rangle)^T \rangle, \tag{1.1}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are random vectors, and $\langle \boldsymbol{x} \rangle, \langle \boldsymbol{y} \rangle$ denote the mean value of $\boldsymbol{x}, \boldsymbol{y}$, respectively. The diagonal elements of $\boldsymbol{K}$ give the variances, while the off-diagonal denote the covariances of two random variables $\boldsymbol{x}, \boldsymbol{y}$, expressing a measure of their correlation strength. The sign of $\boldsymbol{K}$ denotes the linear relationship of $\boldsymbol{x}, \boldsymbol{y}$, and a zero no correlation. The normalised covariance or Pearson correlation coefficient of Eqn. (1.1) is expressed by the coefficients [107]:

$$\boldsymbol{\varrho}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{y})}{\sqrt{\langle \boldsymbol{x} \rangle^2 \langle \boldsymbol{y} \rangle^2}}, \tag{1.2}$$

where its magnitude can vary between $-1 < \boldsymbol{\varrho}(\boldsymbol{x}, \boldsymbol{y}) < 1$; the sign of $\boldsymbol{\varrho}$ stands for a positive or negative correlation and a zero for no correlation.

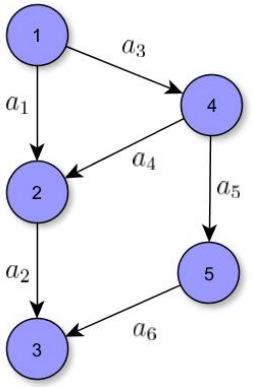
1.3 Mathematical Framework

It is possible to mathematically represent any network by its associated graph $\mathcal{G}(V, E)$, where V is the set of vertices (nodes) and $(i, j) \in E$ the set of edges that connect two nodes i and j [108, 109]. The nodes can represent genes, proteins, transcription factors or other regulatory molecules, and the edges represent the interactions between these molecules. When $\mathcal{G}$ is considered to be directed, each edge has an orientation from node ν_i to ν_j, with a sign corresponding to activation or inhibition. In the event that $\mathcal{G}$ is regarded as undirected, all interactions are bi-directional. A vertex is adjacent to another when there is an edge that connects them. Information about whether or not there is such a connection is provided by the adjacency matrix, which is defined as follows

$$\boldsymbol{A} = \begin{cases} a_{ij} \in \mathbb{R},\ i \neq j \\ a_{ij} = 0,\ i = j \end{cases} . \tag{1.3}$$

The diagonal elements of $\boldsymbol{A}$ are zero, which is related to the fact that the graph $\mathcal{G}$ is loop-less, such that no vertex is connected to itself.

As an example, the following directed graph can be expressed with its corresponding adjacency matrix $\boldsymbol{A}$ accordingly:



$$\boldsymbol{A} = \begin{pmatrix} 0 & a_1 & 0 & a_3 & 0 \\ 0 & 0 & a_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & a_4 & 0 & 0 & a_5 \\ 0 & 0 & a_6 & 0 & 0 \end{pmatrix} .$$

In the case that in a graph all the interactions are bi-directional, the adjacency matrix representation is symmetric.

In the following sections, in order to proceed further in noise modelling, we introduce basic notions and methods that will be used later in calculations.

1.3.1 Conditional Probability

Considering n states $y_1, y_2, \ldots, y_n$ that a system may exist in at times $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_n$, we define the conditional probability $p(y_n, \tau_n \mid y_{n-1}, \tau_{n-1}; y_{n-2}, \tau_{n-2}; \ldots; y_0, \tau_0)$

[107], given the previous states, as follows

$$p(y_n, \tau_n \mid y_{n-1}, \tau_{n-1}; y_{n-2}, \tau_{n-2}; \ldots; y_0, \tau_0) = \frac{p(y_n, \tau_n; y_{n-1}, \tau_{n-1}; y_{n-2}, \tau_{n-2}; \ldots; y_0, \tau_0)}{p(y_{n-1}, \tau_{n-1}; y_{n-2}, \tau_{n-2}; \ldots; y_0, \tau_0)}. \tag{1.4}$$

This probability in Eqn. (1.4) is conditioned on all previous states.

1.3.2 Markov Process

A stochastic process is said to satisfy the Markov property when it is memoryless [106, 107], which means that any conditional probability of a system being in a state y_n at time τ_n, as defined in Eqn. (1.4), in this case depends only on the previous one, and therefore satisfies

$$p(y_n, \tau_n \mid y_{n-1}, \tau_{n-1}) = \frac{p(y_n, \tau_n; y_{n-1}, \tau_{n-1}; y_{n-2}, \tau_{n-2}; \ldots; y_0, \tau_0)}{p(y_{n-1}, \tau_{n-1})}. \tag{1.5}$$

1.3.3 Wiener Process

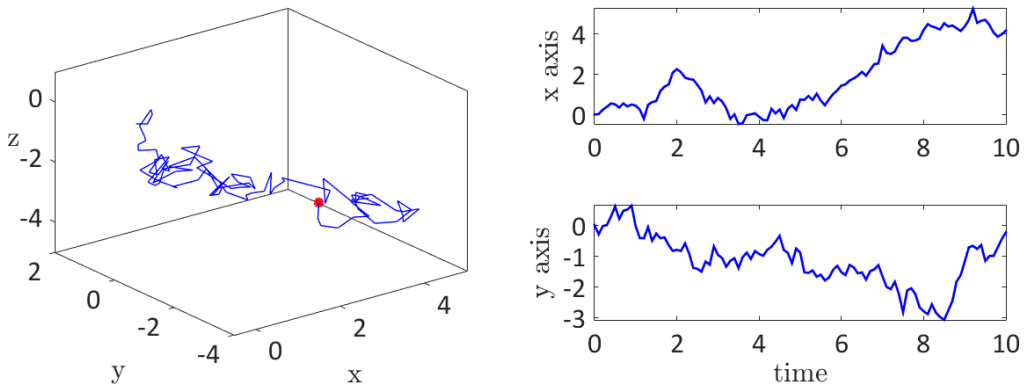


Figure 1.7: The left image shows one realisation of a 3-dimensional Brownian motion, produced by a simulation for final time $T = 1$. The right figures are plots of the BM in axis x (top), and y (bottom) against time.

A well known example of a Markov process is Brownian motion (BM) or the Wiener process [106, 107]. It is observed that when particles are submerged in fluids such as water or oxygen, they move randomly. This motion is caused by the collision of the particles against the atoms in the fluid [110, 111]. The molecules in cells behave in a similar way to these particles [36].

The Wiener process $\boldsymbol{W}(t)$ is a continuous-in-time Markov process, satisfying the property in Eqn. (1.5). The most important properties of Wiener process are highlighted as follows:

1. At time $t = 0$, the Wiener process is $\boldsymbol{W}(0) = 0$ with probability 1.

2. At the time intervals $0 \leq t_1 < t_2 < \ldots < t_{j-1} < t_j \leq T$, the increment

$$\Delta \boldsymbol{W}(t_{j-1}) = \boldsymbol{W}(t_j) - \boldsymbol{W}(t_{j-1}), \tag{1.6}$$

is normally distributed according to $\mathcal{N}(0, t_j - t_{j-1})$, with mean zero and variance $t_j - t_{j-1}$.

3. For different time intervals the increments $\boldsymbol{W}(t_i) - \boldsymbol{W}(t_{i-1})$ and $\boldsymbol{W}(t_j) - \boldsymbol{W}(t_{j-1})$ are independent for $i \neq j$.

1.3.4 Stationary Process

Consider the stochastic processes

$$\boldsymbol{x}(t) = \{x_i(t_i), x_i \in S, t_i \in T, i = 1, 2..., n\} \tag{1.7}$$

$$\boldsymbol{x}(t+\tau) = \{x_i(t_i + \tau), x_i \in S, t_i, \tau \in T, i = 1, 2..., n\}, \tag{1.8}$$

where S and T are the sample space and the index set. Then $\boldsymbol{x}$ is a stationary process when for a time shift τ the probability distribution is unique and independent of time, and hence satisfies:

$$P(x_i, t_i) = P(x_i, t_i + \tau), \text{ for } i = 1, 2, ..., n. \tag{1.9}$$

1.3.5 Ornstein-Uhlenbeck Process

An example of a well studied stochastic process is the *Ornstein-Uhlenbeck* process [106, 107, 112, 113]. It is a continuous in time Markov process with stationary probability distribution and for n-multivariate random variables, denoted by the vector $\boldsymbol{x} \in \mathbb{R}^n$, and defined by the stochastic differential equation (SDE):

$$d\boldsymbol{x}(t) = -\boldsymbol{M}\boldsymbol{x}(t)dt + \boldsymbol{\Sigma}d\boldsymbol{W}(t), \tag{1.10}$$

where matrix $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times m}$. The vector $\boldsymbol{W}(t) = [W_1, W_2, ..., W_m]$ denotes m independent standard Brownian motions. The stationary solution of Eqn. (1.10) at the time limit in infinity is a Gaussian process distributed according to $\mathcal{N}(e^{-\boldsymbol{M}t}\boldsymbol{x}_0, \boldsymbol{K}(t))$, with mean zero and covariance matrix $\boldsymbol{K}(t) \in \mathbb{R}^{n \times n}$, defined as

$$\boldsymbol{K}(t) = \langle \boldsymbol{x}(t) - \langle \boldsymbol{x}(t) \rangle, (\boldsymbol{x}(t) - \langle \boldsymbol{x}(t) \rangle)^T \rangle. \tag{1.11}$$

1.3.6 Itô Calculus

Let us consider a function $h(t)$ and the Brownian motion $\boldsymbol{W}(t)$, as defined in Section 1.3.3 [106, 114]. Then, the stochastic integral of this function $h(t)$ with respect to

Brownian motion can be approximated in a similar way with a deterministic integral, by the following sum

$$\int_0^T h(t)d\boldsymbol{W}(t) = \sum_{j=1}^{N} h(t_{j-1})\left(\boldsymbol{W}(t_j) - \boldsymbol{W}(t_{j-1})\right), \tag{1.12}$$

known as the Itô stochastic integral. The most important properties of Itô integral, as defined in Eqn. (1.12) are the following:

1. The mean of Itô integral is zero, since

$$\left\langle \int_0^T h(t)d\boldsymbol{W}(t) \right\rangle = \sum_{j=1}^{N} \langle h(t_{j-1})\rangle \left\langle \boldsymbol{W}(t_j) - \boldsymbol{W}(t_{j-1})\right\rangle = 0. \tag{1.13}$$

2. The so-called Itô's isometry is satisfied:

$$\left\langle \int_0^T h(t)d\boldsymbol{W}(t) \right\rangle^2 = \int_0^T h(t)^2 dt \tag{1.14}$$

1.3.7 Euler-Maruyama Approximation

In order to find an approximate solution and simulate an SDE, we recall now the Ornstein-Uhlenbeck process in Eqn. (1.10), which can be written as an Itô integral using Eqn. (1.12):

$$\begin{aligned} \boldsymbol{x}(t) &= \boldsymbol{x}_0 + \int_0^t \boldsymbol{M}\boldsymbol{x}(s)ds + \boldsymbol{\Sigma}\int_0^t d\boldsymbol{W}(s) && (1.15)\\ &= \boldsymbol{x}_0 e^{-\boldsymbol{M}t} + \boldsymbol{\Sigma}\int_0^t e^{-\boldsymbol{M}(t-s)} d\boldsymbol{W}(s), && (1.16) \end{aligned}$$

where $\boldsymbol{x}(0) = \boldsymbol{x}_0$ is the initial condition and $t \in [0, T]$.

We take now the discrete interval $\Delta t = T/n$ and consider $t_j = j\Delta t$. Then, we apply the Itô integral in Eqn. (1.15) for t_j and t_{j-1} and subtract to get

$$\boldsymbol{x}(t_j) - \boldsymbol{x}(t_{j-1}) = \int_{j-1}^{t_j} \boldsymbol{M}\boldsymbol{x}(s)ds + \int_{j-1}^{t_j} \boldsymbol{\Sigma} d\boldsymbol{W}(s), \tag{1.17}$$

from which we derive the Euler-Maruyama approximation

$$\boldsymbol{x}(t_j) = \boldsymbol{x}(t_{j-1}) + \boldsymbol{M}\boldsymbol{x}(t_{j-1})\Delta t + \boldsymbol{\Sigma}\left(\boldsymbol{W}(t_j) - \boldsymbol{W}(t_{j-1})\right),\ j = 1, \ldots, n. \tag{1.18}$$

In our simulations in the next chapters we use the random variable $\xi_j \sim \mathcal{N}(0, \Delta t)$, so Eqn. 1.19 can be expressed as

$$\boldsymbol{x}(t_j) = \boldsymbol{x}(t_{j-1}) + \boldsymbol{M}\boldsymbol{x}(t_{j-1})\Delta t + \boldsymbol{\Sigma}\sqrt{\Delta t}\ \xi_{j-1}. \tag{1.19}$$

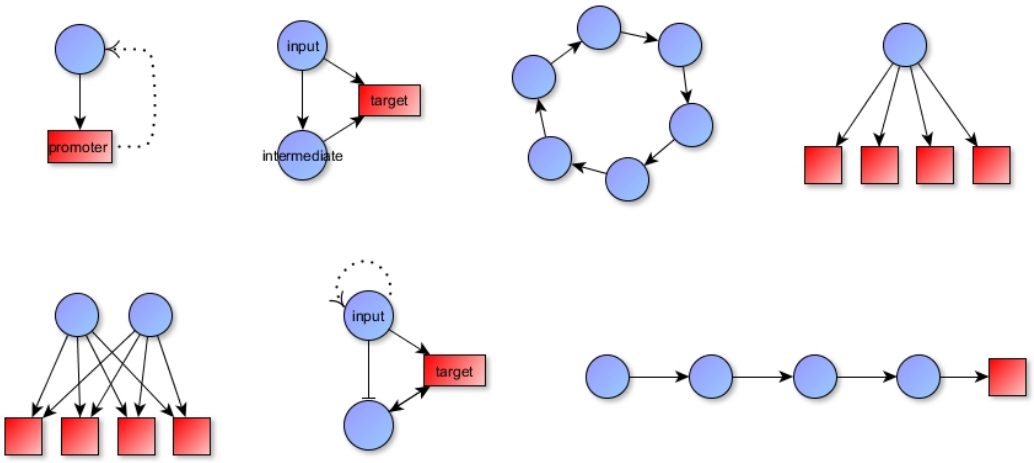


Figure 1.8: Network motifs examples of positive auto-regulation, feed-forward loop, cycle, multiple targets and inputs, feedback loop and a chain.

1.4 Motivation of the Thesis

The aim in this thesis is to try to explain how external fluctuations are mitigated by the underlying network. In particular, we will focus on how extrinsic noise is propagated through a signalling network with linear interactions, from a noisy source to a target of interest. An intriguing question is to answer how the structure of a network is related to its noise processing abilities. Because the cell is complex and consists of a huge number of molecules, it is useful to consider smaller networks, or motifs. Some of these are well studied, such as auto-regulation, positive or negative feedback loops, feed-forward loops, chains, cycles, single or multiple input networks [21, 35]. In Fig. 1.8 we show some examples of common network motifs. In this work, we discuss a non reversible chain and a feed-forward loop in order to relate the topology of the network with the noise propagation and describe the mechanism by which our expression for the variance functions.

The rest of the thesis is organized as follows. Chapter 2 explains the mathematical modelling of intrinsic fluctuations, and presents some computational algorithms. Chapter 3 is focused on the extrinsic noise of a signalling network. In particular, we consider modelling a network from another perspective, by using the adjacency matrix in order to have a connection of the structure of the network with the covariance. This noise is extrinsic and we focus on how it can be propagated through a complex network. In Section 3.1.1 we establish a mathematical formula for the covariance at the equilibrium that relates the noisy sources to any target node of particular interest. There follows a discussion in Section 3.1.2 of different ways that stability can be ensured in a biochemical system. Then, in Sections 3.2 and 3.3 we demonstrate our result with two simple examples, a chain and a feed-forward loop. In Section 3.4, we

apply the expression in a real signalling network of pluripotent cells and identify the most important interactions in terms of the noise processing. Finally, the discussion in Section 1.3 explains the underlying theory as to how the noise processing is affected by the network's architecture.

Chapter 2

Modelling Intrinsic Noise

In this chapter we will discuss different ways of modelling intrinsic noise. The molecules inside the cell are moving randomly, as described by Brownian motion. The aim here is to focus on the probability that a molecule will hit another, and therefore a chemical reaction will occur. Firstly, some basic examples are given in order to calculate the copy number of a species after the next reaction and the chemical master equation is derived in detail. Furthermore, the stationary solution is derived analytically at the equilibrium for the simple example of production-degradation of a species. This solution is represented as a Poisson distribution, followed by figures that illustrate the results. Moreover, the most important moments, expected value and variance are derived, and the use of the generating function as a tool for the higher moments is given. Furthermore, computational algorithms are described and their differences are shown in corresponding figures. In the case of large copy number, we describe approximations to simulate such systems. We derive the stochastic Langevin equation and the deterministic reaction rate equation and simulate the comparison of their solutions with the stochastic simulation algorithms. Finally, the chemical Fokker-Planck is derived from the chemical master equation, and the stationary probability is determined in the case of an one dimensional system. We conclude with a comparison between the probability density functions obtained using the chemical equation with the one obtained using the chemical Fokker-Planck.

2.1 Chemical Master Equation

In this chapter intrinsic noise is modelled by the well studied chemical master equation (CME). Consider a spatially homogeneous biochemical system of N species $S_1, S_2, ..., S_N$ that interact via M chemical reactions. This system can be modelled using a set of linear ordinary differential equations (ODEs): the chemical master equation [27–29, 106, 107, 115]. Due to the inherent random nature of molecular interactions, the noise that is emerging from the biochemical reactions is intrinsic. The CME charac-

terises the time evolution of the probability of the system having n copies of a species at time t. We will denote this probability as $p(\boldsymbol{n}, t)$. Furthermore, we consider the temperature of the system to be constant while reactions take place in a fixed volume V. Then the concentration of the S_i species is given by $X_i(t) = n_i(t)/V$, where n_i is the number of molecules of S_i at t. Let us denote with $\boldsymbol{n}(t) = [n_1(t), n_2(t), ..., n_N(t)]^T, \boldsymbol{n}(t) \in \mathbb{R}^N$, the state vector, where each $n_i(t)$ is the number of molecules of species S_i at time t. Every time a reaction takes place, $\boldsymbol{n}(t)$ changes due to the consumption or production of the interacting molecules. This change is described by the stoichiometric column vectors $\boldsymbol{\nu_j} \in \mathbb{Z}^N, j = 1, 2, ..., M$, and the set of $\nu = \{\nu_{ij}\} \in \mathbb{Z}^{N\times M}$, which forms the stoichiometric matrix. If ν_{ij} is negative then this implies a loss of molecules for the i-th species, while a positive indicates a gain, when the jth reaction occurs. When a species does not take part in the reaction, we have that $\nu_{ij} = 0$. The opposite is not always true, since a zero does not always correspond to not interacting species. For example, for the following reactions:

$$S_1 + S_2 \xrightarrow{r_1} 2S_3, \tag{2.1a}$$

$$S_3 + S_3 \xrightarrow{r_2} S_2 + 2S_3 + 3S_4, \tag{2.1b}$$

the corresponding stoichiometric vectors are $\boldsymbol{\nu_1} = [-1, -1, 2, 0]^T$ and $\boldsymbol{\nu_2} = [0, 1, 0, 3]^T$. In $\boldsymbol{\nu_2}$, the zero entry in the 3rd element indicates that the gain in the molecules of species S_3 is equal to its loss.

