King's College London

Department of Natural and Mathematical Science

Global Stability and Entropy Production for a Stochastic Neural Network Model

Author:
Jacob Delveaux

Supervisor:
Prof. Yan Fyodorov

A report submitted for the degree of

MSc in Non-Equilibrium Systems

April 16, 2019

Abstract

Much of the current research in non-equilibrium systems seeks to classify the dynamical stability. Quantifying such stability continues to be a challenge. Herein, we explore a Hopfield-like neural network in the presence of noise under the probability-flux landscape formalism. We find several measures of stability through Lyapunov function that charactarize the global probability space. Furthermore, we quantify the extent of non-equilibrium behavior through the entropy production of the system in relation to dynamics on the landscape. Lastly, we attempt to derive an analytical expression for the steady-state probability distribution *a priori* from the interaction matrix in the case of a linear response function.

Acknowledgements

I would like to thank Jozsef, Yanik, Rob, Eddie, Martina, Giorgio, Dom, and Lewis being great support in the office during the production of this paper; and to all faculty of the CANES programme for their outstanding excellence in teaching.

Contents

List of Figures

Chapter 1

Introduction

Complex systems with many non-linear interacting degrees of freedom; a phrase that would have induced nightmares in any physicist in Newton's time, has emerged as a topic of wide scale discussion over the last century and is quickly paving the way for future research. These once intractable problems have been tamed by massive developments in computation and the fields of stochastic systems, random matrix theory, and statistical mechanics. The impact of the research has had a broad scope, elucidating problems in insect colony behavior[1], ecological stability[2], feedback in gene-regulation networks[3, 4], and spontaneous order in the economy[5, 6]. We have substituted quantitative results for once only qualitative theories[7], and provided new perspectives on the age old problems of society[8], emergence[9], and evolution[10].

The myriad methods used to study such systems touches the boundary of modern mathematics and physics, though generally they fall into two classes: abstracting out the core components of the system into a simplified framework, gaining insight into the general behavior; and modelling to minute detail, observing the individual transitions and emergence phenomena *ab initio*. Herein, we analyze a representation of a neural network structure, mimicking neuron-neuron interactions in the brain, much in the spirit of the former class. By abstracting the structure one can describe similar dynamics from predator-prey interactions to gene regulatory networks. The model, to be detailed in Section 2.1 is shown here to lead to novel, non-equilibrium states.

Special detail in this report will address the emergence of out-of-equilibrium behavior. As we will see, the non-deterministic trajectory of the neural network structure admits limit cycles and oscillating steady states in the presence of asymmetric interactions between neurons. We will review the existing literature and list key results relating to our model (Section 2.2), develop the probability and flux landscape formalism (Section(s) 2.2.3, 3.1), classify the global stability (Section(s) 3.2, 3.3.2, 3.4), provide a quantitative measure of non-equilibrium behavior (Section 3.3.1), and attempt a method to find steady-states *a priori* in Section 3.5. In Chapter 2 we build the model and review the current literature. Chapter 3 will be dedicated to developing the landscape formalism and deriving main results. We will have a discussion and present simulation results in Chapter 4 with a conclusion in Chapter 5.

Chapter 2

Background

2.1 Model Description

Our model is an extension of the original Hopfield neural network[11], which models the collective action of a neural circuit taking into account firing rate, capacitance, and resistance of the individual synapses; itself a non-linear analogue to the McCulloch-Pitts neural network[12]. Consider a system of N interacting neurons, the state (synaptic potential) of each being encoded in a vector $\boldsymbol{x} = (x_1, ..., x_N) \in R^N$. The state of each neuron evolves according to four possible interactions: 1) discharge of synaptic potential, 2) input from other neurons, 3) coupling to an external magnetic field, and 4) chemical oscillations, structural variability, oxygen concentration, and a several other more minor factors.