The value of the state vector $\boldsymbol{n}(t)$ depends on $\boldsymbol{\nu_j}$, since the former changes from $\boldsymbol{n}(t)$ to $\boldsymbol{n}(t) + \boldsymbol{\nu_j}$. Let dt be the infinitesimal time that it takes a reaction to fire with reaction rate $r \in \mathbb{R}$. The coefficients that connect the rates r to the number of molecules of the reactant species are the propensity functions $\alpha_j(\boldsymbol{n}(t))$. The probability that the j-th reaction will occur in the next dt is given by $\alpha_j(\boldsymbol{n}(t))dt$, namely the transition probability [106]. Depending on the reaction rates and the molecules that interact, it can take various forms. The propensity function $\alpha_j(\boldsymbol{n}(t))$ of a reaction is the product of the constant rate r and all possible combinations of the number of molecules consumed. In the above example, the propensity function in the reaction Eqn. (2.1a) takes the form

$$\alpha_1(\boldsymbol{n}(t)) = r_1 n_1(t) n_2(t), \tag{2.2}$$

while in Eqn. (2.1b) it is given by

$$\alpha_2(\boldsymbol{n}(t)) = r_2 \frac{n_3(t)(n_3(t) - 1)}{2} = r_2 \binom{n_3(t)}{2}. \tag{2.3}$$

Derivation of chemical master equation Let us assume that the event A that characterises the system is in a state with n copies at time $t + dt$. In addition let us assume that $H_j, j = 1, ..., M$ denote the event that only the j-th reaction occurs in the interval $[t, t + dt]$, while in addition H_{M+1} is the event that two or more reactions happen in this interval. The aim is to obtain the transition probability $p(\boldsymbol{n}, t + dt)$,

which satisfies the Markov property (1.5) [107]. Consider the null event H_0 so that at time t the system is at the state with n copies. Then two events are possible: no reactions fired with probability

$$\mathbb{P}(A \mid H_0) = 1 - \sum_{j=1}^{M} \alpha_j(\boldsymbol{n})dt, \tag{2.4}$$

or only the j-th reaction fired with probability:

$$\mathbb{P}(A \mid H_j) = \alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j})dt. \tag{2.5}$$

We also find the following probability, which more than one reaction occurred

$$\mathbb{P}(A \mid H_{M+1}) = 0, \tag{2.6}$$

since only one reaction can occur in $[t, t + dt)$. Combining Eqns. (1.5), (2.5), (2.4) and the law of total probability [28, 106]:

$$\mathbb{P}(A) = \sum_{j=0}^{M+1} \mathbb{P}(A \mid H_j)\mathbb{P}(H_j), \tag{2.7}$$

we obtain the transition probability $p(\boldsymbol{n}, t + dt)$ as follows

$$p(\boldsymbol{n}, t + dt) = \left(1 - \sum_{j=1}^{M} \alpha_j(\boldsymbol{n})dt\right) \times p(\boldsymbol{n}, t) + \sum_{j=1}^{M} \alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j})dt \times p(\boldsymbol{n} - \boldsymbol{\nu_j}, t) \tag{2.8}$$

$$= p(\boldsymbol{n}, t) - \sum_{j=1}^{M} \alpha_j(\boldsymbol{n})dt \times p(\boldsymbol{n}, t) + \sum_{j=1}^{M} \alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j})dt \times p(\boldsymbol{n} - \boldsymbol{\nu_j}, t) \tag{2.9}$$

After simple rearrangements in order to obtain the derivative of the probability $p(\boldsymbol{n}, t)$, the limit of dt to zero gives the general CME in higher dimensions (N):

$$\frac{dp(\boldsymbol{n}, t)}{dt} = \sum_{j=1}^{M} \left(\alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j})p(\boldsymbol{n} - \boldsymbol{\nu_j}, t) - \alpha_j(\boldsymbol{n})p(\boldsymbol{n}, t)\right). \tag{2.10}$$

The CME can be solved given the initial condition as the *Kronecker* delta:

$$p(\boldsymbol{n_0}, 0) = \begin{cases} 1, & \text{if } \boldsymbol{n} = \boldsymbol{n_0}, \\ 0, & \text{if } \boldsymbol{n} \neq \boldsymbol{n_0}. \end{cases} \tag{2.11}$$

Example: We now describe the CME for a specific chemical system, where two species S_1 and S_2 interact with a constant production rate r_+ to produce S_3. Moreover, the reaction goes reversibly with decay rate r_-. The system is described by the following chemical reactions:

$$S_1 + S_2 \underset{r_-}{\overset{r_+}{\rightleftharpoons}} S_3. \tag{2.12}$$

According to the notation introduced, $N = 3$ is the number of the species, taking part in $M = 2$ reactions and the state vector is formed by $\boldsymbol{n}(t) = [n_1(t), n_2(t), n_3(t)]^T$. The corresponding stoichiometric vectors for the system (2.12) are $\boldsymbol{\nu_1} = [-1, -1, 1]^T$ and $\boldsymbol{\nu_2} = [1, 1, -1]^T$. The propensity function for the forward reaction is given by $\alpha_1(\boldsymbol{n}(t)) = r_+ n_1(t) n_2(t)$ and for the backward $\alpha_2(\boldsymbol{n}(t)) = r_- n_3(t)$. The non-negative integers m, n, and l represent the copy numbers that are present at time t, and $p(m, n, l, t)$ the probability that $n_1(t) = m$, $n_2(t) = k$ and $n_3(t) = l$ at t. Applying the formulae (2.10) and (2.12), the time evolution of Eqn. (2.12) can be described as follows:

$$\begin{aligned}
\frac{dp(m,k,l,t)}{dt} &= \alpha_1(m+1,k+1)\; p(m+1,k+1,l-1,t) - \alpha_1(m,k)\; p(m,k,l,t) \\
&\quad + \alpha_2(l+1)\; p(m-1,k-1,l+1,t) - \alpha_2(l)\; p(m,k,l,t) \\
&= r_+(m+1)(k+1)\; p(m+1,k+1,l-1,t) - r_+ mk\; p(m,k,l,t) \\
&\quad + r_-(l+1)\; p(m-1,k-1,l+1,t) - r_- l\; p(m,k,l,t) \\
&= r_+(m+1)(k+1)\; p(m+1,k+1,l-1,t) \\
&\quad + r_-(l+1)\; p(m-1,k-1,l+1,t) - (r_+ mk + \; r_- l)\; p(m,k,l,t).
\end{aligned} \tag{2.13}$$

The CME derived in Eqn. (2.13) for the example (2.12) has a second order reaction, and hence can be difficult to solve it analytically. Therefore, ones needs to resort in numerical simulations to study its properties. This is covered in the next section, where an exact stochastic simulation algorithm is discussed.

Example Stationary solution: We next consider the production-degradation example that has only first order reactions

$$S \underset{r_2}{\overset{r_1}{\rightleftharpoons}} \emptyset. \tag{2.14}$$

Assuming that we are only interested in S with copy number n, the stoichiometric quantities are $\nu_1 = -1$ and $\nu_2 = 1$, with propensities $\alpha_1(n) = r_1 n$ and $\alpha_2 = r_2$. The probability that at time t we have n molecules in the system is given by:

$$\begin{aligned}
\frac{dp(n,t)}{dt} &= \sum_{j=1}^{2} \alpha_j(n-\nu_j) p(n-\nu_j, t) - \sum_{j=1}^{2} \alpha_j(n) p(n,t) \\
&= \alpha_1(n+1) p(n+1,t) - \alpha_1(n) p(n,t) + \alpha_2 p(n-1,t) - \alpha_2 p(n,t) \\
&= r_1(n+1) p(n+1,t) + r_2 p(n-1,t) - (r_1 n + r_2) p(n,t).
\end{aligned} \tag{2.15}$$

In order to find the stationary distribution we take the limit t at the infinity as

$$p^*(n) = \lim_{t\to\infty} p(n,t). \tag{2.16}$$

Here, $p^*(n)$ can be calculated analytically by solving Eqn. (2.15) for stability, which gives

$$r_1(n+1)p^*(n+1) - (r_1 n + r_2)p^*(n) + r_2 p^*(n-1) = 0. \tag{2.17}$$

Now setting the initial condition for having zero copies of a species and assuming that $p^*(-1) \equiv 0$, we have

$$p^*(1) = \frac{r_2}{r_1} p^*(0), \tag{2.18}$$

and for any $n \geq 1$

$$p^*(n+1) = \frac{r_1 n + r_2}{r_1(n+1)} p^*(n) - \frac{r_2}{r_1(n+1)} p^*(n-1). \tag{2.19}$$

Then by setting $p^*(0) \equiv 1$ in Eqn. (2.18) for convenience and by calculating Eqn. (2.19) for $n \geq 1$ we solve the recurrence relation

$$p^*(n) = c\, \frac{1}{n!} \left(\frac{r_2}{r_1}\right)^n, \tag{2.20}$$

where c is a constant. By considering that all probabilities sum to 1, *i.e.*

$$\sum_{n=0}^{\infty} p^*(n) = 1, \tag{2.21}$$

we can calculate $c = e^{-r_2/r_1}$ and derive the stationary distribution of Eqn. (2.15)

$$p^*(n) = \frac{1}{n!} \left(\frac{r_2}{r_1}\right)^n e^{-\frac{r_2}{r_1}}. \tag{2.22}$$

Hence, in the limit of $t \to \infty$, $n \sim \text{Pois}(r_2/r_1)$, where $\text{Pois}\,(\lambda)$ denotes the Poisson distribution with probability mass function $\text{Pois}\,(\lambda) = \lambda^n e^{-\lambda}/n!$. (2.23)

Expected value: The expected value of a species S_i is given by [116]:

$$\langle n_i(t) \rangle = \sum_{k=0}^{\infty} k p(n_i = k, t). \tag{2.24}$$

In order to derive an evolution equation for the mean value of each species, we multiply CME in Eqn. (2.10) with n_i:

$$\frac{d}{dt} n_i p(\boldsymbol{n}, t) = \sum_{j=1}^{M} n_i \alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j}) p(\boldsymbol{n} - \boldsymbol{\nu_j}, t) - \sum_{j=1}^{M} n_i \alpha_j(\boldsymbol{n}) p(\boldsymbol{n}, t). \tag{2.25}$$

Then, by taking the infinite sum of the molecules of a species x_i we write:

$$\frac{d}{dt} \sum_{n_i=0}^{\infty} n_i p(\boldsymbol{n}, t) = \sum_{j=1}^{M} \sum_{n_i=0}^{\infty} n_i \alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j}) p(\boldsymbol{n} - \boldsymbol{\nu_j}, t) - \sum_{j=1}^{M} \sum_{n_i=0}^{\infty} n_i \alpha_j(\boldsymbol{n}) p(\boldsymbol{n}, t). \tag{2.26}$$

Rearranging, we have

$$\boldsymbol{y} = \boldsymbol{n} - \boldsymbol{\nu_j} \Rightarrow \boldsymbol{n} = \boldsymbol{y} + \boldsymbol{\nu_j}, \tag{2.27}$$

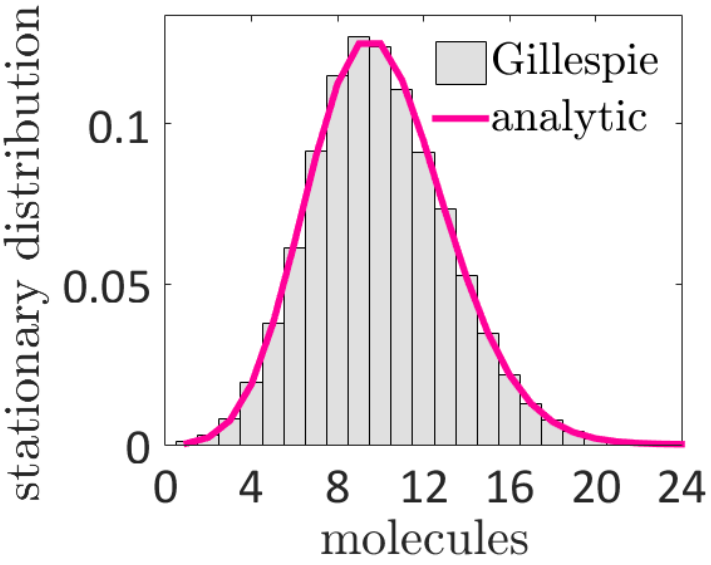


Figure 2.1: The figure shows a comparison of the stationary distribution with the Gillespie algorithm. The pink line is the exact stationary solution derived in Eqns. (2.18) and (2.19). The grey area is the implementation of the Gillespie algorithm for 10^6 realisations, for time $T = 1$. The rates used are $r_1 = 0.004$ and $r_2 = 0.04$, and initial value $n_0 = 0$ molecules.

and hence for every n_i:

$$n_i = y_i + \nu_{ij}. \tag{2.28}$$

Then Eqn. (2.26) takes the form:

$$\begin{aligned} \frac{d}{dt}\sum_{n_i=0}^{\infty} n_i p(\boldsymbol{n},t) &= \sum_{j=1}^{M}\sum_{y_i=0}^{\infty}(y_i+\nu_{ij})\alpha_j(\boldsymbol{y})p(\boldsymbol{y},t) - \sum_{j=1}^{M}\sum_{n_i=0}^{\infty} n_i\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) \\ &= \sum_{j=1}^{M}\sum_{y_i=0}^{\infty}\nu_{ij}\alpha_j(\boldsymbol{y})p(\boldsymbol{y},t). \end{aligned} \tag{2.29}$$

Now summing over the remaining sum, we get

$$\frac{d}{dt}\langle n_i(t)\rangle = \sum_{j=1}^{M}\nu_{ij}\langle\alpha_j(\boldsymbol{n}(t))\rangle, \tag{2.30}$$

and we obtain the evolution equation for the mean of the state vector $\boldsymbol{n}$:

$$\frac{d}{dt}\langle\boldsymbol{n}(t)\rangle = \sum_{j=1}^{M}\boldsymbol{\nu_j}\langle\alpha_j(\boldsymbol{n}(t))\rangle = \nu\langle\alpha(\boldsymbol{n}(t))\rangle. \tag{2.31}$$

Second moment: The evolution equation for the second moment can be calculated in a similar manner [107, 117]. Now, we apply the mean equation from Eqn. (2.24) to $\boldsymbol{n}\boldsymbol{n}^T$

$$\langle \boldsymbol{n}(t)\boldsymbol{n}(t)^T\rangle_{ij} = \sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} n_i n_j p(\boldsymbol{n},t), \tag{2.32}$$

and use the CME in Eqn. (2.10) to get:

$$\begin{aligned}\frac{d}{dt}\langle \boldsymbol{n}\boldsymbol{n}^T\rangle_{ij} &= \frac{d}{dt}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} n_i n_j p(\boldsymbol{n},t) &(2.33)\\ &= \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} n_i n_j \alpha_k(\boldsymbol{n}-\boldsymbol{\nu_k})p(\boldsymbol{n}-\boldsymbol{\nu_k},t)\\ &- \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} n_i n_j \alpha_k(\boldsymbol{n})p(\boldsymbol{n},t) &(2.34)\end{aligned}$$

Changing the variables now by using Eqns. (2.27) and (2.28), we get:

$$\begin{aligned}\frac{d}{dt}\langle \boldsymbol{n}\boldsymbol{n}^T\rangle_{ij} &= \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} (y_i+\nu_{ik})(y_j+\nu_{jk})\alpha_k(\boldsymbol{y})p(\boldsymbol{y},t)\\ &- \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} n_i n_j \alpha_k(\boldsymbol{n})p(\boldsymbol{n},t) &(2.35)\\ &= \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} y_i\nu_{jk}\alpha_k(\boldsymbol{y})p(\boldsymbol{y},t) + \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} \nu_{ik}y_j\alpha_k(\boldsymbol{y})p(\boldsymbol{y},t)\\ &+ \sum_{k=1}^{M}\sum_{n_i=0}^{\infty}\sum_{n_j=0}^{\infty} \nu_{ik}\nu_{jk}\alpha_k(\boldsymbol{y})p(\boldsymbol{y},t). &(2.36)\end{aligned}$$

Therefore, Eqn. (2.36) can be written as follows:

$$\frac{d}{dt}\langle \boldsymbol{n}\boldsymbol{n}^T\rangle_{ij} = \langle \nu\alpha(\boldsymbol{n})\boldsymbol{n}^T\rangle_{ij} + \langle \boldsymbol{n}(\nu\alpha(\boldsymbol{n}))^T\rangle_{ij} + \langle \nu\,\mathrm{diag}(\alpha(\boldsymbol{n}))\nu^T\rangle_{ij}. \tag{2.37}$$

Finally, Eqn. (2.38) can be written in the form [117]

$$\frac{d}{dt}\langle \boldsymbol{n}\boldsymbol{n}^T\rangle_{ij} = \langle h(\boldsymbol{n})\boldsymbol{n}^T\rangle_{ij} + \langle \boldsymbol{n}h(\boldsymbol{n})^T\rangle_{ij} + \langle B(\boldsymbol{n})\rangle_{ij}, \tag{2.38}$$

where, $h(\boldsymbol{n}) = \nu\alpha(\boldsymbol{n})$ and the difussion matrix $B(\boldsymbol{n}) = \nu\,\mathrm{diag}(\alpha(\boldsymbol{n}))\nu^T$.

Generating function: In general, we cannot always obtain analytical solution for the probability distribution with the CME. However, the generating function give us this benefit, and plays an important role in directly estimating the higher moments by using its derivatives [106, 107, 118]. In one dimension it is defined as follows:

$$G(s,t) = \sum_{x=0}^{\infty} s^x p(x,t). \tag{2.39}$$

Considering the random variables driven from a *Poisson* distribution $P(x,\lambda) = \frac{\lambda^x}{x!}e^{-\lambda}$, the generating function is

$$G(s,t) = \sum_{x=0}^{\infty} \frac{(\lambda s)^x e^{-\lambda}}{x!}.$$

The higher moments can be derived by differentiating with respect to s and setting $s = 1$ in Eqn. (2.39)

$$\left[\left(\frac{d}{ds}\right)^n G(s,t)\right]_{s=1} = \left[\left(\frac{d}{ds}\right)^n \sum_{x=0}^{\infty} s^x p(x,t)\right]_{s=1} = \langle x^n \rangle_F, \tag{2.40}$$

where the factorial moment is

$$\langle x^n \rangle_F = \langle x(x-1)...(x-n)(x-n+1) \rangle.$$

The normalization condition is also applied

$$G(1,t) = \sum_{x=0}^{\infty} p(x,t) = 1. \tag{2.41}$$

The probability distribution can also be expressed by the generating function

$$\varphi(n) = \frac{1}{n!}\left[\left(\frac{d}{ds}\right)^n G(s,t)\right]_{s=0}. \tag{2.42}$$

Applying Eqn. (2.40) on the production-degradation example in Eqn. (2.15), we can derive the expected value

$$\langle n \rangle = \left[\frac{dG(s,t)}{ds}\right]_{s=1} = \sum_{n=0}^{\infty} n p_n, \tag{2.43}$$

with evolution equation

$$\frac{d\langle n \rangle}{dt} = \sum_{n=0}^{\infty} (-r_1 n + r_2) p_n = -r_1 \langle n \rangle + r_2, \tag{2.44}$$

where the solution of the mean is

$$\langle n \rangle = \frac{r_2}{r_1}(1 - e^{-r_1 t}). \tag{2.45}$$

The corresponding equations for the variance are:

$$\begin{aligned} \sigma^2 &= \langle n^2 \rangle - \langle n \rangle^2 \\ &= \sum_{n=0}^{\infty} n^2 p_n - \left(\sum_{n=0}^{\infty} n p_n\right)^2 \\ &= \left[\frac{d^2 G(s,t)}{ds^2}\right]_{s=1} + \left[\frac{dG(s,t)}{ds}\right]_{s=1} - \left[\frac{dG(s,t)}{ds}\right]^2_{s=1}, \end{aligned} \tag{2.46}$$

and its evolution equation

$$\frac{d\sigma^2}{dt} = -2r_1\sigma^2 + r_1\langle n \rangle + r_2. \tag{2.47}$$

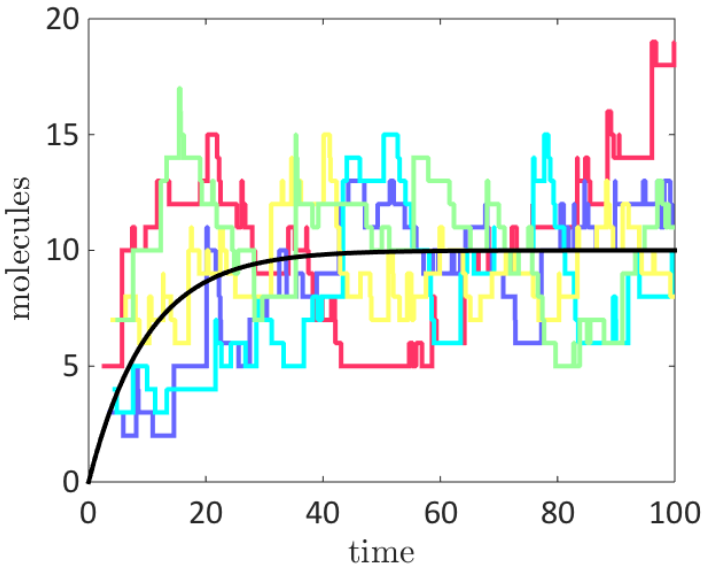


Figure 2.2: The figure illustrates the expected value in comparison with SSA, for the production - degradation example in Eqn. (2.14). Each of the colours represent one the five different realisations, simulated by Gillespie's algorithm. The dark line is the mean equation derived in Eqn. (2.45). The parameters used are similar to the Fig. 2.23, $r_1 = 0.004$ and $r_2 = 0.04$, with initial copy number zero.

2.2 Stochastic Simulation Algorithms

In order to simulate the dynamics of a continuous time Markov chain system described by the CME, in this section we study and compare two methods: the discrete Gillespie algorithm and the τ-leaping method.

2.2.1 Gillespie's Algorithm

Gillespie's algorithm or the stochastic simulation algorithm (SSA) [27–29, 117–120], is a discrete algorithm that simulates the exact solution of the CME by taking into account every reaction that takes place in the underlying chemical system. Instead of solving the CME, *Gillespie's* algorithm calculates realisations of the state vector in such a way that it agrees with the probability distribution from the CME solution. The SSA can simulate the stationary distribution $p_s(n_1, ..., n_N) = \lim_{t\to\infty} p(n_1, ..., n_N, t)$ as time approaches infinity. The SSA algorithm is summarised in the following box.

Gillespie's algorithm

1. Calculate the propensity functions $\alpha_j = \alpha_j(\boldsymbol{n}(t))$ for $i = 1, 2, ..., M$ and

$$\alpha_s = \sum_{i=1}^{M} \alpha_j. \tag{2.48}$$

2. Generate two independent random numbers ξ_1, ξ_2, uniformly distributed between 0 and 1.
3. The time for the next reaction is $t + \tau$. Generate a random variable τ exponentially distributed, with probability density $\alpha_s e^{-\alpha_s \tau}$, such that

$$\tau = {\alpha_s}^{-1} \ln\left(\frac{1}{\xi_1}\right). \tag{2.49}$$

4. In order to choose the next j^{th} reaction with probability $\alpha_j \alpha_s^{-1}$, select an integer $\mu \in [1, M]$ for the random variable ξ_2, which satisfies

$$\sum_{i=1}^{\mu-1} \alpha_j < \xi_2 \alpha_s \leq \sum_{i=1}^{\mu} \alpha_j. \tag{2.50}$$

5. Then substitute the next time $t = t + \tau$ so that $\boldsymbol{n}(t + \tau) = \boldsymbol{n}(t) + \boldsymbol{\nu_j}$, and iterate the algorithm by calculating the new propensity functions up to the final time.

In Fig. 2.3 a simulation of Gillespie algorithm is illustrated for the example in Eqn. (2.12) for the initial conditions $X_1 = 200, X_2 = 150$, and $X_3 = 0$, and rate constants $r_1 = 4 \times 10^{-4}, r_2 = 2 \times 10^{-2}$.

2.2.2 τ-leaping Method

The *Gillespie* algorithm is exact as it draws samples from the CME in an unbiased way. However, it can be slow, since the propensity functions are updated for every new time $t + \tau$. For large copy numbers this could lead to high computational cost, since the updated time random variable τ can be very small.

An alternative providing a faster way to simulate the CME is the τ-leaping method [28, 116, 120]. Here, for every subinterval $[t, t + d\tau]$ we approximate how many reactions are fired until the next one, by choosing a random variable $P(\alpha_j(\boldsymbol{n}(t), \tau)$ from Poisson distribution, with mean $\lambda = \tau \alpha_j(\boldsymbol{n}(t))$. In other words, we are treating *tau*-leaping as continuous in the sense that it counts the number of reactions that have been fired of each type M by jumping from one subinterval $[t, t + \tau]$ to the next, and then

updating the state vector $\boldsymbol{\nu_j}$. Then the algorithm approximates the random variable by $P_j\left(\alpha_j(\boldsymbol{n}(t),\tau)\right), j=1,\ldots,M$, which represents the number of reactions that have fired at that time, as shown in Eqn. (2.51), by calculating their mean $\alpha_j(\boldsymbol{n}(t))\tau$ and then updates all the reactions simultaneously as shown in Eqn. (2.52). In this way, the main idea is to select a small enough τ, so that $\boldsymbol{n}(t+\tau)$ change is infinitesimally small to maintain the propensity functions. The main points of τ-leaping algorithm are summarised in the following box.

τ-leaping algorithm

1. Choose random variables from the Poisson distribution for each reaction j as:

$$\left\{P\left(\alpha_j(\boldsymbol{n}(t),\tau)\right)\right\}_{j=1}^{M}. \tag{2.51}$$

2. For the next time $t=t+\tau$ one obtains

$$\boldsymbol{n}(t+\tau)=\boldsymbol{n}(t)+\sum_{j=1}^{M}\boldsymbol{\nu_j}P(\alpha_j(\boldsymbol{n}(t),\tau)). \tag{2.52}$$

The algorithm then iterates from the first step or stops.

An important difference between these two algorithms is that the SSA updates the system for the next reaction for every subinterval, while in τ-leaping all reactions at the same time. Moreover, the questions for the SSA are the following:

(1) When does the next reaction fire?

(2) Which is the next reaction that fires?

However, in the case of τ-leaping there is only one:

(1) In the interval $[t,t+\tau]$ how many reactions are fired?

In order to compare Gillespie algorithm and τ-leaping method, we calculate the cpu time in each simulation for the top left Figs. 2.3 and 2.4. We can conclude that by using the τ-leaping algorithm with a chosen $\tau=0.04$ the computational cost is reduced from 1.211 sec that calculated in Gillespie, to 0.150 sec in cpu time.

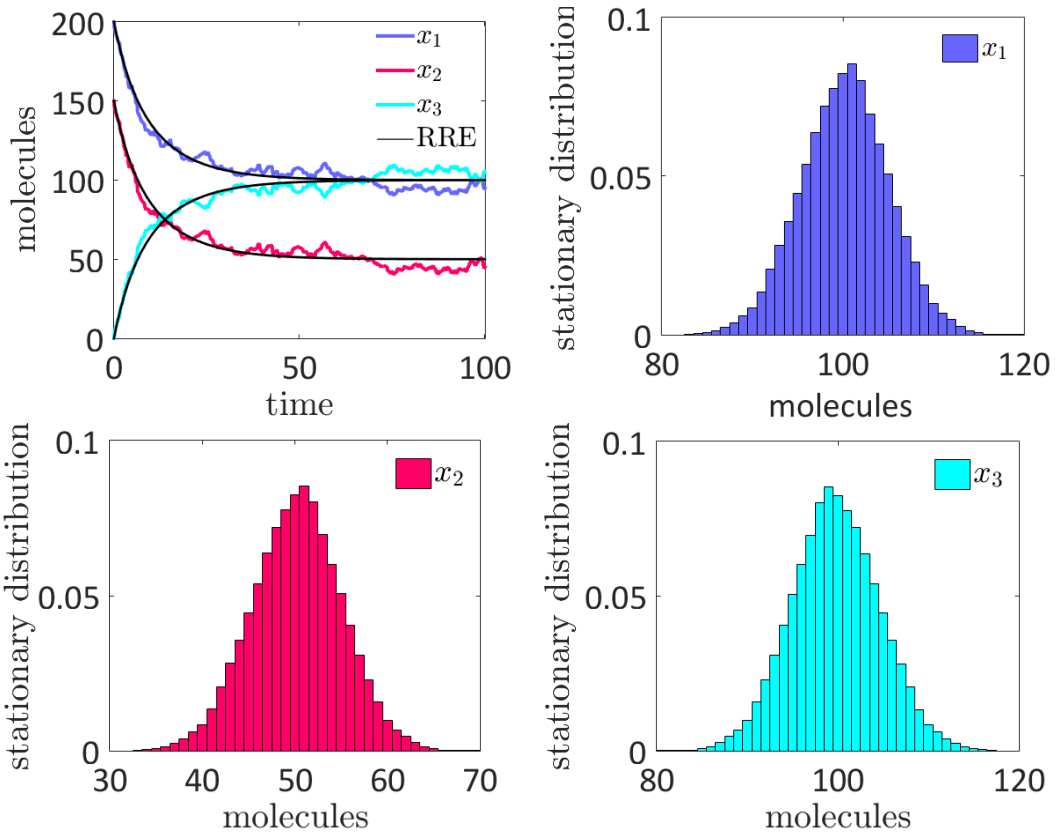


Figure 2.3: The top left figure is a simulation using the *Gillespie's* algorithm for the chemical system (2.12). The parameters used for the simulation are $r_+ = 4\times10^{-4}, r_- = 2\times10^{-2}$, for the initial copy numbers $x_1 = 200, x_2 = 150$ and $x_3 = 0$ for final time 100. The histograms are the stationary distributions for each species for 10^6 realisations for time $T = 50$.

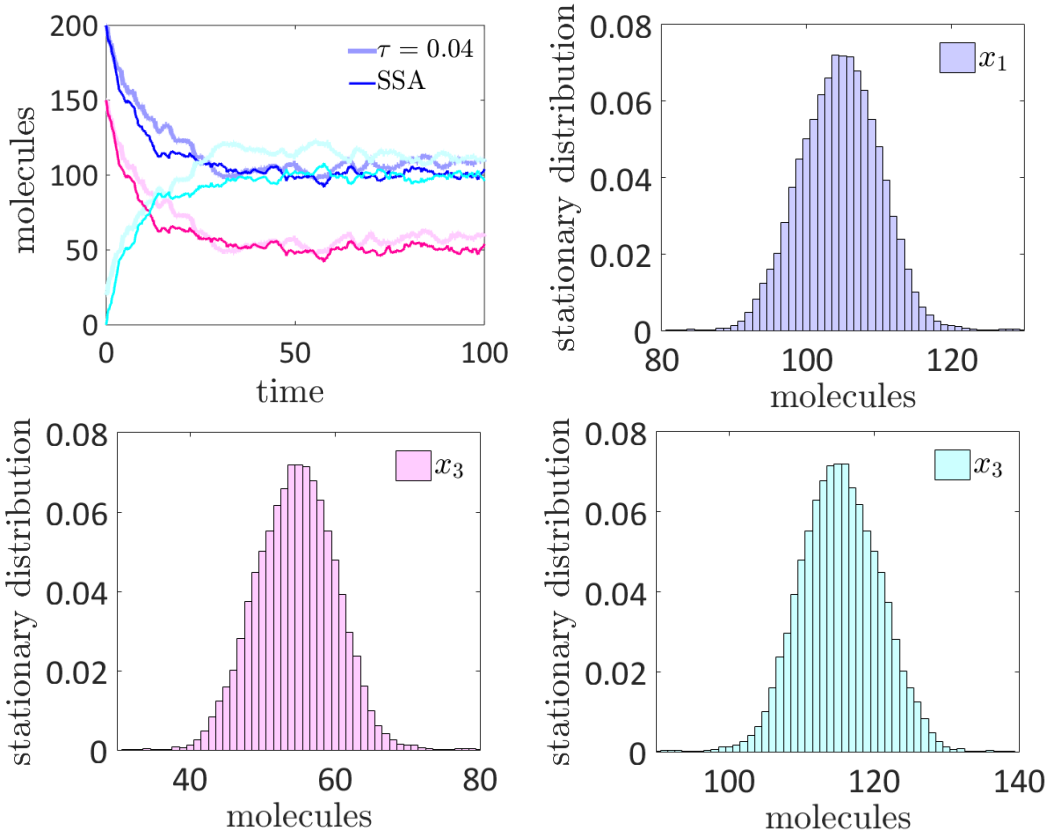


Figure 2.4: The top left figure is a comparison of the Gillespie and the τ-leaping algorithm for the chemical system (2.12). For the τ-leaping simulation is used $\tau = 0.04$. The parameters used for the simulation are $r_+ = 4 \times 10^{-4}, r_- = 2 \times 10^{-2}$, for the initial copy numbers $x_1 = 200, x_2 = 150$ and $x_3 = 0$ for final time 100. The histograms are the stationary distributions for each species for 10^6 realisations for time $T = 50$.

2.3 Chemical Langevin Equation

One way to simulate a stochastic process of chemical reactions with large copy numbers is to consider a continuous process. The chemical Langevin equation (CLE) [28, 106, 107, 121, 122], based on the τ-leaping method, is a system of non-linear stochastic differential equations (SDEs) that describes how the system evolves in time. Its dimension is determined by the number of species involved in the reactions. The solution of each one of the SDEs is associated with the population of a species after time τ. Let τ be infinitesimally small so that the propensity functions will be approximately the same, *i.e.*

$$\alpha_j(t+\tau) \approx \alpha_j(t). \tag{2.53}$$

The random variable $(\alpha_j(\boldsymbol{n}(t)), \tau)$ that represents the number of times a reaction fires after τ has a Poisson distribution Eqn. (2.23) with mean $\tau\alpha_j(\boldsymbol{n}(t))$. According to Eqn. (2.52), the random variable x_i at the next time step τ will be

$$\boldsymbol{n}(t+\tau) = \boldsymbol{n}(t) + \sum_{j=1}^{M} \boldsymbol{\nu_j} P(\alpha_j(\boldsymbol{n}(t), \tau)). \tag{2.54}$$

Although we require τ to be small enough for no changes to occur in the propensities, it is suggested that the approximation is better under the condition that $\alpha_j(\boldsymbol{n}(t))\tau$ can take a large enough value [28, 121, 122]. This assumption allows each random variable driven by the *Poisson* distribution $P(\alpha_j(\boldsymbol{n}(t), \tau))$ to be expressed with a normal random variable $\zeta_j \sim N(0, 1)$, leading to

$$\alpha_j(\boldsymbol{n}(t))\tau + \sqrt{(\alpha_j(\boldsymbol{n}(t))\tau)}\ \zeta_j. \tag{2.55}$$

Then, it follows that Eqn. (2.52) takes the form

$$\boldsymbol{n}(t+\tau) = \boldsymbol{n}(t) + \sum_{j=1}^{M} \boldsymbol{\nu_j} \alpha_j(\boldsymbol{n}(t))\tau + \sum_{j=1}^{M} \boldsymbol{\nu_j} \sqrt{(\alpha_j(\boldsymbol{n}(t))\tau)}\ \zeta_j. \tag{2.56}$$

After rearranging and take the derivative limit of x for $\tau \to 0$, we find

$$d\boldsymbol{n}(t) = \sum_{j=1}^{M} \boldsymbol{\nu_j} \alpha_j(\boldsymbol{n}(t))d\tau + \sum_{j=1}^{M} \boldsymbol{\nu_j} \sqrt{(\alpha_j(\boldsymbol{n}(t))}\ dW_j(t), \tag{2.57}$$

where $d\tau$ is the time step and $W_j(t)$ are independent *Brownian* motions. The formula (2.57) is the *Euler-Maruyama* approximation in *Ito* representation.

Alternatively, chemical *Langevin* equation can be obtained directly from the *Fokker-Planck* Eqn. (2.68) that we discuss in the Section 2.5. Setting the propensity quantities with appropriate matrices

$$\boldsymbol{\nu_j}\alpha_j(\boldsymbol{n}(t)) \equiv A_j(\boldsymbol{n}, t), \tag{2.58}$$

$$\boldsymbol{\nu_j}\sqrt{\alpha_j(\boldsymbol{n}(t))} \equiv \sqrt{B_j(\boldsymbol{n}, t)}, \tag{2.59}$$

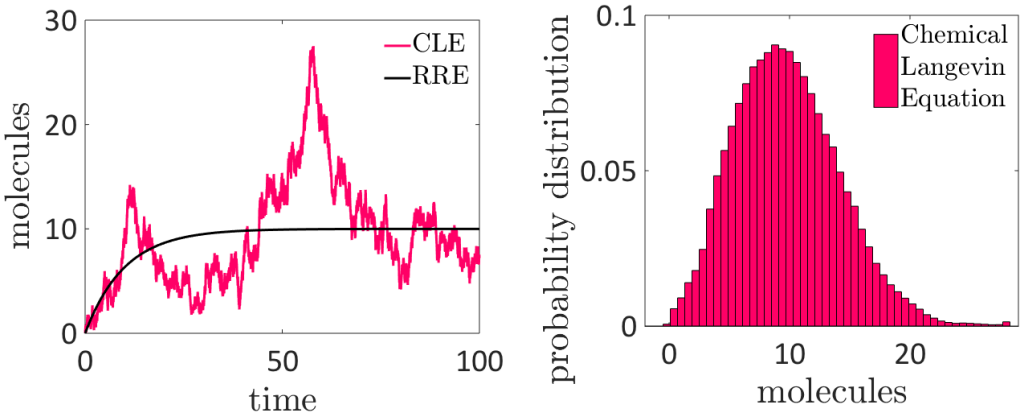


Figure 2.5: The left figure is a simulation of the chemical Langevin equation compared with the reaction rate equation. We use the Euler-Maruyama approximation for 10^6 realisations, with $dt = 0.01$. Parameters used $r_1 = 0.1$ and $r_2 = 1$ and initial condition $n_0 = 0$. The histogram is the probability distribution for a long simulation of 10^6 realisations, for fixed time $T = 10$.

Eqn. (2.57) can be expressed as

$$d\boldsymbol{n}(t) = \sum_{j=1}^{M} A_j(\boldsymbol{n}, t)d\tau + \sum_{j=1}^{M} \sqrt{B_j(\boldsymbol{n}, t)}\; dW_j(t). \tag{2.60}$$

Applying CLE in Eqn. (2.60) we can write the coresponding SDE for the production-degradation example [123] of Eqn. (2.14) as follows

$$dn(t) = (r_2 - r_1 n(t))\, d\tau + \sqrt{r_2}\; dW_1(t) - \sqrt{r_1 n(t)}\; dW_2(t), \tag{2.61}$$

where $dW_1(t), dW_2(t)$ are two independent Brownian motions. In order to simulate the CLE we use Euler-Maruyama method and a simulation is illustrated in Figs. 2.5 and 2.6.

2.4 Reaction Rate Equations

The reaction rate equation (RRE) is a set of ODEs, each one for a molecular species that can be derived by setting the stochastic part in the Langevin equation (2.57) to zero.

$$d\boldsymbol{n}(\tau) = \sum_{j=1}^{M} \boldsymbol{\nu_j} \alpha_j(\boldsymbol{n}(t))d\tau. \tag{2.62}$$

Applying the reaction rate equations in production-degradation example [123] described in Eqn. (2.14) the RRE from Eqn. (2.62) takes the form

$$dn(\tau) = (r_2 - r_1 n(t))\, d\tau. \tag{2.63}$$

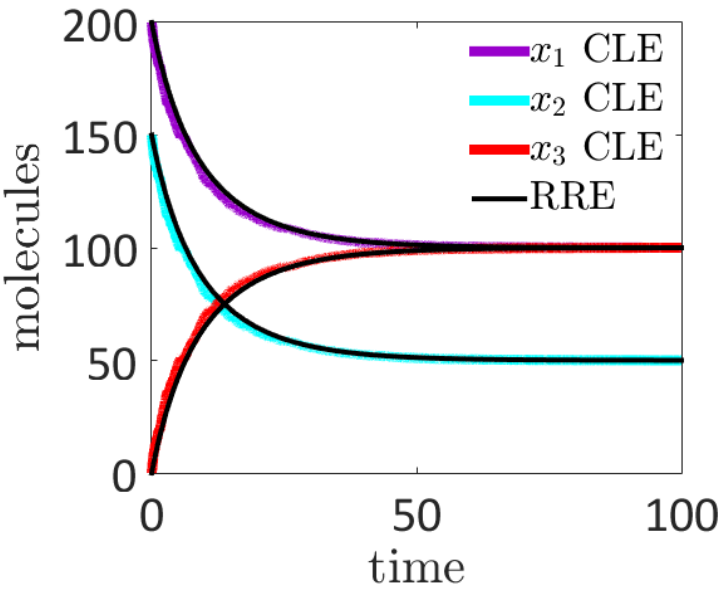


Figure 2.6: The figure is a simulation of the chemical Langevin equation compared with the reaction rate equation, described in detail in Section 2.3. We use the Euler-Maruyama approximation for 10^4 realisations, with $dt = 0.01$. Parameters and initial conditions similar to Fig. 2.3. The comparison with the simulation in Gillespie algorithm in Fig. 2.3 shows that the approximation applied in CME obtains more robust results, by running for a long enough time.

The solutions are shown in Fig. 2.3 against the stochastic simulations of the *Gillespie's* algorithm, and in Fig. 2.6 against the τ- leaping method. Another way to derive the RRE is to follow the law of mass action. The rate of a reaction is proportional to the product of the copy number of the reactants [28, 107], such that the following equation is satisfied:

$$\frac{c_+}{c_-} = \prod_j \frac{np_j^{p_j}}{nr_j^{d_j}}, \tag{2.64}$$

where p_j, d_j are the production and the decay coefficient and np, nr the copy numbers of products and reactants.

2.5 Chemical Fokker-Planck Equation

The chemical *Fokker-Planck* equation (CFPE) is a second order partial differential equation that can approximate the chemical master equation [106, 107, 124]. Using the *Taylor* expansion on

$$\alpha_j(\boldsymbol{n} - \boldsymbol{\nu_j})p(\boldsymbol{n} - \boldsymbol{\nu_j}, t)$$

up to second order, we obtain

$$\alpha_j(\boldsymbol{n}-\boldsymbol{\nu_j})p(\boldsymbol{n}-\boldsymbol{\nu_j},t) = \alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) - \boldsymbol{\nu_j}\frac{\partial}{\partial \boldsymbol{n}}\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) + \frac{\boldsymbol{\nu_j^2}}{2!}\frac{\partial^2}{\partial \boldsymbol{n}^2}\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) + \mathcal{O}(\boldsymbol{\nu_j^3}). \tag{2.65}$$

In the large volume limit an appropriate scaling shows that it is reasonable to ignore the higher order terms. The *Fokker-Planck* equation for a multivariate system of M reaction types is obtained as follows:

$$\begin{aligned}\frac{dp(\boldsymbol{n},t)}{dt} &= \sum_{j=1}^{M}\left(\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) - \boldsymbol{\nu_j}\frac{\partial}{\partial n_j}\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) - \alpha_j(\boldsymbol{n})p(\boldsymbol{n},t)\right) \\ &+ \sum_{i=1}^{M}\sum_{j=1}^{M}\frac{\boldsymbol{\nu_i}\boldsymbol{\nu_j}}{2!}\frac{\partial^2}{\partial n_i \partial n_j}\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) \qquad (2.66) \\ &= \sum_{j=1}^{M} -\boldsymbol{\nu_j}\frac{\partial}{\partial n_j}\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M}\boldsymbol{\nu_i}\boldsymbol{\nu_j}\frac{\partial^2}{\partial n_i \partial n_j}\alpha_j(\boldsymbol{n})p(\boldsymbol{n},t) \qquad (2.67)\end{aligned}$$

Setting $\boldsymbol{\nu_j}\alpha_j(\boldsymbol{n}) = A_j(\boldsymbol{n})$ and $\boldsymbol{\nu_i}\boldsymbol{\nu_j}\alpha_j(\boldsymbol{n}) = B_{ij}(\boldsymbol{n})$ in Fig. (2.66), we get the general form of the *Fokker-Planck* equation

$$\frac{\partial p(\boldsymbol{n},t)}{\partial t} = -\sum_{i=1}^{M}\frac{\partial}{\partial n_i}(A_i(\boldsymbol{n})p(\boldsymbol{n},t)) + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M}\frac{\partial^2}{\partial n_i \partial n_j}(B_{ij}(\boldsymbol{n})p(\boldsymbol{n},t)), \tag{2.68}$$

where the matrices A_i, B_{ij} are real, with B_{ij} non-negative and symmetric.