Interaction strength between neuron i and j are given by an element M_{ij} in a random matrix $\boldsymbol{M}$. The matrix is comprised of mean-zero Gaussian random variables with variance $\langle M_{ij}^2 \rangle = M^2/N$. There is no self-self interaction so diagonal terms are set to zero: $M_{ii} = 0, \forall i$. A parameter, η, controls the degree of asymmetry for the matrix; where $\eta = 1$ implies a fully symmetric matrix, $\eta = 0$ for a fully random, and $\eta = -1$ for an anti-symmetric matrix. The less significant contributions to the dynamics, in interaction (4), are grouped together as a Gaussian white noise term modeled by an N independent Brownian motions each at temperature T and represented as the vector $\boldsymbol{\xi} = (\xi_1, ..., \xi_N) \in R^N$ such that $\langle \xi(t_1)\xi(t_2) \rangle = 2T\boldsymbol{D}\delta(t_1 - t_2)$. $\boldsymbol{D}$ is the diffusion matrix, here taken to be the identity $\boldsymbol{I}$. We take our system to be free from a magnetic field. The evolution of the system is represented as a set of N coupled Langevin equations:

$$\frac{dx_i}{dt} = -x_i + \sum_{j \neq i} M_{ij} g(x_j) + \xi_i(t) \tag{2.1}$$

The function g acts to preserve synaptic non-linearity, here taken to be the hyperbolic tangent $g(x) = tanh(x)$.

2.2 Review of Literature

This model has been extensively researched in both its mathematical properties of stability[13], chaotic behavior[14], steady-states[15], &c. and applications in computing[11, 16], biology[10], finance[17], &c. Without the noise term, the right hand size of Eqn. (2.1) are Kirchhoff equations[14] for circuits.

2.2.1 Stability

For the symmetric case in the presence of a magnetic field, $\boldsymbol{I}$, and absence of noise, the dynamics of the network are fully deterministic, following the energy function[11]:

$$E = -\frac{1}{2} \sum_{i \neq j} \sum_{j \neq i} M_{ij} g(x_i) g(x_j) - \sum I_i g(x_i) \tag{2.2}$$

For computational applications, networks contain an update procedure for $\boldsymbol{M}$; a non-negative definite interaction matrix leads to a steady-state with the serial (stationary)[18], parallel (bi-stable/stationary)[16], and partial simultaneous (limit cycle) updates[19].

There have been several papers that have established the sufficient conditions for local stability and equilibrium points[20, 21, 22, 23], number of equilibrium points[13, 15] and exponential global stability[24, 25, 26] extended to the case of asymmetric networks. For use in signal processing, the stability of time delayed asymmetric Hopfield networks has been investigated[27, 28]. Absolute stability criteria has been determined for symmetric[29], normal[30], and Lyapunov diagonally (semi)stable[31, 29] matrices.

2.2.2 Chaotic Behavior

Chaos has been both predicted by[32, 33] and studied within neural networks. It has been shown that in the limit of linear responses, the dynamics follow fixed-point behaviour[34]. In high degrees of non-linearity and random interaction matrices, Hopf bifurcations can be observed[34]. Sompolinsky *et* al. (1988)[14] demonstrated the emergence of chaotic behaviour in the absence of a zero fixed point in the infinite node limit. Babcock and Westervelt (1987)[35] analyzed driven circuits and found individual overdamping can produce a global underdamped behavior. Chaotic behaviour has been demonstrated rarely in small systems, however, a notable example illustrates the existence of relaxation oscillation in a three neuron Hopfield model[36]. Likewise, a detailed numerical simulation of a four-neuron network[37] indicated chaotic attractors, positive Lyapunov exponents, and Hopf bifurcations.

2.2.3 Probability Landscape

When the force of a system follows the gradient of a potential, it is natural to visualize the dynamics as paths along a surface leading to local and global minima; as is the case for equilibrium systems with deterministic trajectories. This technique is widely used in biology[38] and transition state theory[39], where the construction of the landscape can be performed a priori by the known interactions between parts of the system. However, for non-equilibrium, stochastic trajectories, interaction energies cannot often be predicted a priori, and the dynamics generally do not follow the gradient of a potential[40]. Several other formalisms have been derived to confront the dynamics associated with non-deterministic trajectories, such as the Janssen - De Dominicis[41] Langevin field theoretic techniques that make use of QFT methods and diagrammatic perturbation. However, most fail to capture the global properties and stability of the system.

Essentially, one attempts to mimic the potential landscape approach by decomposing the force into a gradient and remainder term[42]. The method rests on the existence of a steady-state solution to the Fokker-Planck equation governing the dynamics of the system in the form $P_{ss} = e^{-\Phi}$. The first application of this method occurs in a paper by H. Haken[43] in the quantum description of lasers. Φ in the non-equilibrium quantum description drew an immediate analogue to the thermodynamic potential in the equilibrium Landau theory, a realization expanded upon in a series of papers by R. Graham (1971, 1983)[44, 45, 46] who laid the foundations for the non-equilibrium landscape framework. His main focus was on systems approaching equilibrium, though he did derive several major qualitative general results, such as that the dynamics result in a flux around surfaces of constant Φ and minimum energy states become degenerate if not invariant under symmetry transformations of the system (leading to 'Mexican Hat' style landscapes). Similar decompositional frameworks were developed for the master equation by Hu[47] (with later applications in biology [48] [3]) and for the Langevin by Ao using a symmetric/anti-symmetric decomposition[49].