We now fix the starting time $t_0 = 0$ and set the initial condition for all x as

$$\lim_{t\to 0} p(\boldsymbol{n},t) = \prod_j \delta(n_j - n_{j0}). \tag{2.69}$$

By choosing reflecting boundary conditions [106, 107] and defining with J_j the probability flux J_j as

$$J_j(\boldsymbol{n},t) = A_j(\boldsymbol{n})p(\boldsymbol{n},t) - \frac{1}{2}\sum_j \frac{\partial}{\partial n_j}(B_{ij}(\boldsymbol{n})p(\boldsymbol{n},t)), \tag{2.70}$$

the *Fokker-Planck* equation (2.68) takes the form

$$\frac{\partial p(\boldsymbol{n},t)}{\partial t} = -\sum_j \frac{\partial}{\partial n_j}J_j(\boldsymbol{n},t), \tag{2.71}$$

where we consider $J_j(\boldsymbol{n},t) = 0$. The stationary distribution $p_s(\boldsymbol{n})$ can be derived as follows

$$0 = A_j(\boldsymbol{n})p_s(\boldsymbol{n}) - \frac{1}{2}\sum_j p_s(\boldsymbol{n})\frac{\partial}{\partial n_j}B_{ij}(\boldsymbol{n}) - \frac{1}{2}\sum_j B_{ij}(\boldsymbol{n})\frac{\partial}{\partial n_j}p_s(\boldsymbol{n}). \tag{2.72}$$

Rearranging we get

$$p_s(\boldsymbol{n})[2A_j(\boldsymbol{n}) - \sum_j \frac{\partial}{\partial n_j}B_{ij}(\boldsymbol{n})] = \sum_j B_{ij}(\boldsymbol{n})\frac{\partial}{\partial n_j}p_s(\boldsymbol{n}). \tag{2.73}$$

Dividing by $p_s(\boldsymbol{n})$ and assuming that $B_{ij}(\boldsymbol{n})$ is invertible as to

$$\frac{\partial}{\partial n_j} p_s(\boldsymbol{n}) \frac{1}{p_s(\boldsymbol{n})} = \sum_k B^{-1}_{jk}(\boldsymbol{n})[2A_j(\boldsymbol{n}) - \sum_j \frac{\partial}{\partial n_j} B_{ij}(\boldsymbol{n})], \tag{2.74}$$

so that we can write

$$\frac{\partial}{\partial n_j} \ln p_s(\boldsymbol{n}) = \sum_k B^{-1}_{jk}(\boldsymbol{n}) \left[2A_k(\boldsymbol{n}) - \sum_j \frac{\partial}{\partial n_j} B_{ki}(\boldsymbol{n}) \right] = F_l(A, B, \boldsymbol{n}), \tag{2.75}$$

where the right term in Eqn. (2.75) needs to satisfy the potential conditions [106] :

$$\frac{\partial F_j}{\partial n_i} = \frac{\partial F_i}{\partial n_j}.$$

Stationary probability in one dimension: In order to calculate analytically the probability density we consider the one dimensional *Fokker-Planck* equation

$$\frac{\partial p(n,t)}{\partial t} = -\frac{\partial}{\partial n}(A(n)p(n,t)) + \frac{1}{2}\frac{\partial^2}{\partial n^2}(B(n)p(n,t)). \tag{2.76}$$

After combining the results from Eqns. (2.69), (2.75) and (2.76), and taking $\dfrac{\partial p(n,t)}{\partial t} = 0$, the stationary solution $p_s(n)$ is:

$$p_s(n) = cB^{-1}(n) \exp\left(2\int_0^n A(s)B^{-1}(s)ds \right), \tag{2.77}$$

where c is a positive constant of integration. By setting

$$2\int_0^n A(s)B^{-1}(s)ds = -\Phi(n), \tag{2.78}$$

$p_s(n)$ is proportional to the function $\Phi(n)$, which is defined as the potential, and we get the stationary probability in the form

$$p_s(n) = cB^{-1}(n)e^{-\Phi(n)} \tag{2.79}$$

$$= B^{-1}(n)e^{-\Phi(n)} \left(\int_0^{+\infty} B^{-1}(s)e^{-\Phi(s)}ds \right). \tag{2.80}$$

This is called the stationary partition function, where c is obtained by applying the normalized $p_s(n)$ for all $t > 0$:

$$\int_0^{+\infty} p_s(s)ds = 1, \tag{2.81}$$

as

$$c = \int_0^{+\infty} B^{-1}(s)e^{-\Phi(s)}ds, \tag{2.82}$$

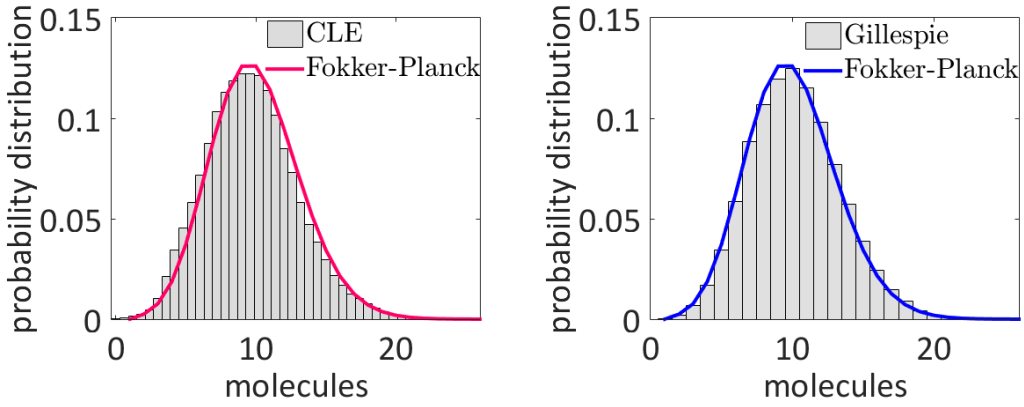


Figure 2.7: The two figures represent a comparison of the chemical Langevin equation (left image), and the Gillespie algorithm (right image), for a long simulation for 10^6 realisations, with the exact solution derived from the Fokker-Planck for the production-degradation example in Eqn. (2.86). The parameters used are $r_1 = 0.004, r_2 = 0.04$, starting with initial condition $x_0 = 0$ molecules, for a fixed time $T = 1$.

and Eqn. (2.80) is finally expressed as

$$p_s(n) = B^{-1}(n) \exp\left(\int_0^n 2A(s)B^{-1}(s)ds\right) \int_0^{+\infty} B^{-1}(s) \exp\left(\int_0^n 2A(s)B^{-1}(s)ds\right) ds. \tag{2.83}$$

The derivation of the Fokker-Planck equation as shown in Eqns.(2.80) and (2.81) leads to the stationary distribution. The distribution can be derived from the Fokker-Planck Eqn. (2.76) for the production-degradation example in Eqn. (2.14) as follows:

$$\frac{\partial p(n,t)}{\partial t} = -\frac{\partial}{\partial n}\left((r_2 - r_1 n)\, p(n,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial n^2}\left((r_2 + r_1 n)\, p(n,t)\right). \tag{2.84}$$

Now substitute in Eqn. (2.77) $A(n) = r_2 - r_1 n$ and $B(n) = (r_2 + r_1 n)/2$ from the Fokker-Planck Eqn. (2.84) to obtain the potential in Eqn. (2.78)

$$\Phi(n) = -2\int_0^n \frac{r_2 - r_1 s}{r_2 - r_1 s} ds = -\left[\frac{4r_2}{r_1}\log(r_2 + r_1 n) - 2n\right], \tag{2.85}$$

and finally derive the stationary distribution

$$P_s(n) = 2c \exp\left[\left(\frac{4r_2}{r_1} - 1\right)\log(r_2 + r_1 n) - 2n\right], \tag{2.86}$$

where the constant c is given by

$$2c = \left(\int_0^{+\infty} \exp\left[\left(\frac{4r_2}{r_1} - 1\right)\log(r_2 + r_1 n) - 2n\right] ds\right)^{-1}. \tag{2.87}$$

Chapter 3

Noise Processing by Signalling Networks

In this chapter we interpret a network from another perspective; we model the biochemical reactions with the Ornstein-Uhlenbeck process and describe the network by its equivalent graph. Whether or not two species interact is decided by the adjacency matrix. The noise here is considered extrinsic and we focus on how it can be propagated through a complex network. One question that we are interested to answer is: '*What is the strength and the manner that extrinsic noise is circulated and transmitted on nodes of particular interest, through a directed graph?*' In order to investigate the effect of noise on any node of interest in a signalling network, we derive a mathematical formula for the variance at the equilibrium. This equation relates the structure of the underlying graph and the noise processing from a noisy input(s) to the output that we are interested in. We examine the stability criteria of the underlying deterministic system that is essential for the existence of an invariant measure. Under the assumption of stability and that the noise is coming from one input we derived a mathematical formula that allows us to calculate the variance of any node, depending on the paths to that node from the input.

In the next section, in order to demonstrate the general expression of the variance, we look at two common biological motifs; a non-reversible chain and a feed-forward loop. Applying the result from the variance in these examples, we can explain with our formula and simulations how the structure determines the effect of noise. In the last section we focus on the pluripotent signalling network. Our mathematical expression for the limiting variance is applied to a simplified version of the transcriptional regulatory network for naive pluripotency. We justify and deduce the importance of the length and the coherence of the walks, from the input to the target in the noise propagation.

3.1 Noise and Structure of a Graph

Our concern here is with how a network G processes noise from an external source. In the context of signalling networks the nodes in the network denote molecules in the signalling cascades, and edges are regulatory interactions (e.g. phosphorylation etc.) between molecules. Since signalling networks pass information from the cell exterior to the nucleus, we assume that the network G is inherently directed: the presence of an edge (i, j) indicates that node i exerts a regulatory effect on node j but not necessarily vice versa. Since regulatory interactions may be activatory or inhibitory we also allow each edge to have positive or negative weight representing the strength of activation or inhibition respectively. We denote the weight of edge (i, j) by a_{ij}. Assuming that there are n nodes in G the $n \times n$ adjacency matrix $\boldsymbol{A}$ then describes the strength of all interactions in the system, as is defined in Eqn. (1.3).

The higher powers of $\boldsymbol{A}$ contribute essentially in the network's organisation. The square matrix $\boldsymbol{A}^2$ gives the number of the paths of length two from node i to j, when i and j are connected. In general, the higher powers $\boldsymbol{A}^n$ of the adjacency $\boldsymbol{A}$ of a graph $\mathcal{G}$ describe the strength of two nodes i and j are connected via a walk of length n [125], by counting the number of the corresponding length of walks. Considering now the sum of all the powers of $\boldsymbol{A}$ scaled by $1/n!$ as the weighted sum of all the possible walks between a pair of nodes, we come up eventually to the communicability measure of the exponential of $\boldsymbol{A}$ [125–128] defined as:

$$\exp(\boldsymbol{A}) = \sum_{n=0}^{\infty} \frac{1}{n!} \boldsymbol{A}^n. \tag{3.1}$$

The exponential matrix of $\boldsymbol{A}$ is essential for the analytical calculations for the covariance matrix $\boldsymbol{K}$, as well as for our conclusions about significant walks and on important interactions.

As an important 'centrality' measure of $\mathcal{G}$ we consider the weighted degree d_i [125, 129, 130] of a node i, defined by

$$d_i = \sum_{j=1}^{N} |a_{ij}| + \sum_{j=1}^{N} |a_{ji}|, \tag{3.2}$$

which contains the number of adjacent nodes i is connected, *i.e.* d_i indicates the number of edges i is involved. In particular, in a directed graph $\mathcal{G}$ we distinguish between the in-degree $d_i^{\text{in}} = \sum_{j=1}^{N} |a_{ji}|$ and out-degree $d_i^{\text{out}} = \sum_{j=1}^{N} |a_{ij}|$ of a node i, which counts the adjacent nodes entering and leaving i, respectively [9, 131, 132].

In general, regulatory interactions may be highly non-linear. However, to better understand the relationship between network structure and function we will assume here that that the dynamics are linear. By doing so we are effectively considering the linearisation near to a fixed point in the non-linear dynamics; this rationale for

studying the linear case has been taken elsewhere [133]. We focus now on how noise from one or multiple inputs is processed through the network described by graph $\mathcal{G}$. In order to investigate this we model the process of biochemical reactions with the Ornstein-Uhlenbeck process that has been defined in Eqn. (1.10):

$$d\boldsymbol{x}(t) = -\boldsymbol{M}\boldsymbol{x}(t)dt + \boldsymbol{\Sigma}d\boldsymbol{W}(t),$$

where $\boldsymbol{x} \in \mathbb{R}^n$ represents the species vector, and $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ is defined as

$$\boldsymbol{M} = d\boldsymbol{I} - \boldsymbol{A}, \tag{3.3}$$

where $\boldsymbol{I}$ is the $n \times n$ identity matrix, and we have assumed that all nodes decay at the same rate d, which is meant to represent the decay time scale for the dynamics. Without loss of generality we may take $d = 1$ since this may always be achieved by suitable re-scaling.

Given the linearity of this system there are only two possible types of long-term behaviour: convergence toward a stable fixed point or divergence to infinity. We will assume that only the first behaviour can happen, *i.e.* convergence to a stable fixed point is the only physically realistic scenario. This occurs whenever the real parts of the eigenvalues of $\boldsymbol{M}$ are all strictly positive, *i.e.* $\boldsymbol{M}$ is a positive definite. Then, the SDE in Eqn. (1.10) is ergodic: there is a unique stationary distribution to which any initial conditions Eqn. (1.10) can converge and therefore admits a unique invariant measure[106]. Properties of the network $\mathcal{G}$ for which the associated deterministic of Eqn. (1.10) has a stable solution have been discussed at length, and it is known that sparse modular networks confer stability, for example [133, 134].

The matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times m}$ has the noise coefficients, and $\boldsymbol{W}(t)$ is the m-dimensional standard Brownian motion, as defined in Section. 1.3.3. The solution of Eqn. (1.10) for initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$ is given by

$$\boldsymbol{x}(t) = e^{-\boldsymbol{M}t}\boldsymbol{x}_0 + \int_0^t e^{-\boldsymbol{M}(t-s)}\boldsymbol{\Sigma}d\boldsymbol{W}(s), \tag{3.4}$$

with mean given by

$$\langle \boldsymbol{x}(t)\rangle = e^{-\boldsymbol{M}t}\boldsymbol{x}_0, \tag{3.5}$$

and by applying Itô calculus [106, 112] covariance matrix $\boldsymbol{K}(t)$:

$$\boldsymbol{K}(t) = \langle \boldsymbol{x}(t) - \langle \boldsymbol{x}(t)\rangle, (\boldsymbol{x}(t) - \langle \boldsymbol{x}(t)\rangle)^T\rangle = \int_0^t e^{-\boldsymbol{M}(t-s)}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T e^{-\boldsymbol{M}(t-s)^T}ds. \tag{3.6}$$

If Eqn. (1.10) describes an ergodic process then

$$\lim_{t\to\infty} \boldsymbol{K}(t) = \boldsymbol{K}, \tag{3.7}$$

where $\boldsymbol{K}$ satisfies the Lyapunov equation [106]:

$$\boldsymbol{M}\boldsymbol{K} + \boldsymbol{K}\boldsymbol{M}^T = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T. \tag{3.8}$$

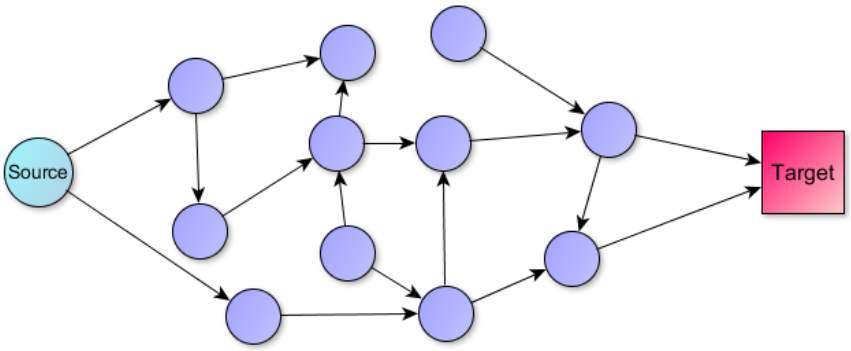


Figure 3.1: An acyclic graph that the source is passing a noisy signal to the target through a number of paths.

Although this is the standard formulation [106], instead of working with Eqn. (3.8) we will work directly with the Eqn. (3.6) as it ultimately allows for a more transparent assessment on the effects of network structure on the stationary covariance of the system.

Since our purpose is to determine the way in which input noise is processed by the network $\mathcal{G}$ it is natural to consider a single noisy input to the system, which represents the fluctuating extra-cellular environment, and a single output, representing the computational core of the network. To do so we may, without loss of generality, chose a labelling of the nodes such that the first node is the noisy input or the source, and the n-th node is the output, namely the target. Thus, we set

$$\mathbf{\Sigma} = (\sigma, 0, \cdots, 0)^T \tag{3.9}$$

and we are interested in calculating the variance of the n-th node in the network, which is given by K_{nn}, relative to the magnitude of the input noise. Fig. 3.1 illustrates an example of a network of interest.

The environmental noisy signals can be bearing essential information, thus can be thought as a measure of the degree the target node can capture; or just random environmental fluctuations that can be thought as a measure of the degree of the network amplifies or suppresses that fluctuations.

3.1.1 Derivation of Limiting Variance

We will derive a formula that relates the limiting variance of the target with the structure of the network of interest.

In particular, the $K_{ij}(t)$ element of the covariance in Eqn. (3.6), in index notation is given by

$$K_{ij}(t) = \int_0^t (e^{-M(t-s)})_{il}(\Sigma)_{lk}(\Sigma^T)_{km}(e^{-M(t-s)^T})_{mj}ds. \tag{3.10}$$

Due to the fact that Σ in Eqn. (3.9) has only one non-zero entry $(\Sigma\Sigma^T)_{lm} = \sigma^2$ for $l = m = 1$ and by using Eqn. (3.3), we obtain

$$K_{ij}(t) = \sigma^2 \int_0^t e^{-2(t-s)} \left(e^{A(t-s)}\right)_{i1} \left(e^{A^T(t-s)}\right)_{1j} ds. \tag{3.11}$$

For the target of interest, *i.e.* the n-th node we replace $i = j = n$

$$K_{nn}(t) = \sigma^2 \int_0^t e^{-2(t-s)} \left(e^{A(t-s)}\right)_{n1} \left(e^{A^T(t-s)}\right)_{1n} ds, \tag{3.12}$$

which is equivalent with

$$K_{nn}(t) = \sigma^2 \int_0^t e^{-2(t-s)} \left(e^{A(t-s)}\right)_{n1}^2 ds. \tag{3.13}$$

This expression connects strongly to the structure of the network via the fact that (i, j)-th entry of the exponential of the adjacency matrix of a network is a weighted sum of all walks between nodes i and j, and so is a simple measure of network 'communicability' [125], as described in Section 3.1.

Before proceeding further with the Eqn. (3.13), we will consider the weight $w(P)$ of a walk P from node i to node j as the product of its edge weights,

$$w(P) = \prod_{(i,j)\in P} a_{ij}. \tag{3.14}$$

If P_k is the set of walks of length k between nodes i and j then the total weight of all walks from node i to node j with length k is

$$\beta_{ijk} = a_{ij}^k = \sum_{P_k} w(P_k). \tag{3.15}$$

Combining Eqn. (3.1) and according to [99, 125], the exponential of the adjacency matrix $\boldsymbol{A}$ in Eqn. (3.11) can be written as

$$\left(e^{A(t-s)}\right)_{ij} = \sum_{k=0}^{\infty} \frac{\beta_{ijk}}{k!} (t-s)^k. \tag{3.16}$$

Using Eqn. (3.16) and the formulation of Eqn. (3.15), and by considering the convention $\beta_{ik} = \beta_{1ik}$ and $\beta_{ij0} = \delta_{ij}$, Eqn. (3.11) takes the form

$$K_{ij}(t) = \sigma^2 \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \int_0^t \frac{\beta_{ik}\beta_{jl}}{k!l!} (t-s)^{k+l} e^{-2(t-s)} ds, \tag{3.17}$$

or for the n-th node Eqn. (3.13) can be expressed as

$$K_{nn}(t) = \sigma^2 \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \int_0^t \frac{\beta_{nk}\beta_{nl}}{k!l!} (t-s)^{k+l} e^{-2(t-s)} ds. \tag{3.18}$$

If now we take the limit at $t \to \infty$ and consider the fact that

$$\lim_{t\to\infty} \int_0^t (t-s)^n e^{-2(t-s)} ds = \frac{n!}{2^{n+1}} \tag{3.19}$$

we simplify Eqn. (3.17) and obtain the limiting covariance

$$K_{ij} = \frac{\sigma^2}{2} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{2^{k+l}} \binom{k+l}{l} \beta_{ik}\beta_{jl}, \tag{3.20}$$

or more specific the limiting variance of the target node n:

$$K_{nn} = \frac{\sigma^2}{2} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{2^{k+l}} \binom{k+l}{l} \beta_{nk}\beta_{nl}. \tag{3.21}$$

Since the variance of the noisy input is given by $K_{11} = \sigma^2/2$, we may investigate the noise-processing ability of the network by considering the ratio

$$R = \frac{K_{nn}}{K_{11}} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{2^{k+l}} \binom{k+l}{l} \beta_k\beta_l, \tag{3.22}$$

where we have further simplified notation by setting $\beta_{nk} = \beta_k$.

When $R > 1$ noise is amplified by the network; when $R < 1$ noise is suppressed by the network. Importantly, if the process described by Eqn. (1.10) is ergodic then R is finite and depends only on the structure of the network $\mathcal{G}$. This formula therefore provides an explicit connection between network architecture and noise-processing; our interest is to determine how R is affected by different network architectures. To do so we note that Eqn. (3.22) has a natural interpretation in terms of random walks on $\mathcal{G}$, as follows.

Since each walk P from the input to the target has an associated weight $w(P)$, a pair of possibly intersecting walks P, Q from the input to the target also has an associated weight $w(P,Q) = w(P)w(Q)$, the product of the edge weights involved. If we write P_k and P_l for arbitrary walks from the input to the target of length k, and l respectively, then the product $\beta_k\beta_l$ (see Eqn. (3.21)) can be written as

$$\beta_k\beta_l = \sum_{P_k} w(P_k) \sum_{P_l} w(P_l) = \sum_{P_k,P_l} w(P_k, P_l). \tag{3.23}$$

Substituting this into Eqn. (3.22) and rewriting the second sum in terms of $m = k + l$ gives,

$$R = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \sum_{P_k,P_{m-k}} \frac{1}{2^m} \binom{m}{k} w(P_k, P_{m-k}), \tag{3.24}$$

from which it can be seen that R is a weighted sum of all pairs of walks through $\mathcal{G}$ from the source node to the target, with the relative importance of each walk-pair determined by a coefficient drawn from a binomial distribution $\mathcal{B}(m, 1/2)$, where m is the length of the walk-pair.

The appearance of binomial probabilities arises as the natural probability measure for pairs of random walkers on the network $\mathcal{G}$. This can be discerned by a simple example: let us consider two independent walks starting at the same time at the input node. At each time step a node is selected with equal probability 1/2; this walker decides through available edges in $\mathcal{G}$ the with equal probability; *i.e.* if the walker is at node i each possible edge is chosen with probability $1/d_i^{\text{out}}$, where d_i^{out} is the out-degree of node i, as defined throughout Eqn. (3.2). Then, after $m = k + l$ steps both walks reach the target node with probability $\binom{m}{k}2^{-m}$. Now, we can think the sum $\sum_{P_k, P_{m-k}}$ in Eqn. (3.24) as the expected weight of a pair of walks from input to target, with respect to the probability for two independent random walks; since $\boldsymbol{K}(t)$ in Eqn. (3.11) depends on two exponential terms, $e^{A(t-s)}$ and $e^{A^T(t-s)}$, is attributed to the fact that we have two random walks.

How the ratio in Eqn. (3.24) grows depends strongly on the graph's $\mathcal{G}$ structure; a directed acyclic graph for example will have no feedback loops, and thus the sum $\sum_{m=0}^{\infty}$ in Eqn. (3.24) is finite, due to the finite length of each walk. On the other hand, a graph with cycles may have arbitrarily long walks, and one needs to take infinite terms into account in Eqn. (3.24). Therefore, positive feedback adds infinitely positive terms to Eqn. (3.24), and deliver more noise in the target than a corresponding acyclic; similarly, negative feedback adds infinitely both positive and negative terms to this sum; however, may amplify or diminish noise, depending on the distinct patterns of inhibitory interactions in the network.

3.1.2 Stability of a Graph

Stability of the underlying deterministic system is necessary for the existence of an invariant measure. Let us consider the simplest case in a directed graph $\mathcal{G}$ when all weights have the same magnitude α, for the matrix $\boldsymbol{M} = \alpha\boldsymbol{A} - \boldsymbol{I}$ (here $d = 1$, see Eqn. (3.3)). One way to ensure stability and hence ergodicity is to consider the strength of the weights α as follows

$$\alpha < \frac{1}{\max_i(d_i)}, \tag{3.25}$$

where here $\max_i(d_i)$ is the maximum degree of a node i (d_i defined in Eqn. (3.2)). Although this is a sufficient but not a necessary condition for stability.

Another sufficient criterion for stability is an acyclic graph, as is depicted in Fig. 3.1. In this case, the adjacency matrix $\boldsymbol{A}$, which is strictly upper triangular, guarantees stability and Eqn. (1.3) takes the form

$$\boldsymbol{A} = \begin{cases} a_{ij} \in \mathbb{R},\ i < j \\ a_{ij} = 0,\ i \geq j \end{cases}, \tag{3.26}$$

where its eigenvalues are the entries on the main diagonal. That ensures that all the eigenvalues of matrix $-\boldsymbol{M}$ from Eqn. (3.3) have negative real parts, *i.e.* $\lambda_i = -1 \quad \forall i$.

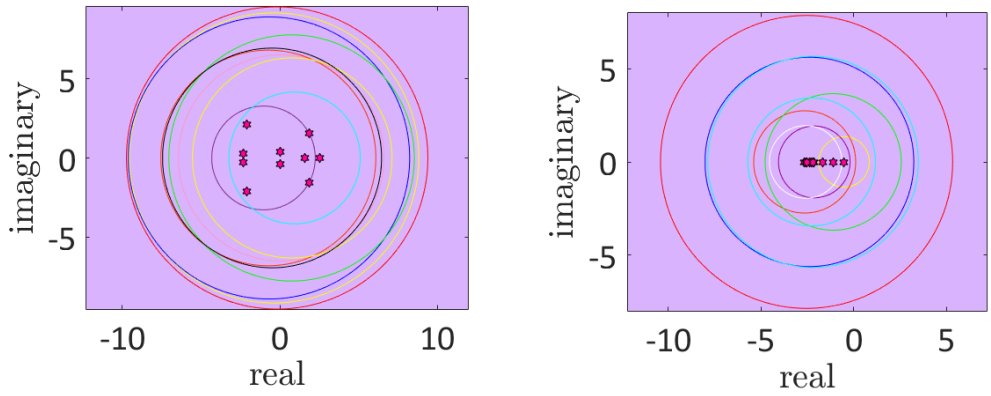


Figure 3.2: The images illustrate the area the eigenvalues are bounded in the complex plane, according to the Gershgorin circle theorem. The left figure shows the eigenvalues of a random matrix with normally distributed entries, while the right represents the eigenvalues of an upper triangular matrix (with distinct diagonal entries), as it is in an acyclic graph.

A way of eigenvalues estimation of a matrix is by using the Gershgorin circle theorem, which gives the area the eigenvalues are bounded. For real square matrix $\boldsymbol{A}$, each eigenvalue λ_i is localized in the disk

$$|\lambda_i - a_{ii}| < \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}|, \tag{3.27}$$

and is centred on the diagonal element a_{ii}, with the radius the sum of the absolute values of the non-diagonal entries in every row. Therefore, since every eigenvalue λ_i of A lies on $\mathbb{C}$, can be represented as a point in the complex plane of distance $dist_i$ from a_{ii} as follows

$$dist_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}, \tag{3.28}$$

where r_i is $\sum_{i=1 i \neq j}^{n} |a_{ij}|$. The Gershgorin disks are illustrated in Fig. 3.2 for two random matrices with normally distributed elements. The right figure indicates that all the real parts of the eigenvalues of an upper triangular matrix are negative, as could be in an acyclic graph. The images show that least one eigenvalue lies in each disk. Gershgorin circle theorem indicates a sufficient but not necessary condition for stability, since every stable system satisfies the theorem, but the reverse is not always true.

In the case of a real square $n \times n$ matrix $\boldsymbol{A}$, with normally distributed elements and eigenvalues λ_i, Girko's circular law is satisfied [135, 136]. As n approaches infinity the normalised quantity

$$\frac{\lambda_i}{\sqrt{n}} \tag{3.29}$$

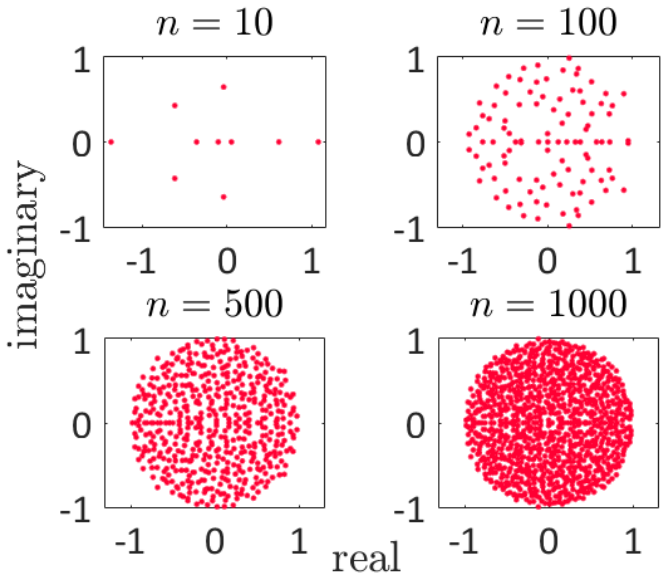


Figure 3.3: The images illustrate how the normalised eigenvalues of a square real matrix form the unit disk, as n gets large. The figure describes the Girko circular law.

is uniformly distributed on a disk with centre zero and unit radius. This property is illustrated in Fig. 3.3; as n gets larger the unit disk that the quantities from Eqn. (3.29) form can be obviously seen.

The mathematical equation proved in Section 3.1.1 in Eqn. (3.20) can be applied in any system that is considered stable. In the next sections we consider two examples that the graphs are acyclic, and therefore satisfy the stability criterion, as well as an example of the pluripotent network considering condition in Eqn. (3.25).

3.2 Chemical Reaction Chain

In order to illustrate the general expression of the variance as is derived in Eqn. (3.20), we give two examples, starting with the simplest and considering a non-reversible chain. A species x_i produces x_{i+1} with production rate a_i. The chemical chain in Fig. 3.4 of n species is modelled as follows

$$\begin{aligned} dx_1(t) &= -x_1(t)dt + \sigma dW(t) \\ dx_i(t) &= (a_{i-1}x_{i-1}(t) - x_i(t))dt, \end{aligned} \tag{3.30}$$

or is equivalent with the Eqn. (1.10), where matrix $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{A}$ takes the form

$$\boldsymbol{M} = \begin{cases} 1, & \text{for } i = j \\ -a_{i-1}, & \text{for } j = i+1 \\ 0, & \text{otherwise.} \end{cases} \tag{3.31}$$

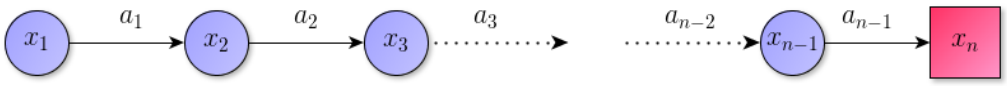


Figure 3.4: An acyclic graph that the source x_1 is passing a noisy signal to the target x_n through a weighted path of length n_1.

Here there is only one path P from the source to the target of length $n-1$, of weight given by

$$w(P) = \begin{cases} \beta_i = \beta_j = \prod_{i=i}^{n-1} a_i, & \text{for } i = j = n \\ \beta_i = \beta_j = 0, & \text{for } i \neq j \neq n. \end{cases} \tag{3.32}$$

The limiting variance derived in Eqn. (3.20), expanded the binomial coefficient, takes the form

$$K_{nn} = \sigma^2 \frac{(2n)!}{n!^2} \frac{1}{2^{2n+1}} \prod_{i=1}^{n-1} a_i^2. \tag{3.33}$$

Our result concurs with the expression derived in [99], where $a_i = k$, $\forall i$. Now by using the ration R from Eqn. (3.22) we obtain

$$R = \frac{(2n)!}{n!^2} \frac{1}{2^{2n}} \prod_{i=1}^{n-1} a_i^2. \tag{3.34}$$

In Eqns. (3.34) and (3.33), we keep the decay rates equal and allow positive and negative distinct values for the production rates a_i for each node. The result indicates that for a fixed noise coefficient σ, for small enough a_i, satisfy $a_i < 1$, K_{nn} is decreasing logarithmically; thus the longer is the chain the smaller K_{nn} becomes. Nevertheless, larger $a_i > 1$ can amplify K_{nn}; the product can exceed the binomial term and hence reinforce the variance of the target. The signs do not affect the convergence of the variance, as indicated from the square weights in the product $\prod_{i=1}^{n-1} a_i^2$, and this symmetry is illustrated in Fig. 3.5. Fig. 3.5 depicts the variance of the n-th node against the weights for $a_i = a$ and the number of nodes n in the chain.

We now proceed with analysing more the case when $a_i = a$, in order to investigate how the magnitude of a affects the noise processing abilities of the underlying network. Then for large n we can use the Stirling's approximation and can be shown that the ratio in Eqn. (3.34) can be simplified in the form

$$R \sim \frac{a^{2(n-1)}}{\sqrt{\pi n}}. \tag{3.35}$$

Now we can show that if

$$|a| < (\pi n)^{1/4(n-1)}, \tag{3.36}$$

the signalling pathway suppresses noise. By using Stirling's approximation:

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n, \tag{3.37}$$

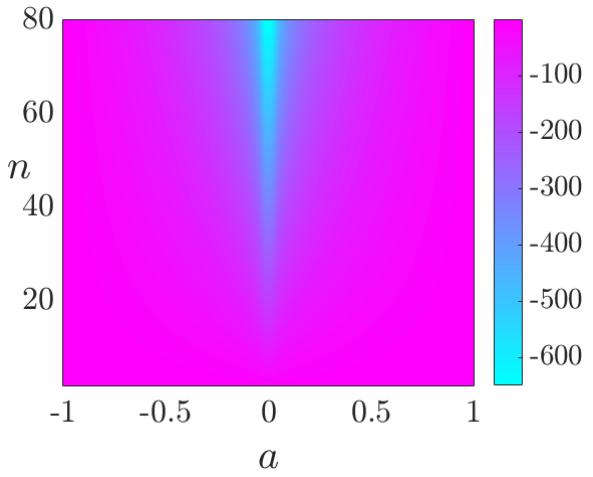


Figure 3.5: The log of the variance of the n-th node in a linear chain against the weights with same magnitude a and the number of nodes n.

we can prove that the ratio R converges for a large n as follows:

$$R \sim \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{2\pi n\left(\frac{n}{e}\right)^{2n}}\frac{1}{2^{2n}}a^{2(n-1)} \tag{3.38}$$

$$= \frac{a^{2(n-1)}}{\sqrt{\pi n}}. \tag{3.39}$$

Then, *if and only if* $a^{4(n-1)} < \pi n$, and by taking the limit to the infinity:

$$\lim_{n\to\infty} R = \lim_{n\to\infty}\frac{(2n)!}{n!^2}\frac{a^{2(n-1)}}{2^{2n}} = \lim_{n\to\infty}\frac{a^{2(n-1)}}{\sqrt{\pi n}} = 0. \tag{3.40}$$

The limit at which n approaches infinity is less that one, hence R converges.

3.2.1 Uncorrelated Normal Distributed Random Weights

We have shown that for fully correlated weights $a_i = a$, and for any $a_i < 1$ the chain can suppress the noise in the n-th node, under specific conditions. It is interesting now to assume the random weights taken from a normal distribution. In the following calculations the weights a_i are considered independent and identically distributed (i.i.d.), so that $a_i \sim \mathcal{N}(0, \gamma^2)$. In order to draw conclusions for the weights a_i, we derive the mean and the variance of the limiting variance in Eqn. (3.33). Then, the second central moment of their product is given by

$$\mathbb{E}[\prod_{i=1}^{n-1} a_i^2] = \prod_{i=1}^{n-1}\mathbb{E}[a_i^2] = \gamma^{2(n-1)}. \tag{3.41}$$

Then the mean of the variance K_{nn} is given by:

$$\mathbb{E}[K_{nn}] = \sigma^2\frac{(2n)!}{n!^2}\frac{1}{2^{2n+1}}\mathbb{E}[\prod_{i=1}^{n-1} a_i^2] = \sigma^2\frac{(2n)!}{n!^2}\frac{1}{2^{2n+1}}\gamma^{2(n-1)}. \tag{3.42}$$

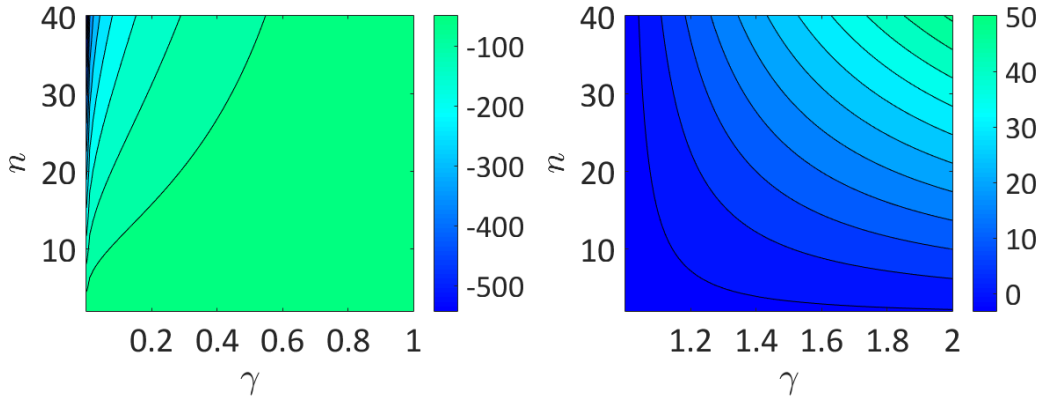


Figure 3.6: The figures are the contour plots of the log of the variance K_{nn} against $\gamma < 1$ (left) and $\gamma > 1$ (right), respectively. The weights a_i are considered independent and identically distributed (i.i.d.), so that $a_i \sim \mathcal{N}(0, \gamma^2)$.

In order to have a finite expected value of variance K_{nn}, variance γ^2, of the weights a_i needs to satisfy the condition $\gamma^2 < 1$, so that $\lim_{n\to\infty} \mathbb{E}[K_{nn}] = 0$. Then the variance of K_{nn} is derived as follows

$$\mathrm{Var}(K_{nn}) = \mathbb{E}[K_{nn}^2] - \mathbb{E}^2[K_{nn}] \tag{3.43}$$

$$= \sigma^4 \left(\frac{(2n)!}{n!^2} \frac{1}{2^{2n+1}} \right)^2 \left((3\gamma^4)^{n-1} - \gamma^{4(n-1)} \right), \tag{3.44}$$

where $3\gamma^4$ is the forth order central moment [106]. By using now Stirling's approximation we can rewrite Eqn. (3.43) as follows

$$\mathrm{Var}(K_{nn}) \sim \frac{\sigma^4}{4\pi n} \left[(3\gamma^4)^{n-1} - (\gamma^4)^{n-1} \right] = \frac{\sigma^4}{4\pi n} (3^{1/4}\gamma)^{4(n-1)}. \tag{3.45}$$

Interestingly, we found that while the mean of the variance K_{nn} in the chain is finite, as it is shown in Eqn. (3.41), the variance can be infinite, for γ satisfying the condition $3^{1/4}\gamma > 1$.

As a measure of variability we consider now the coefficient of variation $CV(K_{nn})$, which describes the degree of variability relative to the mean, and is calculated as

$$CV(K_{nn}) = \frac{\sqrt{Var(K_{nn})}}{\mathbb{E}[K_{nn}]} = \sqrt{3^{n-1} - 1} \simeq 3^{(n-1)/2}. \tag{3.46}$$

Eqn. (3.46) shows that as well as the variance of K_{nn} approaches infinity for the limit of $n \to \infty$, and is independent of γ. We conclude that normally i.d.d. weights *i.e.* $a_i \sim \mathcal{N}(0, \gamma^2)$ can amplify the noise, since for a finite mean the magnitude of K_{nn} variance can be burst to infinity for $|\gamma| > 3^{-1/4}$.

These results are demonstrated in Fig. 3.6, which is a contour plot of the *logarithm* of variance of the target node, K_{nn} for $\gamma < 1$ (left) and $\gamma > 1$ (right),

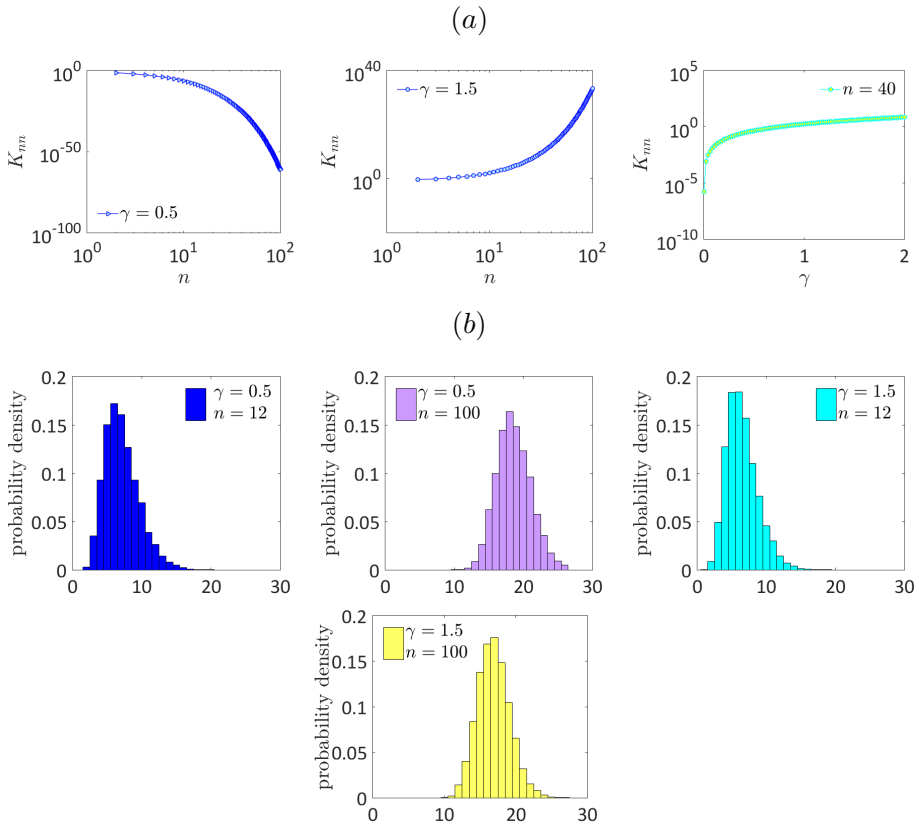


Figure 3.7: Figure (a) plots the variance against n for $\gamma = 0.5$ and $\gamma = 1.5$ and variance against γ for a fixed $n = 40$. Histograms (b) are depicting the probability density functions produced by fixing the number of nodes in a chain of length $n = 12$, (blue) and of length $n = 100$, (purple) for $\gamma = 0.5$. The square mean of the weights is fixed to $\gamma = 1.5$ and n considered $n = 12$, (cyan) and $n = 100$ (yellow). The y axis is the number of molecules. The histograms are produced for a simulation of 10^4 realisations.

respectively. The shapes in the contour plot are the level sets that the function K_{nn} is equal for the different combinations of matrix dimension and variance γ. Fig. 3.6 illustrates that for $\gamma < 1$ the chain suppresses the noise for large n, while $\gamma > 1$ amplifies the noise as n gets larger. This results are also illustrated in Fig. 3.7; these plots depict the variance against the length of chain n and the probability density functions for $\gamma < 1$ and $\gamma > 1$, in Figs. 3.7 (a) and (b) respectively. For $\gamma < 1$ the variance of the target goes to zero as the chain becomes longer. However, for $\gamma > 1$ the variance K_{nn} goes to infinity for larger n. The plots also confirm the conclusions in Eqn. (3.40).

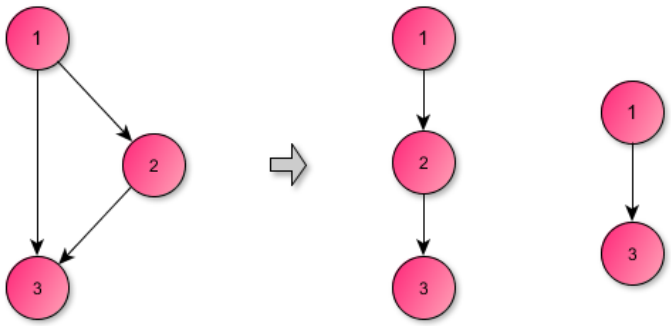


Figure 3.8: The image illustrates the ffl motif ('OR' cascade), where the target node (node 3) is regulated from the source (node 1) and an intermediate node (node 2).