2.2.4 Entropy Production in Non-Equilibrium Systems

It is well known that equilibrium systems tend to a state of maximum entropy[50]. This is generally not the case for non-equilibrium systems, whose sisyphean behaviour incites a continual production of entropy. Numerous attempts have been made to quantify the entropics of non-equilibrium steady-states, beginning with the minimum entropy production principle [51] first formulated by Prigogine[52] in the mid 1900s; claiming the non-equilibrium steady state is characterized by minimizing entropy production. This was slowly overthrown by the maximum entropy production principle (MEPP), originating from the work of Ziegler[53]. Although the two may seem contradictory, it can be shown that Prigogine's principle is a special case of the more general Ziegler principle dealing with a stationary system with free forces near equilibrium[54]. For general, dissipative, non-equilibrium systems, MEPP asserts that the steady-state is that which maximizes entropy production. The microscopic formulation arises immediately from the solution of the linearized Boltzmann equation by Enskog[55] and later Hellund and Uehling[56], where the solution of the velocity distribution function is shown to be that which maximizes entropy production. The method has been used to describe classical gasses[57], electron and phonon transport[58], and how Earth retains atmospheric stability[59]. A secondary formulation of MEPP comes from Jaynes'[60, 61] 'MaxENT' formalism, which is a generalization to the method Gibbs used in his classification of the canonical ensembles. Dewar[62, 63] shows how MEPP arises as a consequence of Jaynes' maximization procedure. More recently, MEPP has gained considerable interest over the last decade, however only few studies exist that describe our current network[64, 65].

Chapter 3

Methods

In this chapter, we derive the probability landscape and flux formalism using the decomposition described in Section 2.2.3 and apply it to our system for various degrees of network asymmetry. Next, we will present a Lyapunov function applicable in the weak fluctuation limit. Afterwards, the toolbox of non-equilibrium thermodynamics will allow us to quantify the level of out-of-equilibrium behaviour in the form of entropy production, and yield two more Lyapunov functions. Finally, we attempt an analytical solution of an *a priori* steady-state distribution.

3.1 Probability Landscape and Flux Formalism

Let $F_i(\boldsymbol{x}) = -x_i + \sum_{j\neq i} M_{ij} g(x_j)$ represent the deterministic portion of our driving force, our equations of motion then become:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{\xi} \tag{3.1}$$

We can then derive the Fokker-Planck equation for the time-dependent probability distribution[66][67]. Using the Ito convention, this takes the form:

$$\frac{\partial P(\boldsymbol{x},t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\big[F_i(\boldsymbol{x})P(\boldsymbol{x},t)\big] + T\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j} D_{ij} P(\boldsymbol{x},t) \tag{3.2}$$

$$\frac{\partial P(\boldsymbol{x},t)}{\partial t} = -\nabla \boldsymbol{J} \tag{3.3}$$

Where $\boldsymbol{J}$ is the probability flux, $\boldsymbol{J} = \boldsymbol{F}(\boldsymbol{x})P(\boldsymbol{x},t) - \nabla\cdot(T\boldsymbol{D}P(\boldsymbol{x},t))$. As stated in Section 2.2.3, we seek a steady-state distribution in the form $P(\boldsymbol{x})_{ss} = e^{-\Phi}$. Sections 5 and 6 of Risken, H. (1996)[68], outline the necessary conditions for a steady-state probability distribution, briefly they are: 1) initial time probabilities are non-negative and sum to unity, 2) the diffusion tensor must be positive semi-definite, 3) $\boldsymbol{x}\cdot\boldsymbol{D}\cdot\boldsymbol{x} > 0$, 4) $\lim_{\boldsymbol{x}\to\infty} P^{(n)}(\boldsymbol{x},t) = 0 : P^{(n)} = \frac{\partial^n P}{\partial \boldsymbol{x}^n}, n = 1, 2, ...$; and 5) the Langevin dynamics are free from singularities and do not tend to infinity. At the steady-state, $\partial P_{ss}/\partial t = 0$ we have $\nabla \boldsymbol{J_{ss}} = 0$ and our probability flux is divergence-free.