3.3 Stochastic Feed-Forward Loop

Another example to consider is the well studied motif [21, 137], feed-forward loop (ffl), which is a directed graph of three nodes that is commonly found in transcription regulatory networks within the cell. Two nodes regulate the third and each other, as is shown in Fig. 3.8. A positive sign in a rate a_i corresponds to an activation, while a negative to a repression. ffl is known for acting as a molecular timer, since depending on the signs of the constant rates they accelerate or delay the response time of the target [21]. Depending on the sign of the total product of the direct and the indirect path from the source to the target, there are 8 different types. The four incoherent types of ffl accelerate the response time, while the other four coherent cause a delay [21]. Here we consider the '*OR*' cascade of the ffl motif, where the target is either regulated by the source directly, or indirectly by an intermediate node with the following linear ODEs

$$\begin{aligned} dx_1(t) &= -x_1(t)dt + \sigma dW(t) && (3.47)\\ dx_2(t) &= (a_1 x_1(t) - x_2(t))dt && (3.48)\\ dx_3(t) &= (a_2 x_2(t) + a_3 x_1(t) - x_3(t))dt. && (3.49) \end{aligned}$$

The adjacency matrix A takes the form

$$A = \begin{pmatrix} 0 & a_1 & a_3 \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.50}$$

There are two paths from the source to the target as shows Fig. 3.8, one of length two and one of length one. Applying the formula in Eqn. (3.18) the variance of the

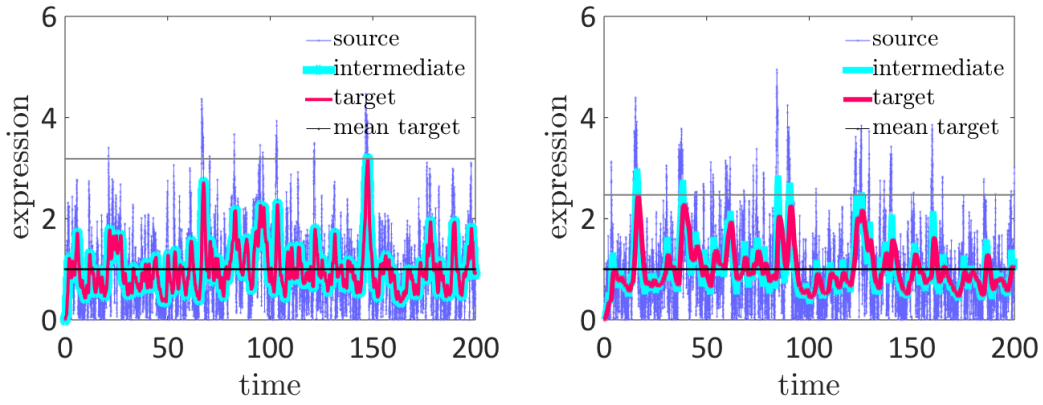


Figure 3.9: The left figure is a simulation of the stochastic ffl. The right figure is the stochastic chain for three nodes. Noise from the source (blue line) is delivered to the target (red line). As we can see from the comparison the the variance of the target in the chain (red line) fluctuations do not exceed expression magnitude for 2.5 (grey line). However, the ffl process to the target node more noise (red line), since fluctuations can reach a higher maximum magnitude, *i.e.* close to 4 (grey line).

target node is given by

$$\begin{aligned} K_{33}(t) &= \sigma^2 \int_0^t \beta_1^2 (t-s)^2 e^{-2(t-s)} ds \\ &+ \sigma^2 \int_0^t \beta_2^2 \frac{1}{2!^2} (t-s)^4 e^{-2(t-s)} ds \\ &+ \sigma^2 \int_0^t 2\beta_1\beta_2 \frac{1}{2!} (t-s)^3 e^{-2(t-s)} ds, \end{aligned} \tag{3.51}$$

and taking the limit to $t \to \infty$ and using Eqn. (3.20) we derive

$$K_{33} = \sigma^2 \beta_1^2 \frac{1}{2^2} + \sigma^2 \beta_2^2 \frac{3}{2^4} + \sigma^2 \beta_1 \beta_2 \frac{3}{2^3}. \tag{3.52}$$

Equivalently expressed with the weights a_1, a_1 and a_3

$$K_{33} = \sigma^2 a_3^2 \frac{1}{2^2} + \sigma^2 (a_1 a_2)^2 \frac{3}{2^4} + \sigma^2 a_1 a_2 a_3 \frac{3}{2^3}. \tag{3.53}$$

The ratio now R can be expressed as

$$R = \beta_1^2 \frac{1}{2} + \beta_2^2 \frac{3}{2^3} + \beta_1 \beta_2 \frac{3}{2^2}. \tag{3.54}$$

If all edges are positive then β_1, $\beta_2 > 0$, then all paths in the network are positive and the target receives a consistent signal from the source. In this case the ffl is coherent. However, if $\beta_1 < 0$ or $\beta_2 < 0$, which occurs if either one or three of the edges is negative, then the target receives a inconsistent signal from the source. In this case, the ffl is incoherent. Denoting the noise processing ratios in the coherent and incoherent cases by R^+ and R^- respectively, it follows from Eqn. (3.54) that $R^+ > R^-$, and therefore

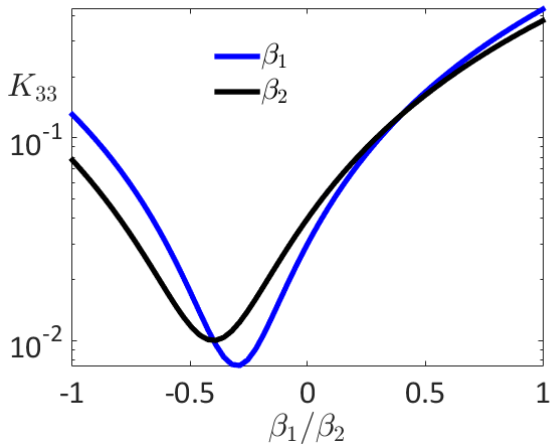


Figure 3.10: The log of the variance of the target node against the direct β_1 (blue) and indirect β_2 (black) weighted paths. A positive coherent fll is illustrated by the figure in the positive x axes, while an incoherent for the negative values of β_1 and β_2. The figure is produced for a fixed values of β_2 (blue) or β_1 (black) to 0.4.

that the incoherent feedback loop has the ability to suppress the noise in the target node.

In Fig. 3.9 the stochastic fll is plotted using the Euler-Maruyama approximation for 10^4 realisations. The noise from the source has strength $\sigma = 1$. Each line represents the expression of each species for final time $T = 200$. In order to show the variation in the number of molecules in the target, Fig. 3.9 shows the mean of the target that is represented by the black line. In Fig. 3.10, we plot the variance of the target K_{33} against the weighted paths β_1 and β_2. We fix one of the paths β_1 or β_2 and plot the variance against the other. The changes in the contribution of the noise in the variance occur when the weighted products take the same value, here at 0.5. Therefore, we conclude that in this particular magnitude for the weighted path of length 1 $|\beta_1| < 0.5$, the path of length 2 contributes more towards the variance K_{33}, while β_1 processes more noise to the target for $|\beta_2| > 0.5$. According to our formula in Eqn. (3.52), this is explained by the length of the direct path from node 1 to 3, with contribution $\beta_1^2/4$, since the coefficient is larger than the one of the indirect path. Inhibition in one of the paths leads to reduction of the magnitude in the variance of the target. This result is explained by the combination of the paths $\beta_1\beta_2$, as described in Eqn. (3.52). This result is not only in this special case but in general; the incoherent paths in the network tend to play the most important role in noise processing.

Now, we consider the weights i.i.d., where $a_i \sim \mathcal{N}(0, \gamma^2)$. In the next Fig. 3.11 is shown how the variance K_{33} changes with the variance γ^2 of the random normally distributed weights a_2 and a_3. Here the value for γ varied from 0.001 to 2. It seems that after a value that γ exceeds 1, variance is amplified. The probability density in

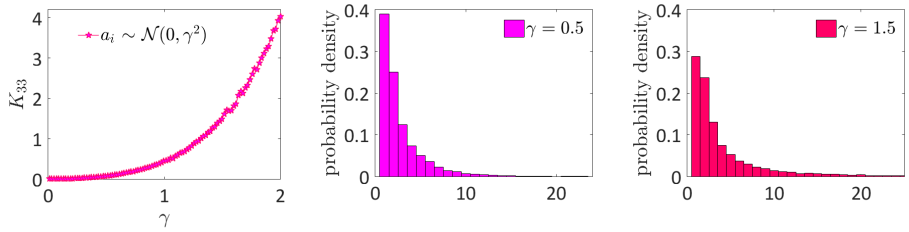


Figure 3.11: The left figure illustrates the variance of the target K_{nn} of ffl, given in Eqn. (3.53), against the square mean γ of the normally i.i.d. random variables a_1, a_2, a_3, so that $a_i \sim \mathcal{N}(0, \gamma^2)$. The following figures on the right, illustrate the probability density of K_{nn} for a fixed value of $\gamma = 0.5$ and $\gamma = 1.5$, respectively. In these central and right figures y axis depicts number of molecules.

Fig. 3.11 shows a higher probability the variance will be close to zero for $\gamma = 0.5$, than in the right for $\gamma = 1.5$.

3.4 Application to Transcription Regulatory Network

In this Section we consider the pluripotent network as defined and simplified in [8], and demonstrated in Fig. 3.12. The underlying interactions have been selected according to a Pearson correlation threshold, through which the assigned interacting factors are highly expressed [8]. The behaviour of this network is determined by the input stimulation of three extra-cellular factors commonly added to ES cells in culture media preparations: the cytokine Leukemia inhibitory factor (Lif), and selective inhibitors of glycogen synthase kinase 3 (TGF-β) (Chiron99021, denoted CH) and mitogen-activated protein kinase (Mek) (PD0325901, denoted PD) [8]. In this thesis, in order to explore the relationship between network structure and noise transduction from the noisy source Lif to the target Oct4, we modify the architecture regarding specific interactions. In particular, the role of distinguished interactions in noise propagation is examined in greater detail. The results are explained by using our mathematical expression and the numerical simulations.

We consider the noise to be arising from a molecule in the membrane of the cell, and processing through a signalling network to the output in the nucleus. As the noisy source we choose Lif that is essential for the pluripotent cells to remain undifferentiated *in vitro*. The protein encoded by the Lif gene binds onto the receptor of the signal transducer and activator of transcription 3 (Stat3) [138, 139], and activates signalling pathways that induce the transcriptional regulatory network [62, 69]. We select as the target the octamer DNA binding transcription factor 4 (Oct4), which in coordination with the factors Sox2 and Nanog interact accordingly to maintain the pluripotent circuit [68, 140, 141]. The underlying network is illustrated in Fig. 3.12 B, which is

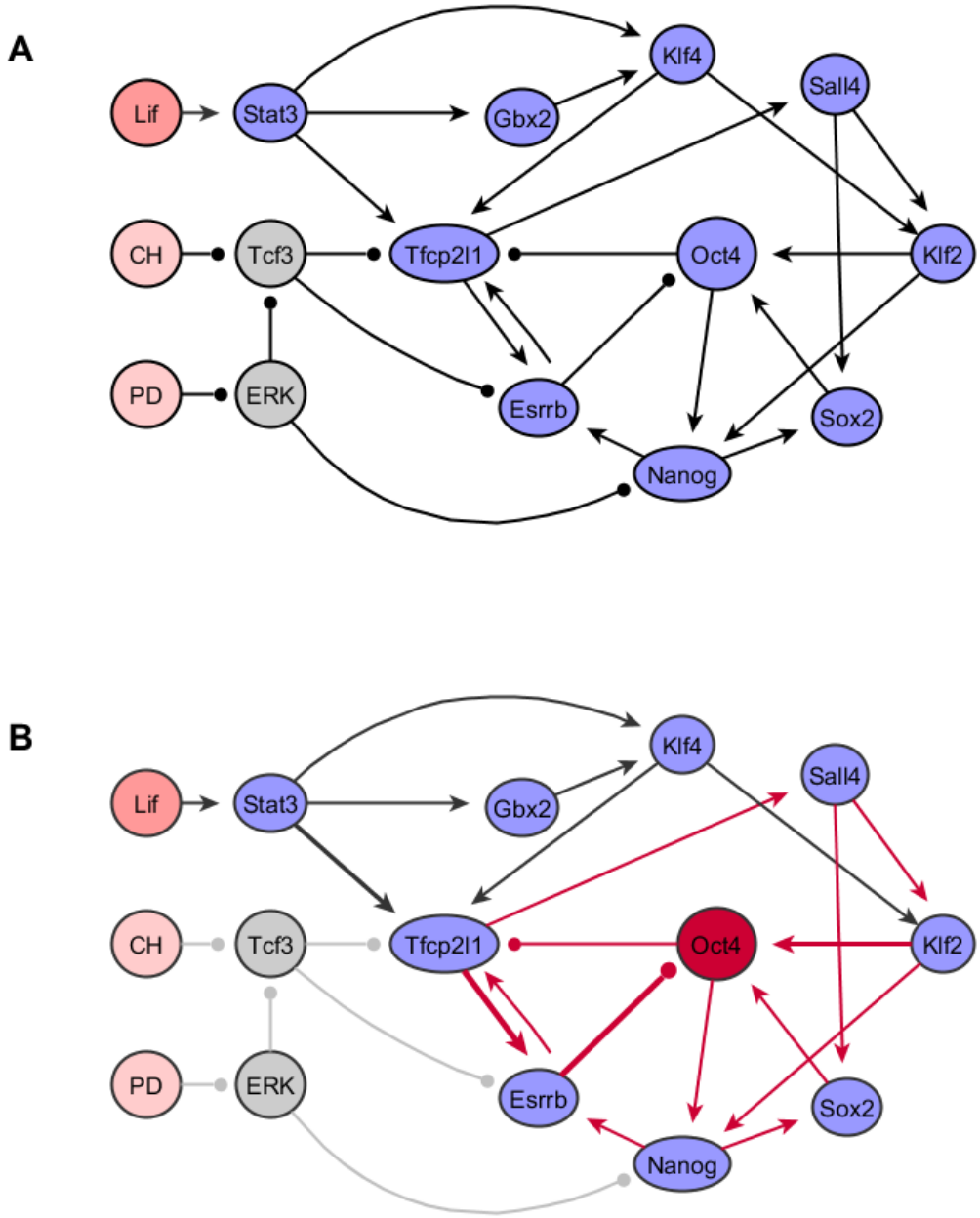


Figure 3.12: The graph of the simplified regulatory network [8] at the top, and on the bottom the version with LIF as the noisy source and the target Oct4. The red edges denote the interactions that give feedback at the network. The arrows activation of a factor, while circles indicate inhibition. The colour on the edges, ranged from light grey to black, along with the thickness indicates the importance of the related interaction. The variance of Oct4 takes it minimum value, when noise is processing from one source, Lif (Table B.1).

composed of 11 factors and 20 interactions. There are 26 paths from Lif to Oct4, including 3 cyclic paths.

	Variance of genes			
Genes	Noise processing R	All possible paths from Lif	Length of shortest path	Number of cycles/feedback
Stat3	0.0112	1	1	0
Tfcp2l1	$2.5450e-04$	13	2	9
Klf4	$2.4086e-04$	2	2	0
Gbx2	$1.8984e-04$	1	2	0
Klf2	$5.8979e-06$	5	3	5
Esrrb	$5.0270e-06$	20	3	7
Sall4	$4.8499e-06$	13	3	7
Sox2	$1.2849e-07$	24	4	4
Nanog	$1.2431e-07$	16	4	7
Oct4	$5.9950e-09$	23	4	9

Table 3.1: The normalized variance of each gene by the variance of Lif, is denoted by the ratio R, in the second column. The third column shows the number of all possible paths from Lif to each node; the forth column indicates the length of the shortest path each node participates; the fifth column identifies the number of cycles/feedback loops that each node take part in the network. Factors are ordered by decreasing order of the magnitude of R (column 2).

Although the sign of the interactions (*i.e.* whether they are activatory or inhibitory) is defined [8] their strength is unknown. In the absence of this information we assume that all interactions are of equal unit strength since this represents the most economic model. We consider the weights to magnitude a small enough value $a < 1/6$, such that the stability condition in Eqn. (3.25) is satisfied. Then, the normalised variance $R = K_{factor}/K_{Lif}$ by the source Lif of each node is calculated and demonstrated in the Table 3.1. Additionally, in the same table we count all the possible paths from Lif to each factor. The transcription factors with the least noise are the ones that are in the core of the nucleus of the transcriptional circuit in ES cells; Oct4, Nanog and Sox2. The more connected factors with Lif are Sox2, Oct4, Esrrb and Nanog which are all involved into cyclic paths with Oct4. In the last column of Table 3.1 are listed the number of cycles (feedback) in the network; Tfcp2l1 and Oct4 are engaged into 9 out of 11 cycles; Tfcp2l1 into 3 negative and Oct4 into 4 negative cycles.

There are particular edges that are responsible for the development of feedback in the network (see red lines in Fig. 3.12). In order for the flow to follow the same direction, we construct the corresponding acyclic graph by changing the orientation of three interactions: Oct4 $\rightarrow$ Nanog, Oct4 $\dashv$ Tfcp2l1 and Esrrb $\rightarrow$ Tfcp2l1. The new acyclic version includes 19 interactions and 29 paths from Lif to Oct4. To compare the effect of no-cycles in the graph we consider the mean path length [9, 12, 142], as a

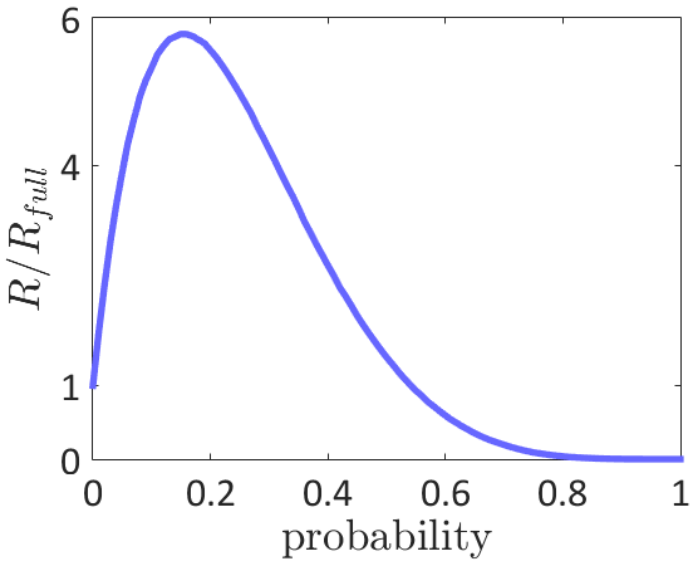


Figure 3.13: The figure shows the ratio R/R_{full} for Oct4 against the probability of removing each interaction, for 10^6 realisations.

measure of 'navigability' in the graph. The smaller is the magnitude of the mean path length, the faster is the flow of noise in the signalling network. The mean path length for the unperturbed (denoted by $\langle l_{\text{org}} \rangle$) and the acyclic (denoted by $\langle l_{\text{acyc}} \rangle$) graph is calculated as $\langle l_{\text{org}} \rangle = 0.0576$ and $\langle l_{\text{acyc}} \rangle = 0.0171$, respectively. This result shows that the acyclic version has the shortest paths from Lif to Oct4. Therefore, the noise in the acyclic graph is mitigating faster than the original. There are also formed four extra feed-forward loops in the acyclic case, illustrated in Fig. A.1. One conclusion can be summarised from this outcome: the pluripotent ES cells have a preference to interact in a distinctive way that form circuits and feedback loops in order to minimize the noise.

Results

To investigate the effect of network structure on noise processing, we look how the ratio R (Eqn. (3.22)) changes upon targeted removal of different interactions, compared to the unperturbed network, denoted by R_{full}. In order to visualise how the magnitude of the variance of Oct4 is affected by the structure of the network, we introduce the probability p that an interaction is removed. Fig. 3.13 illustrates the evolution of Oct4 ratio R/R_{full} against the probability p that an edge is removed. Interestingly, we note here that the maximum in Fig. 3.13, at probability $p = 0.15$ corresponds to the removal of an average of three interactions. When all the edges removed simultaneously, the network is disconnected, and this explains the zero in the magnitude of R/R_{full} for probability one.

Interaction removed	Noise-processing (R/R_{full})	Coherence (c/c_{full})	Feedback (f/f_{full})
Tfcp2l1 → Sall4	0.02	0.68	0.64
Sox2 → Oct4	0.12	0.60	0.37
Sall4 → Klf2	0.18	0.98	0.45
Sall4 → Sox2	0.18	0.86	0.18
Stat3 → Gbx2	0.32	0.98	0
Gbx2 → Klf4	0.32	0.98	0
Nanog → Sox2	0.83	0.94	0.18
Klf2 → Nanog	1	1.10	0.27
Esrrb → Tfcp2l1	1.12	0.95	0.36
Nanog → Esrrb	1.19	1.25	0.45
Klf4 → Tfcp2l1	1.99	1.06	0
Stat3 → Klf4	4.70	0.98	0
Klf4 → Klf2	8.68	0.92	0
Klf2 → Oct4	12.10	0.86	0.18
Stat3 → Tfcp2l1	12.18	1.03	0
Tfcp2l1 → Esrrb	24.62	1.23	0.18
Esrrb ⊣ Oct4	25.68	1.53	0.27

Table 3.2: The effect of the removal of interactions on network's noise processing. The first column identifies the edge removed from the network; the second column shows the effect of removal of the given edge on the ratio R by comparison with that of the unperturbed network R_{full}; the third column shows the effect of the removal of the given edge on network coherence; the fourth column shows the effect of the removal of the given edge on network feedback. Edges that emanate from Oct4 do not contribute to the noise processing capacity of the network and their removal does not affect R so they are excluded from this table. Since all paths from Lif to Oct4 pass through the edge Lif → Stat3 its removal disconnects the network; this edge is also accordingly excluded from the table. Interactions are ordered by column 2.