For a divergence free $\boldsymbol{J_{ss}}$, there can be two cases: 1) $\boldsymbol{J_{ss}} = 0$ or 2) $\boldsymbol{J_{ss}} \neq 0$. If $\boldsymbol{J_{ss}} = 0$, the probability distribution is stationary throughout the system, implying detailed balance holds and the system is in equilibrium. In this case, $\boldsymbol{F} = T\boldsymbol{D}\cdot\nabla P_{ss}/P_{ss} = -T\boldsymbol{D}\cdot\Phi + T\boldsymbol{D}$ with $\Phi = -\ln P_{ss}$. Thus, the driving force is represented as moving down the gradient of a potential energy landscape Φ with an additional constant term. However, if $\boldsymbol{J_{ss}} \neq 0$, then J is a recurrent solenoidal field and $\boldsymbol{F} = -T\boldsymbol{D}\cdot\Phi + T\boldsymbol{D} + \boldsymbol{J_{ss}}/P_{ss}$. Thus, the force can be described as both traversing down a gradient and rotating in a curl vector field; similar to an electron moving in both an electric and magnetic field.

The leading author in this research has compared this to a Hodge/Helmholtz decomposition[40], however, this would require the quantity $\boldsymbol{J_{ss}}/P_{ss}$ to be divergence free and not just $\boldsymbol{J_{ss}}$, which in general does not hold[42]. From Eqn. (3.3), we have $\boldsymbol{J_{ss}}/P_{ss}\cdot\nabla\Phi = \nabla\cdot\boldsymbol{J_{ss}}/P_{ss}$, implying a a true Hodge decomposition only in the case the gradient and 'flux' terms are orthogonal.

There are several other decompositions, such as the aforementioned Ao symmetric/anti-symmetric decomposition for the Langevin equation [49], a true Hodge decomposition, and a normal decomposition, that follow a similar ideology of constructing the force to be both a gradient and remainder term. A breif comparison of the methods can be found in [42].

3.2 Lyapunov Stability in the Low Fluctuation Limit

Local stability of a complex system is often assessed by considering the dynamics around a fixed point after a small perturbation. Global stability, on the other hand, must represent the entire phase-space converging to a steady-state. This is often done by finding a Lyapunov function[69] that monotonically changes with time as the system evolves to some minimum/maximum, indicating convergent behaviour.

In the spirit of the WKB method, we expand our intrinsic potential with respect to T as $\lim_{T\to 0}$:

$$\Phi = -\ln P_{ss} = \frac{1}{T}\sum_i T^i \phi_i(\boldsymbol{x})$$
$$\implies P_{ss} = \frac{1}{Z} e^{-\frac{1}{T}\sum_i T^i \phi_i(\boldsymbol{x})}$$

Where Z is a normalization factor. Substituting into the RHS of Eqn. (3.2) at the steady-state yields:

$$0 = -\sum_{j=1}^{N} \frac{\partial}{\partial x_j}\big[F_i(\boldsymbol{x})\frac{1}{Z}e^{-\frac{1}{T}\sum_{i=0}^{\infty} T^i\phi_i(\boldsymbol{x})}\big] + T\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial^2}{\partial x_i \partial x_j} D_{ij}\frac{1}{Z}e^{-\frac{1}{T}\sum_{i=0}^{\infty} T^i\phi_i(\boldsymbol{x})} \tag{3.4}$$

For the first term in the sum:

$$= -\sum_{j=1}^{N}\frac{\partial}{\partial x_j}\big[F_i(\boldsymbol{x})P_{ss}(\boldsymbol{x})\big]$$
$$= -\sum_{j=1}^{N} F_j'(\boldsymbol{x})P_{ss}(\boldsymbol{x}) + F_j(\boldsymbol{x})P_{ss}'(\boldsymbol{x})$$
$$= -\sum_{j=1}^{N} P_{ss}\big[F_j'(\boldsymbol{x}) - \frac{1}{T}\sum_{i=0}^{\infty} T^i \frac{\partial \phi_i(\boldsymbol{x})}{\partial x_j}F_j(\boldsymbol{x})\big]$$

by assumption, the second term dominates

$$= \frac{1}{T}P_{ss}\sum_{j=1}^{N}\big[\sum_{i=0}^{\infty} T^i\frac{\partial \phi_i(\boldsymbol{x})}{\partial x_j}F_j(\boldsymbol{x})\big]$$