To reveal the structural influence on noise processing we also calculated two simple network measures based upon our interpretation of Eqn. (3.22): (1) $c = p^+/(p^+ + p^-)$, where p^+ and p^- are the number of positive and negative paths from Lif to Oct4 respectively. Therefore, c gives a measure of coherence in the network, and compared with values of the coherence in the unperturbed network, denoted by c_{full}; and (2) f, is the total number of feedback loops in the network, as a measure of network complexity, and compared with the unperturbed network, denoted by f_{full}. These results along with the modification in the magnitude of the variance of Oct4 for removing an interaction are demonstrating in detail in Table 3.2.

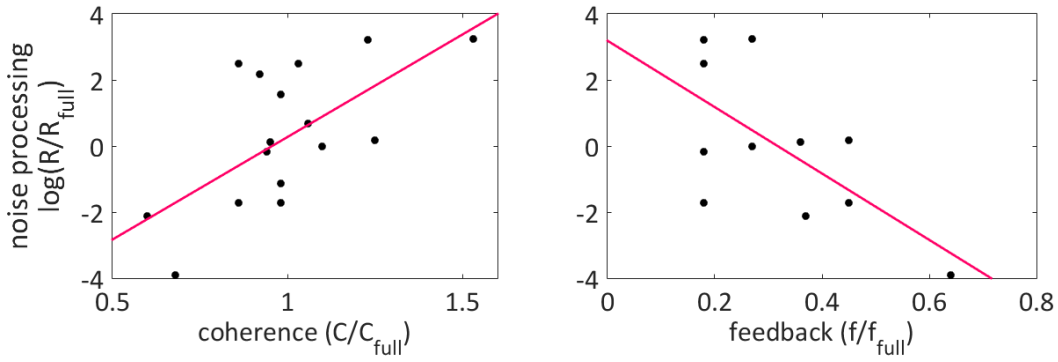


Figure 3.14: Plots of the data from Table 3.2. Removal of edges that result in an increase of coherence in the network tend to diminish the system's noise processing ability, while removal of edges which reduce the overall feedback structure of the network tend to improve the system's noise-processing ability. Red lines show linear regression. In the right figure interactions with no contributions to feedback are excluded.

The shortest paths from Lif to Oct4 in this network have length 4; there are two such paths: (1) Lif $\rightarrow$ Stat3 $\rightarrow$ Tfcp2l1 $\rightarrow$ Esrrb $\dashv$ Oct4; and (2) Lif $\rightarrow$ Stat3 $\rightarrow$ Klf4 $\rightarrow$ Klf2 $\rightarrow$ Oct4. The first of these paths in negative due to the inhibitory interaction Esrrb $\dashv$ Oct4, while the second is positive. If we consider together this pair of paths forms an incoherent feed-forward loop; since these are the shortest paths in the network, we anticipate from Eqn. (3.22) that this incoherent feed-forward loop will have an important role in noise-processing in this network. Indeed, according to Table 3.2, if any of the elements of these paths are removed, then the noise-processing capacity of this network is severely inhibited and the ratio R increases substantially (Fig. 3.14).

It is well known that Tfcp2l1, Sall4 and Nanog play a critical role since their existence reinforces the pluripotent state in ES cells [34, 72]. Although the interaction Tfcp2l1 $\rightarrow$ Sall4 has the greater effect in the noise amplification, Tfcp2l1 is a typical target induced by the source Lif, where its absence jeopardises the receptivity from the protein Lif [88]. Conversely, the presence of the inhibitory interaction Esrrb $\dashv$ Oct4 shows a substantial suppression in the Oct4 variance. This interaction takes part in three cycles with Esrrb, in two activated by Nanog; furthermore, Esrrb is stimulated by the core factor Tfcp2l1 in order to form a circuit with Oct4, and is activated directly by Nanog, sustaining the pluripotent state [143, 144]. Therefore, the interaction is connected to the main body factors and a failure of this suppression induces greater noise in the target Oct4. These two representative examples imply that particular interactions existence tends to balance extreme noise amplifications; while core interactions such as Tfcp2l1 $\rightarrow$ Sall4 and Sox2 $\rightarrow$ Oct4 have a high aggregated effect in noise processing, others such as Esrrb $\dashv$ Oct4 and Tfcp2l1 $\rightarrow$ Esrrb influence

to harmonise this effect.

Discussion and Conclusions

In this Chapter, we focused on the signalling network, in which molecular factors govern the pluripotent state in mouse ES cells [11, 34, 61–63, 69, 70]. Under the assumption of stability and that the extrinsic environmental noise is coming from a single input, we derived a mathematical formula that allows us to relate the structure and the noise processing ability of the network; this is achieved by calculating the variance of any node, depending on the weighted walks connecting that node to the noisy input. Although, the conclusions can apply to any network with more noisy inputs and the results can be also expressed numerically, following the same methodology. The mathematical formula in Eqn. (3.20) for the variance at the equilibrium is a bridge between the topology of the underlying network and the noise processing from a noisy source(s).

We deduce that not only the number of components, but the length of the paths from the noisy source to the target, plays a critical role in the noise propagation. Collectively, results from the limiting variance in the examples of a signalling cascade and a feed-forward loop lead to the conclusion that the structure does determine the effect in noise processing. We report that under conditions such as fixed weight magnitude a, satisfying $|a| < 1$, the shortest pathways in a signalling cascade bear the most decisive role in the amplification of the ability of noise processing; therefore, the smaller the length of a path the greater is its contribution in the variance of the target. Additionally, incoherent feed-forward loops suppress noise processing. Interestingly, applying our formula in Eqn. (3.20) on the transcriptional pluripotent network we found that coherent feedback is influenced in principal by the core transcriptional factors essential for pluripotency maintenance [11, 145]. Nevertheless, other interactions between auxiliary factors, which form incoherent feed-forward loops, function as a balance in noise amplification, and ensure that environmental signals are robustly mediated in this core circuit.

As has been anticipated from Eqn. (3.20), positive cycles and feed-forward loops add extra positive terms to the variance of the target node, and then the sum in Eqn. (3.24) can be infinite; that affects the target receiving constantly infinite noise. However, negative feedback adds both positive and negative terms to this sum, while may amplify or diminish noise processing, depending on the distinct organisation of

inhibitory interactions in the network. We conclude that the special structure in pluripotent ES cells tries to maximize the noise at an early level of interactions close to the noisy source Lif. Nevertheless, this topology enhances the suppression of the noise as interactions approach the transcriptional core and hence the target Oct4 (see Table. 3.2, second column).

Generally, Eqn. (3.20) can be applied to any factor in the network shown in Fig. 3.12, with the intention to examine and determine their variability with the noisy source(s). In order to find the effect of noise from all the three sources in the pluripotent network, we implemented the calculations computationally and the outcome is presented in the Tables B.1–B.7. All the combinations of possible noisy inputs are taken into account; the variance of Oct4 minimises its magnitude, when noise is processing only from Lif, as indicated in Table B.1. This outcome occurs despite the fact that the other noisy inputs, CH and PD, have a suppressing role in the graph.

Appendix A

Motifs in the pluripotent network

Feed-forward loops

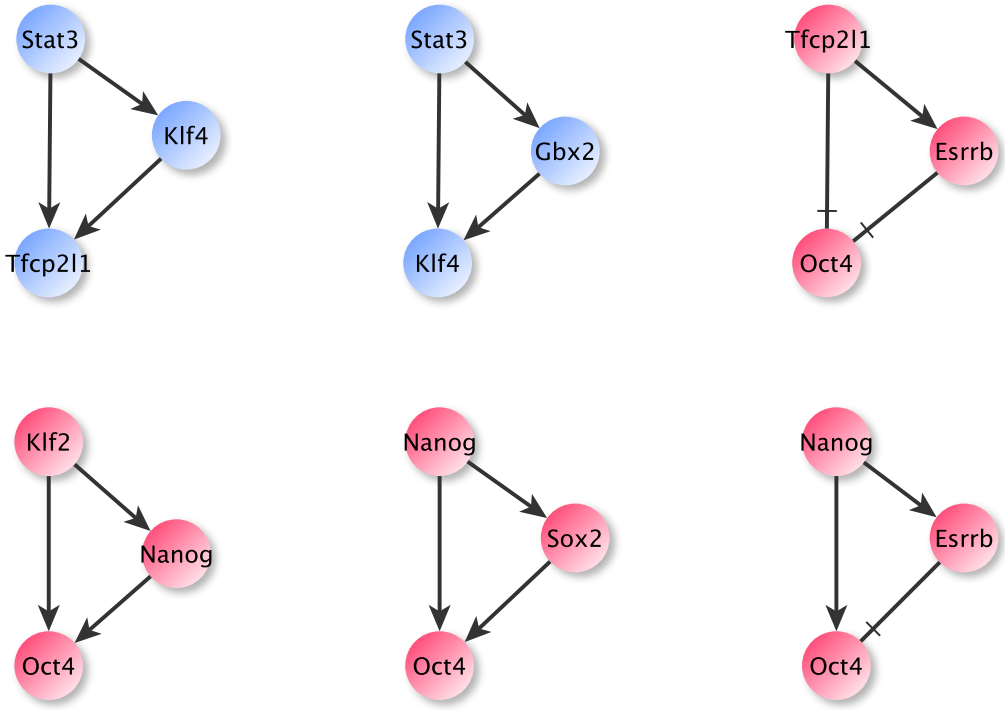


Figure A.1: The first two blue ffl are in both graphs, while the other four red are added in the acyclic.

Shortest-paths

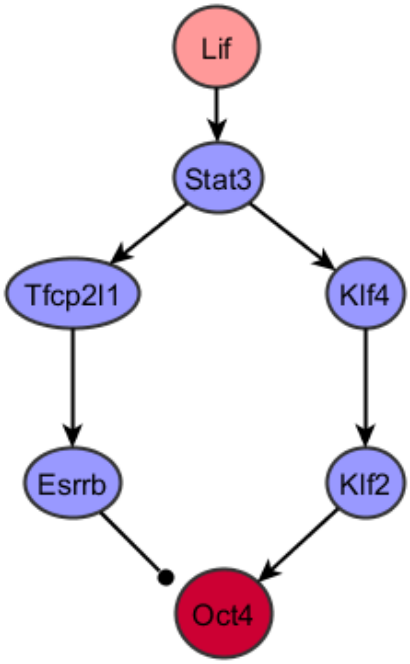


Figure A.2: The shortest paths from Lif to Oct4 in the network (1) Lif → Stat3 → Tfcp2l1 → Esrrb ⊣ Oct4 and (2) Lif → Stat3 → Klf4 → Klf2 → Oct4; together form an incoherent feed-forward loop.

Appendix B

Tables

Variance Oct4-different noisy sources

Variance of genes			
Noisy sources	Ratio R Oct4	Length of shortest path	Sign of the shortest path
CH	$4.4306e-06$	3	$-$
Lif & CH	$4.1756e-06$	3 from CH	$-$
Lif & PH	$4.1756e-06$	4 from Lif / PD	$-$ / $+$
CH & PD	$3.3596e-06$	3 from CH	$-$
Lif, CH & PD	$3.1477e-06$	3 from CH	$-$
PD	$8.5537e-08$	4	$+$
Lif	$5.9950e-09$	4	$-$

Table B.1: The normalized variance of Oct4, by the variance of Lif, CH and PD, in all different combinations of the noisy source(s).

Three noisy sources

Variance of genes		
Genes	Ratio R	Length of shortest path
Stat3	0.0113	1 from Lif
ERK	0.0113	1 from PD
Tcf3	0.0089	1 from CH
Tfcp2l1	$9.0682e-04$	2 from Lif / CH
Esrrb	$3.0991e-04$	2 from CH
Klf4	$2.4086e-04$	2 from Lif
Nanog	$1.9184e-04$	2 from PD
Gbx2	$1.8984e-04$	2 from Lif
Sall4	$1.7204e-05$	3 from Lif / CH
Klf2	$7.2470e-06$	3 from Lif
Sox2	$5.9985e-06$	3 from PD
Oct4	$3.1477e-06$	3 from CH

Table B.2: The normalized variance of each gene, by the variance of Lif, CH and PD.

Noisy source CH

Variance of genes		
Genes	Ratio R	Length of shortest path
Tcf3	0.0113	1
Esrrb	$2.4806e-04$	2
Tfcp2l1	$2.5513e-04$	2
Oct4	$4.4306e-06$	3
Sall4	$4.8638e-06$	3
Nanog	$6.4816e-08$	4
Sox2	$7.6209e-08$	4
Klf2	$9.6749e-08$	4

Table B.3: The normalized variance of each gene, by the variance of the source CH.

Noisy source PD

Variance of genes		
Genes	Ratio R	Length of shortest path
PD	1	-
ERK	0.0113	1
Nanog	$1.9039e-04$	2
Tcf3	$1.8984e-04$	2
Tfcp2l1	$3.8019e-06$	3
Sox2	$3.4581e-06$	3
Sall4	$7.5199e-08$	4
Esrrb	$7.2805e-08$	3
Oct4	$8.5537e-08$	4
Klf2	$1.5278e-09$	5

Table B.4: The normalized variance of each gene, by the variance of PD.

Noisy sources Lif and CH

Variance of genes		
Genes	Ratio R	Length of shortest path
Stat3	0.0113	1 from Lif
Tcf3	0.0113	1 from CH
Tfcp2l1	0.0010	2 from Lif / CH
Sox2	$3.1751e-04$	4 from Lif / CH
Klf4	$2.4086e-04$	2 from Lif
Esrrb	$1.8984e-04$	2 from CH
Gbx2	$7.4267e-06$	2 from Lif
Oct4	$4.1756e-06$	3 from CH
Nanog	$1.9427e-05$	4 from Lif
Sall4	$4.0203e-07$	3 from Lif / CH
Klf2	$1.1046e-08$	3 from Lif

Table B.5: The normalized variance of each gene, by the variance of Lif and CH.

Noisy sources Lif and PD

Variance of genes		
Genes	Ratio R	Length of shortest path
Stat3	0.0113	1 from Lif
ERK	0.0113	1 from PD
Tfcp2l1	$0,0011$	2 from Lif
Sox2	$3.1751e-04$	3 from PD
Klf4	$2.4086e-04$	2 from Lif
Esrrb	$1.8984e-04$	3 from Lif / PD
Nanog	$1,9427e-05$	2 from PD
Gbx2	$7.4267e-06$	2 from Lif
Sall4	$4.0203e-07$	3 from Lif
Klf2	$1.1046e-08$	3 from Lif
Oct4	$4.1756e-06$	4 from Lif / PD

Table B.6: The normalized variance of each gene, by the variance of Lif and PD.

Noisy sources CH and PD

Variance of genes		
Genes	Ratio R	Length of shortest path
ERK	0.0113	1 from PD
Tcf3	0.0089	1 from CH
Sox2	$2.4162e-04$	3 from PD
Tfcp2l1	$2.0076e-04$	2 from CH
Klf2	$1.8523e-04$	4 from CH
Sall4	$4.4816e-06$	3 from CH
Esrrb	$3.7884e-06$	2 from CH
Oct4	$3.3596e-06$	3 from CH
Nanog	$7.4903e-08$	2 from PD

Table B.7: The normalized variance of each gene, by the variance of CH and PD.

Bibliography

[1] S. Kontogeorgaki, R. Sanchez-Garcia, R. Ewing, K. Zygalakis, and B. D. MacArthur, "Noise-processing by signaling networks." biorxiv:10.1101/075366, Submitted in *Sci. Rep.*, 2016.

[2] C. O'Connor and J. U. Adams, "Essentials of cell biology.," *MA: NPG Education*, pp. 1–100, 2010.

[3] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai, "Lethality and centrality in protein networks.," *Nature*, vol. 411, no. 6833, pp. 41–42, 2001.

[4] G. D. Winter, "Three waves of innovation in vertebrate evolution.." http://www.science20.com/curious_cub/three_waves_innovation_vertebrate_evolution-81858, 2011.

[5] https://ebbailey.wordpress.com/general-information/dna-to-protein/.

[6] J. Wang, S. Rao, J. Chu, X. Shen, D. N. Levasseur, T. W. Theunissen, and S. H. Orkin, "A protein interaction network for pluripotency of embryonic stem cells.," *Nature*, vol. 444, no. 7117, pp. 364–368, 2006.

[7] J. Kim, J. Chu, X. Shen, J. Wang, and S. H. Orkin, "An extended transcriptional network for pluripotency of embryonic stem cells.," *Cell*, vol. 132, no. 6, pp. 1049–1061, 2008.

[8] S. J. Dunn, G. Martello, B. Yordanov, S. Emmott, and A. G. Smith, "Defining an essential transcription factor program for naïve pluripotency.," *Science*, vol. 344, no. 6188, pp. 1156–60, 2014.

[9] A. L. Barabási and Z. N. Oltvai, "Network biology: understanding the cell's functional organization.," *Nat. Rev. Genet.*, vol. 5, no. 2, pp. 101–113, 2004.

[10] B. J. Merrill, "Wnt pathway regulation of embryonic stem cell self-renewal.," *Cold Spring Harb. Perspect Biol.*, vol. 4, no. 9, 2012.

[11] B. D. Macarthur, A. Ma'ayan, and I. R. Lemischka, "Systems biology of stem cell fate and cellular reprogramming.," *Nat. Rev. Mol. Cell Biol.*, vol. 10, no. 10, pp. 672–681, 2009.

[12] G. Tkačik and W. Bialek, "Cell biology: Networks, regulation, pathways.," in *Encyclopedia of complexity and systems science.*, pp. 719–741, Berlin: Springer, 2009.

[13] Y. Sun, H. Li, H. Yang, M. S. Rao, and M. Zhan, "Mechanisms controlling embryonic stem cell self-renewal and differentiation.," *Crit. Rev. Eukaryot. Gene Expr.*, vol. 16, no. 3, pp. 211–232, 2006.

[14] B. D. MacArthur, C. P. Please, and R. O. C. Oreffo, "Stochasticity and the molecular mechanisms of induced pluripotency.," *PLoS One*, vol. 3, no. 8, pp. 1–11, 2008.

[15] J. E. Ladbury and S. T. Arold, "Noise in cellular signaling pathways: causes and effects.," *Trends Biochem. Sci.*, vol. 37, no. 5, pp. 173–178, 2012.

[16] M. Buchanan, *Networks in cell biology.* CUP, 2010.

[17] H. Niwa, "Wnt: What's needed to maintain pluripotency?," *Nat. Cell Biol.*, vol. 13, no. 9, pp. 1024–1026, 2011.

[18] H. Kim, J. Wu, S. Ye, C. Tai, X. Zhou, H. Yan, P. Li, M. Pera, and Q. Ying, "Modulation of β-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal.," *Nat. Commun.*, vol. 4, p. 2403, 2013.

[19] L. Attisano and J. L. Wrana, "Signal integration in tgf-β, wnt, and hippo pathways.," *F1000Prime Rep.*, vol. 5, 2013.

[20] A. Niida, T. Hiroko, M. Kasai, Y. Furukawa, Y. Nakamura, Y. Suzuki, S. Sugano, and T. Akiyama, "Dkk1, a negative regulator of wnt signaling, is a target of the beta-catenin/tcf pathway.," *Oncogene*, vol. 23, no. 52, pp. 8520–8526, 2004.

[21] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, vol. 10. Chapman HallCRC mathematical and computational biology series, 2007.

[22] R. Lu, F. Markowetz, R. D. Unwin, J. T. Leek, E. M. Airoldi, B. D. MacArthur, A. Lachmann, R. Rozov, A. Ma'ayan, L. A. Boyer, O. G. Troyanskaya, A. D. Whetton, and I. R. Lemischka, "Systems-level dynamic analyses of fate change in murine embryonic stem cells.," *Nature*, vol. 462, no. 7271, pp. 358–62, 2009.

[23] J. M. Pedraza and A. van Oudenaarden, "Noise propagation in gene networks.," *Science*, vol. 307, no. 5717, pp. 1965–1969, 2005.

[24] J. Paulsson, "Summing up the noise in gene networks.," *Nature*, vol. 427, no. 6973, pp. 415–418, 2004.

[25] G. Hornung and N. Barkai, "Noise propagation and signaling sensitivity in biological networks: A role for positive feedback.," *PLoS Comput. Biol.*, vol. 4, no. 1, p. e8, 2008.

[26] S. Hooshangi and R. Weiss, "The effect of negative feedback on noise propagation in transcriptional gene networks.," *Chaos*, vol. 16, no. 2, p. 026108, 2006.

[27] D. T. Gillespie, "A rigorous derivation of the chemical master equation.," *Phys. A Stat. Mech. its Appl.*, vol. 188, no. 1-3, pp. 404–425, 1992.

[28] D. J. Higham, "Modeling and simulating chemical reactions.," *SIAM*, vol. 50, no. 2, pp. 347–368, 2008.

[29] R. Erban, S. J. Chapman, and K. P. Maini, "A practical guide to stochastic simulations of reaction-diffusion processes.," p. 35, 2007.

[30] M. E. J. Newman, *Networks. An introduction.* OUP, 2010.

[31] C. Bachmaier, U. Brandes, and F. Schreiber, "Biological networks.," *Handbook of Graph Drawing and Visualization*, pp. 621–651, 2013.

[32] X. Zhu, M. Gerstein, and M. Snyder, "Getting connected: analysis and principles of biological networks.," vol. 21, no. 9, pp. 1010–1024, 2007.

[33] E. Estrada, *The structure of complex networks.* OUP, 2011.

[34] G. Martello and A. Smith, "The nature of embryonic stem cells.," *Annu Rev Cell Dev Biol*, vol. 30, pp. 647–675, 2014.

[35] A. Blais and B. D. Dynlacht, "Constructing transcriptional regulatory networks.," *Genes Dev.*, vol. 19, no. 13, pp. 1499–1511, 2005.

[36] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular biology of the cell.* Garland science, Taylor & Francis Group, LLC, 2002.

[37] D. Bray, "Protein molecules as computational elements in living cells.," vol. 376, no. 6538, pp. 307–312, 1995.

[38] A. Zhang, *Protein interaction networks: computational analysis.* CUP, 2009.

[39] J. F. Rual, K. Venkatesan, T. Hao, T. Hirozane-Kishikawa, A. Dricot, N. Li, G. F. Berriz, F. D. Gibbons, M. Dreze, N. Ayivi-Guedehoussou, N. Klitgord, C. Simon, M. Boxem, S. Milstein, J. Rosenberg, D. S. Goldberg, L. V. Zhang, S. L. Wong, G. Franklin, S. Li, J. S. Albala, J. Lim, C. Fraughton, E. Llamosas,

S. Cevik, C. Bex, P. Lamesch, R. S. Sikorski, J. Vandenhaute, H. Y. Zoghbi, A. Smolyar, S. Bosak, R. Sequerra, L. Doucette-Stamm, M. E. Cusick, D. E. Hill, F. P. Roth, and M. Vidal, "Towards a proteome-scale map of the human protein-protein interaction network.," *Nature*, vol. 437, no. October, pp. 1173–1178, 2005.

[40] U. Stelzl, U. Worm, M. Lalowski, C. Haenig, F. H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koeppen, J. Timm, S. Mintzlaff, C. Abraham, N. Bock, S. Kietzmann, A. Goedde, E. Toksöz, A. Droege, S. Krobitsch, B. Korn, W. Birchmeier, H. Lehrach, and E. E. Wanker, "A human protein-protein interaction network: a resource for annotating the proteome.," *Cell*, vol. 122, no. 6, pp. 957–968, 2005.

[41] C. Von Mering, R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields, and P. Bork, "Comparative assessment of large-scale data sets of protein-protein interactions," *Nature*, vol. 417, pp. 399–403, 5 2002.

[42] A. Wagner, "Robustness against mutations in genetic networks of yeast.," *Nat. Genet.*, vol. 24, no. 4, pp. 355–61, 2000.

[43] L. Hunter, "Artificial intelligence and molecular biology," pp. 1–46, 1993.

[44] J. D. Jordan, E. M. Landau, and R. Iyengar, "Signaling networks: the origins of cellular multitasking.," *Cell*, vol. 103, no. 2, pp. 193–200, 2000.

[45] N. J. Eungdamrong and R. Iyengar, "Modeling cell signaling networks.," *Changes*, vol. 29, no. 6, pp. 997–1003, 2012.

[46] E. M. Adler, N. R. Gough, and L. B. Ray, "2015: Signaling breakthroughs of the year.," *Sci. Signal.*, vol. 2016, no. 409, pp. eg1–eg1, 2016.

[47] J. M. Irish, R. Hovland, P. O. Krutzik, O. D. Perez, O. Bruserud, B. T. Gjertsen, and G. P. Nolan, "Single cell profiling of potentiated phospho-protein networks in cancer cells.," *Cell*, vol. 118, no. 2, pp. 217–228, 2015.

[48] F. Verkaar, K. M. Cadigan, and V. R. Amerongen, "Celebrating 30 years of wnt signaling.," *Science Signaling*, vol. 5, no. 254, p. mr2, 2012.

[49] E. Levine and T. Hwa, "Stochastic fluctuations in metabolic pathways.," *PNAS*, vol. 104, no. 22, pp. 9224–9229, 2007.

[50] S. L. Schreiber and B. E. Bernstein, "Signaling network model of chromatin.," *Cell*, vol. 111, no. 6, pp. 771–778, 2015.

[51] J. Song, Z. Du, M. Ravasz, B. Dong, Z. Wang, and R. M. Ewing, "A protein interaction between β-catenin and dnmt1 regulates wnt signaling and dna methylation in colorectal cancer cells.," *Mol. Cancer Res.*, vol. 13, no. 6, pp. 969–981, 2015.

[52] T. Valenta, G. Hausmann, and K. Basler, "The many faces and functions of β-catenin.," *EMBO J.*, vol. 31, no. 12, pp. 2714–36, 2012.

[53] S. A. Vlahopoulos, O. Cen, N. Hengen, J. Agan, M. Moschovi, E. Critselis, M. Adamaki, F. Bacopoulou, J. A. Copland, I. Boldogh, M. Karin, and G. P. Chrousos, "Dynamic aberrant NF-κB spurs tumorigenesis: a new model encompassing the microenvironment.," *Cytokine Growth Factor Rev.*, vol. 26, no. 4, pp. 389–403, 2015.

[54] H. Clevers, K. M. Loh, and R. Nusse, "Stem cell signaling. an integral program for tissue renewal and regeneration: Wnt signaling and stem cell control.," *Science*, vol. 346, no. 6205, p. 1248012, 2014.

[55] M. van den Heuvel, *Signal transduction pathways in development: hedgehog proteins and their receptors.* John Wiley & Sons, Ltd, 2001.

[56] L. Attisano and J. L. Wrana, "Signal integration in TGF-β, WNT, and Hippo pathways.," *F1000Prime Rep.*, vol. 5, no. June, p. 17, 2013.

[57] J. E. Ladbury and S. T. Arold, "Noise in cellular signaling pathways: causes and effects.," *Trends Biochem. Sci.*, vol. 37, no. 5, pp. 173–178, 2013.

[58] B. K. Brott and S. Y. Sokol, "Regulation of Wnt/LRP signaling by distinct domains of Dickkopf proteins.," *Mol. Cell. Biol.*, vol. 22, no. 17, pp. 6100–10, 2002.

[59] Bryan T. M. and Keiko T. and Xi He, "Wnt/β-catenin signaling: components, mechanisms, and diseases.," *Dev. Biol.*, vol. 17, no. 1, pp. 9–26, 2010.

[60] J. J. Tyson, K. C. Chen, and B. Novak, "Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell.," *Curr. Opin. Cell Biol.*, vol. 15, no. 2, pp. 221–231, 2003.

[61] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.," *Cell*, vol. 126, no. 4, pp. 663–676, 2006.

[62] G. Huang, S. Ye, X. Zhou, D. Liu, and Q. L. Ying, "Molecular basis of embryonic stem cell self-renewal: From signaling pathways to pluripotency network.," *Cell. Mol. Life Sci.*, vol. 72, no. 9, pp. 1741–1757, 2015.

[63] I. Glauche, M. Herberg, and I. Roeder, "Nanog variability and pluripotency regulation of embryonic stem cells - Insights from a mathematical model analysis.," *PLoS One*, vol. 5, no. 6, pp. 1–12, 2010.

[64] H. Niwa, "How is pluripotency determined and maintained?," *Development (Cambridge, England)*, vol. 134, no. 4, pp. 635–46, 2007.

[65] A. Blais and B. D. Dynlacht, "Constructing transcriptional regulatory networks.," *Genes Dev.*, no. 212, pp. 1499–1511, 2005.

[66] H. Ng and M. A. Surani, "The transcriptional and signalling networks of pluripotency.," *Nat. Cell Biol.*, vol. 13, no. 5, pp. 490–496, 2011.

[67] M. J. Evans and M. H. Kaufman, "Establishment in culture of pluripotential cells from mouse embryos.," vol. 292, pp. 154–156, 1981.

[68] L. A. Boyer, T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, and R. A. Young, "Core transcriptional regulatory circuitry in human embryonic stem cells.," vol. 122, 2005.

[69] J. Yu and J. Thomson, "Pluripotent stem cell lines.," *Genes Dev.*, pp. 1987–1997, 2008.

[70] L. Dahéron, S. L. Opitz, H. Zaehres, M. W. Lensch, P. W. Andrews, J. Itskovitz-Eldor, and G. Q. Daley, "LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells.," *Stem Cells*, vol. 22, no. 5, pp. 770–778, 2004.

[71] J. Yu, M. A. Vodyanik, K. Smuga-Otto, J. Antosiewicz-Bourget, J. L. Frane, S. Tian, J. Nie, G. A. Jonsdottir, V. Ruotti, R. Stewart, *et al.*, "Induced pluripotent stem cell lines derived from human somatic cells.," *Science*, vol. 318, no. 5858, pp. 1917–1920, 2007.

[72] S. Ye, P. Li, C. Tong, and Q. Ying, "Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1.," *EMBO J.*, vol. 32, no. 19, pp. 2548–60, 2013.

[73] P.-Y. Bourillot and P. Savatier, "Krüppel-like transcription factors and control of pluripotency.," *BMC Biol.*, vol. 8, p. 125, 2010.

[74] K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, and S. Yamanaka, "Induction of pluripotent stem cells from adult human fibroblasts by defined factors.," *Cell*, vol. 131, no. 5, pp. 861–872, 2007.

[75] Q.-L. Ying, J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, and A. Smith, "The ground state of embryonic stem cell self-renewal," *Nature*, vol. 453, no. 7194, pp. 519–523, 2008.

[76] P. S. Stumpf, R. Ewing, and B. D. MacArthur, "Single-cell pluripotency regulatory networks.," *Proteomics*, vol. 16, pp. 2303–2312, 2016.

[77] R. L. Williams, D. J. Hilton, and N. A. Nicolai, "Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.," *Nature*, vol. 336, p. 15, 1988.

[78] A. G. Smith, J. K. Heath, D. D. Donaldson, G. G. Wong, J. Moreau, M. Stahl, and D. Rogers, "Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.," *Nature*, vol. 336, no. 6200, pp. 688–690, 1988.

[79] Q.-L. Ying, J. Nichols, I. Chambers, and A. Smith, "Bmp induction of id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with stat3.," *Cell*, vol. 115, no. 3, pp. 281–292, 2003.

[80] T. Burdon, C. Stracey, I. Chambers, J. Nichols, and A. Smith, "Suppression of shp-2 and erk signalling promotes self-renewal of mouse embryonic stem cells.," *Dev. Biol.*, vol. 210, no. 1, pp. 30–43, 1999.

[81] M. P. Stavridis, J. S. Lunn, B. J. Collins, and K. G. Storey, "A discrete period of fgf-induced erk1/2 signalling is required for vertebrate neural specification.," *Development*, vol. 134, no. 16, pp. 2889–2894, 2007.

[82] T. Kunath, M. K. Saba-El-Leil, M. Almousailleakh, J. Wray, S. Meloche, and A. Smith, "Fgf stimulation of the erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment.," *Development*, vol. 134, no. 16, pp. 2895–2902, 2007.

[83] N. Sato, L. Meijer, L. Skaltsounis, P. Greengard, and A. H. Brivanlou, "Maintenance of pluripotency in human and mouse embryonic stem cells through activation of wnt signaling by a pharmacological gsk-3-specific inhibitor.," *Nat. Med.*, vol. 10, no. 1, pp. 55–63, 2004.

[84] M. Amit, M. K. Carpenter, M. S. Inokuma, C.-P. Chiu, C. P. Harris, M. A. Waknitz, J. Itskovitz-Eldor, and J. A. Thomson, "Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture.," *Dev. Biol.*, vol. 227, no. 2, pp. 271–278, 2000.

[85] R.-H. Xu, R. M. Peck, D. S. Li, X. Feng, T. Ludwig, and J. A. Thomson, "Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human es cells.," *Nat. Methods*, vol. 2, no. 3, pp. 185–190, 2005.

[86] Y. H. Loh, Q. Wu, J. L. Chew, V. B. Vega, W. Zhang, G. Chen, X.and Bourque, J. George, B. Leong, J. Liu, K. Y. Wong, K. W. Sung, C. W. H. Lee, X. Zhao, K. P. Chiu, L. Lipovich, V. Kuznetsov, P. Robson, L. W. Stanton, C. L. Wei,

Y. Ruan, B. Lim, and H. H. Ng, "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells.," *Nat. Genet.*, vol. 38, no. 4, pp. 431–440, 2006.

[87] Liping N., A. E. Vázquez and E. N. Yamoah, "Identification of transcription factor DNA interactions using chromatin immunoprecipitation assays.," *Methods Mol Biol.*, vol. 493, p. 311, 2009.

[88] G. Martello, P. Bertone, and A. Smith, "Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor.," *EMBO J.*, vol. 32, no. 19, pp. 2561–2574, 2013.

[89] A. Eldar and M. B. Elowitz, "Functional roles for noise in genetic circuits.," *Nature*, vol. 467, no. 7312, pp. 167–173, 2010.

[90] R. Losick and C. Desplan, "Stochasticity and cell fate.," *Science*, vol. 320, no. 5872, pp. 65–68, 2008.

[91] J. M. Raser and E. K. O'Shea, "Noise in gene expression: origins, consequences, and control.," *Science*, vol. 309, no. 5743, pp. 2010–2013, 2005.

[92] M. Thattai and A. van Oudenaarden, "Intrinsic noise in gene regulatory networks.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 15, pp. 8614–8619, 2001.

[93] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell.," *Science*, vol. 297, no. 5584, pp. 1183–1186, 2002.

[94] P. S. Swain, M. B. Elowitz, and E. D. Siggia, "Intrinsic and extrinsic contributions to stochasticity in gene expression.," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. 20, pp. 12795–12800, 2002.

[95] M. Kaern, C. E. Elston, W. J. Blake, and J. J. Collins, "Stochasticity in gene expression: from theories to phenotypes.," *Nat. Rev. Genet.*, vol. 6, no. 6, pp. 451–464, 2005.

[96] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, "Regulation of noise in the expression of a single gene.," *Nat. Genet.*, vol. 31, no. 1, pp. 69–73, 2002.

[97] N. Maheshri and E. K. O'Shea, "Living with noisy genes: how cells function reliably with inherent variability in gene expression.," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 36, pp. 413–34, 2007.

[98] V. Shahrezaei and P. S. Swain, "The stochastic nature of biochemical networks.," *Curr. Opin. Biotechnol.*, vol. 19, no. 4, pp. 369–374, 2008.

[99] D. F. Anderson, J. C. Mattingly, H. F. Nijhout, and M. C. Reed, "Propagation of fluctuations in biochemical systems i: linear ssc networks.," *Bull. Math. Biol.*, vol. 69, no. 6, pp. 1791–1813, 2007.

[100] L. Cai, C. K. Dalal, and M. B. Elowitz, "Frequency-modulated nuclear localization bursts coordinate gene regulation.," *Nature*, vol. 455, no. 7212, pp. 485–90, 2008.

[101] H. Shankaran, D. L. Ippolito, W. B. Chrisler, H. Resat, N. Bollinger, L. K. Opresko, and H. S. Wiley, "Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor.," *Mol. Syst. Biol.*, vol. 5, no. 332, p. 332, 2009.

[102] J. E. Ferrell, "Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability.," *Curr. Opin. Cell Biol.*, vol. 14, no. 2, pp. 140–148, 2002.

[103] T. Suda, J. Suda, and M. Ogawa, "Single-cell origin of mouse hemopoietic colonies expressing multiple lineages in variable combinations.," *PNAS*, vol. 80, no. 21, pp. 6689–93, 1983.

[104] H. J. Beaumont, J. Gallie, C. Kost, G. C. Ferguson, and P. B. Rainey, "Experimental evolution of bet hedging.," *Nature*, vol. 462, no. 1476-4687, pp. 90–93, 2009.

[105] M. Richard and G. Yvert, "How does evolution tune biological noise?," *Front. Genet.*, vol. 5, no. OCT, pp. 1–8, 2014.

[106] C. W. Gardiner, *Stochastic methods: a handbook for the natural and social sciences.* Berlin: Springer, 2009.

[107] N. G. V. Kampen, *Stochastic processes in physics and chemistry.* Amsterdam: North-Holland, 1981.

[108] F. Harary, *Graph theory.* Addison-Wesley, London, 1969.

[109] G. A. Pavlopoulos, M. Secrier, C. N. Moschopoulos, T. G. Soldatos, S. Kossida, J. A., R. Schneider, and P. G. Bagos, "Using graph theory to analyze biological networks.," *BioData Min.*, vol. 4, no. 1, p. 10, 2011.

[110] A. Einstein, "On the motion of small particles suspended in a stationary liquid, as required by the molecular kinetic theory of heat.," *Ann. Phys. (Berlin)*, vol. 322, no. 8, pp. 549–560, 1905.

[111] E. Frey and K. Kroy, "Brownian motion: a paradigm of soft matter and biological physics.," *Ann. Phys. (Berlin)*, vol. 14, no. 1-3, pp. 20–50, 2005.

[112] L. J. S. Allen, *An introduction to stochastic processes with applications to biology.* London: CRC Press, 2011.

[113] J. L. Doob, "The brownian movement and stochastic equations.," *Ann. Math.*, vol. 43, no. 2, pp. 351–369, 1942.

[114] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations.," *SIAM Rev.*, vol. 43, no. 3, pp. 525–546, 2001.

[115] T. Jahnke and W. Huisinga, "Solving the chemical master equation for monomolecular reaction systems analytically.," *J Math Biol*, vol. 54, no. 1, pp. 1–26, 2007.

[116] D. T. Gillespie, "Stochastic simulation of chemical kinetics.," *Ann. Rev. OF Phys. Chem.*, vol. 58, pp. 35–55, 2007.

[117] B. Mlykti, K. Burrage, and K. C. Zygalakis, "Fast stochastic simulation of biochemical reaction systems by alternative formulations of the chemical langevin equation.," *J. Chem. Phys.*, vol. 132, 2010.

[118] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.," *J. Comput. Phys.*, vol. 22, no. 4, pp. 403–434, 1976.

[119] D. F. Anderson, "A modified next reaction method for simulating chemical systems with time dependent propensities and delays.," *J. Chem. Phys.*, vol. 127, no. 21, 2007.

[120] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems.," *J. Chem. Phys.*, vol. 115, no. 4, pp. 1716–1733, 2001.

[121] D. T. Gillespie, "The multivariate langevin and fokker-planck equations.," *Am. J. Phys.*, vol. 64, no. 10, p. 1246, 1996.

[122] D. T. Gillespie, "The chemical langevin equation.," vol. 297, no. May 2013, pp. 297–306, 2000.

[123] L. Szpruch and D. J. Higham, "Comparing hitting time behaviour of Markov jump processes and their diffusion approximations.," vol. 8, no. 2, pp. 605–621, 2010.

[124] D. T. Gillespie, "The chemical Langevin and Fokker-Planck equations for the reversible isomerization reaction.," *J. Phys. Chem. A*, vol. 106, no. 20, pp. 5063–5071, 2002.

[125] E. Estrada and D. J. Higham, "Network properties revealed through matrix functions.," *SIAM Rev.*, vol. 52, no. 4, pp. 696–714, 2010.

[126] E. Estrada and N. Hatano, "Communicability in complex networks.," *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, vol. 77, no. 3, pp. 1–12, 2008.

[127] E. Estrada and J. A. Rodríguez-Velázquez, "Subgraph centrality in complex networks.," *Phys. Rev. E*, vol. 71, p. 056103, May 2005.

[128] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*. CUP, 1994.

[129] L. C. Freeman, "Centrality in social networks conceptual clarification.," *Soc. Networks*, vol. 1, no. 3, pp. 215–239, 1978.

[130] U. V. Luxburg, "A tutorial on spectral clustering.," *Stat. Comput.*, vol. 17, pp. 395–416, 2006.

[131] J. A. Bondy and U. S. R. Murty, *Graph theory with applications.*, vol. 290. Elsevier Science Ltd/North-Holland, 1976.

[132] J. G. Restrepo, E. Ott, and B. R. Hunt, "Approximating the largest eigenvalue of network adjacency matrices.," *Phys. Rev. E*, vol. 76, no. 056119, pp. 1–7, 2013.

[133] R. M. May, "Will a large complex system be stable?," *Nature*, vol. 238, no. 5364, pp. 413–414, 1972.

[134] S. Allesina and S. Tang, "Stability criteria for complex ecosystems.," *Nature*, vol. 483, no. 7388, pp. 205–208, 2012.

[135] A. Edelman, "The circular law and the probability that a random matrix has k real eigenvalues.," *Preprint*, vol. 232, pp. 1–20, 1993.

[136] V. L. Girko, "Circular law.," *Theory Probab. Appl.*, vol. 29, no. 4, pp. 694–706, 1984.

[137] S. Mangan and U. Alon, "Structure and function of the feed-forward loop network motif.," *PNAS*, vol. 100, no. 21, pp. 11980–11985, 2003.

[138] T. Matsuda, T. Nakamura, K. Nakao, T. Arai, M. Katsuki, T. Heike, and T. Yokota, "STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells," *EMBO J.*, vol. 18, no. 15, pp. 4261–4269, 1999.

[139] H. Niwa, T. Burdon, I. Chambers, and A. Smith, "Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3.," *Genes Dev.*, vol. 12, no. 13, pp. 2048–2060, 1998.

[140] H. Niwa, J. Miyazaki, and A. G. Smith, "Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.," *Nat. Genet.*, vol. 24, no. 4, pp. 372–376, 2000.

[141] J. Nichols, B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Scholer, and A. Smith, "Formation of pluripotent stem cells in the mammalian embryo dependes on the POU transcription factor Oct4.," *Cell*, vol. 95, pp. 379–391, 1998.

[142] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks.," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, 2002.

[143] X. Zhang, J. Zhang, T. Wang, M. A. Esteban, and D. Pei, "Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells.," vol. 283, 2008.

[144] N. Festuccia, R. Osorno, F. Halbritter, V. Karwacki-Neisius, P. Navarro, D. Colby, F. Wong, A. Yates, S. R. Tomlinson, and I. Chambers, "Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells.," *Cell Stem Cell*, vol. 11, no. 4, pp. 477–490, 2012.

[145] B. D. MacArthur, A. Sevilla, M. Lenz, F.-J. Müller, B. M. Schuldt, A. A. Schuppert, S. J. Ridden, P. S. Stumpf, M. Fidalgo, A. Maayan, *et al.*, "Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity.," *Nat. Cell Biol.*, vol. 14, no. 11, pp. 1139–1147, 2012.