For the second term, assuming a constant diffusion matrix:

$$= T\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial^2}{\partial x_i\partial x_j}D_{ij}P_{ss}(\boldsymbol{x})$$
$$= T\sum_{i=1}^{N}\sum_{j=1}^{N} -\frac{1}{T}D_{ij}P_{ss}(\boldsymbol{x})\big[\sum_{i=0} T^i\frac{\partial^2\phi_i(\boldsymbol{x})}{\partial x_j \partial x_k}\big] + \frac{1}{T^2}P_{ss}(\boldsymbol{x})\big[\sum_{i=0} T^i \frac{\partial\phi_i(\boldsymbol{x})}{\partial x_j}\frac{\partial\phi_i(\boldsymbol{x})}{\partial x_k}\big]$$

like above, the second term dominates

$$= \frac{1}{T}P_{ss}(\boldsymbol{x})\sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij}\big[\sum_{i=0} T^i\frac{\partial\phi_i(\boldsymbol{x})}{\partial x_j}\frac{\partial\phi_i(\boldsymbol{x})}{\partial x_k}\big]$$

Adding the two terms and truncating the series at T^{-1} gives a Hamilton-Jacobi equation

$$\boldsymbol{F}\cdot\nabla\phi_0 + T\nabla\phi_0\cdot\boldsymbol{D}\cdot\nabla\phi_0 = 0 \tag{3.5}$$

Using the chain rule to find the time derivative, we find

$$\begin{aligned}\frac{d\phi_0}{dt} &= \frac{d\boldsymbol{x}}{dt}\cdot\nabla\phi_0 \\ &= \boldsymbol{F}\cdot\nabla\phi_0 = -T\nabla\phi_0\cdot\boldsymbol{D}\cdot\nabla\phi_0 \leq 0\end{aligned}$$

Hence, $\phi_0 = \lim_{T\to 0} T\Phi$ serves as a Lyapunov function for the system in the low fluctuation / deterministic limit. Solutions to the Hamiltonian-Jacobi equation are highly non-trivial; often numerical techniques such as the Eikonal equation or Newton-Raphson method can be used.

3.3 Non-Equilibrium Stochastic Thermodynamics

Working with global probability distributions, it is only natural to consider the thermodynamic properties of our system. Stochastic thermodynamics extends the traditional thermodynamic quantities, work, entropy, and free energy, to the level of individual non-equilibrium trajectories[70] and thermal fluctuations[71]. We will first derive an expression for entropy production from the individual trajectories, then consider free energy as a Lyapunov function.

3.3.1 Entropy Production Rate

The non-equilibrium thermodynamic entropy is given simply by the continuous Shannon entropy:

$$S(t) \equiv -\int d\boldsymbol{x}, p(\boldsymbol{x},,t)\ln p(\boldsymbol{x},,t) \equiv \langle s(t)\rangle \tag{3.6}$$

Where $s(t) = -\ln p(\boldsymbol{x}(t),t)$ represents a trajectory-dependent entropy[72], obtained via evaluating the Fokker-Planck equation along the trajectory $\boldsymbol{x}(t)$. Differentiating with respect to time yields:

$$\begin{aligned}\frac{ds(t)}{dt} &= -\boldsymbol{x}'(t)\frac{\frac{\partial}{\partial\boldsymbol{x}}p(\boldsymbol{x},t)}{p(\boldsymbol{x},t)}_{|\boldsymbol{x}=\boldsymbol{x}(t)} - \frac{\frac{\partial}{\partial t}p(\boldsymbol{x},t)}{p(\boldsymbol{x},t)}_{|\boldsymbol{x}=\boldsymbol{x}(t)} \\ &= -\boldsymbol{x}'(t)\frac{F(\boldsymbol{x})}{T}_{|\boldsymbol{x}=\boldsymbol{x}(t)} + \boldsymbol{x}'(t)\frac{\boldsymbol{J}(\boldsymbol{x},t)}{Tp(\boldsymbol{x},t)}_{|\boldsymbol{x}=\boldsymbol{x}(t)} - \frac{\frac{\partial}{\partial t}p(\boldsymbol{x},t)}{p(\boldsymbol{x},t)}_{|\boldsymbol{x}=\boldsymbol{x}(t)}\end{aligned}$$

We see that the first term corresponds to the heat dissipation of the medium $\frac{dq}{dt} = \boldsymbol{x}'(t)F(\boldsymbol{x}) \equiv T\frac{ds_m(t)}{dt}$. Therefore, the total trajectory-dependent entropy production is given by:

$$\frac{ds_{tot}(t)}{dt} = \frac{ds_m(t)}{dt} + \frac{ds(t)}{dt} = -\frac{\frac{\partial}{\partial t}p(\boldsymbol{x},t)}{p(\boldsymbol{x},t)}_{|\boldsymbol{x}=\boldsymbol{x}(t)} + \boldsymbol{x}'(t)\frac{\boldsymbol{J}(\boldsymbol{x},t)}{Tp(\boldsymbol{x},t)}_{|\boldsymbol{x}=\boldsymbol{x}(t)} \tag{3.7}$$

Macroscopic total entropy, $S_{tot}(t)$, is obtained by first averaging over all trajectories:

$$\langle\boldsymbol{x}'(t)|x,t\rangle = \langle\boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{\xi}|x,t\rangle$$

The first term can be simplified with the Fokker-Planck equation, the noise term averages to zero, leaving us with $\langle\boldsymbol{x}'(t)|x,t\rangle = \boldsymbol{J}(\boldsymbol{x},t)/p(\boldsymbol{x},t)$, which is then averaged over all states x for the total entropy production:

$$\frac{dS_{tot}}{dt} \equiv \left\langle\frac{ds_{tot}}{dt}\right\rangle = \int\frac{\boldsymbol{J}(\boldsymbol{x},t)^2}{Tp(\boldsymbol{x},t)}d\boldsymbol{x} \tag{3.8}$$

Similarly we the average entropy of the medium is:

$$\frac{dS_m}{dt} \equiv \left\langle\frac{ds_m}{dt}\right\rangle = \int F(\boldsymbol{x})\boldsymbol{J}(\boldsymbol{x},t)/T d\boldsymbol{x} \tag{3.9}$$

Hence the system entropy becomes:

$$\frac{dS}{dt} = \frac{dS_{tot}}{dt} - \frac{dS_m}{dt} = \int\frac{\boldsymbol{J}(\boldsymbol{x},t)^2}{Tp(\boldsymbol{x},t)}d\boldsymbol{x} - \int F(\boldsymbol{x})\boldsymbol{J}(\boldsymbol{x},t)/T d\boldsymbol{x} \tag{3.10}$$

An equivalent result is derived via a macroscopic approach by Qian (2018)[73], though it relies on knowning local temperature in the system which is difficult to classify in non-equilibrium cases. The first term on the RHS is the entropy production rate (EPR) and the second is the dissipative entropy. Our system does not break any of the restrictions outlined in [74], hence, the MEPP discussed in Chapter 2 should be satisfied. This implies that the EPR tends towards a maximum as our system evolves.

3.3.2 Free Energy as a Lyapunov Function

The non-equilibrium free energy is defined as

$$\mathcal{F} = \mathcal{E} - T\mathcal{S} \tag{3.11}$$

Where $\mathcal{E} = \int \phi_0 \mathcal{P}(\boldsymbol{x}, t) d\boldsymbol{x} = -\mathcal{T} \int \ln(\mathcal{Z}\mathcal{P}_{ss})\mathcal{P}(\boldsymbol{x}, t) d\boldsymbol{x}$, is the average energy, following from $\mathcal{P}_{ss}(\boldsymbol{x}) = P_{ss}(\boldsymbol{x})_{|T\to 0} = \frac{1}{\mathcal{Z}} e^{-\phi_0/\mathcal{T}}$ with $\mathcal{T} = T_{|T\to 0}$ and $\mathcal{Z}$ the non-equilibrium partition function $\mathcal{Z} = \int e^{-\phi_0/\mathcal{T}} d\boldsymbol{x}$. Likewise, let $\mathcal{J} = \boldsymbol{J}_{|T\to inf}$, substituting everything into Eqn. (3.11) yields[10][75]:

$$\mathcal{F} = \mathcal{T}\Big[\int \mathcal{P}(\boldsymbol{x}, t) \ln(\mathcal{P}(\boldsymbol{x}, t)/\mathcal{P}_{ss}) d\boldsymbol{x} - \ln \mathcal{Z}\Big] \tag{3.12}$$

The partition function in this case serves as a constant, the remaining term can be shown to be non-negative by[76]:

$$\begin{aligned}
\int \mathcal{P}(\boldsymbol{x}, t) \ln(\mathcal{P}(\boldsymbol{x}, t)/\mathcal{P}_{ss}) d\boldsymbol{x} &= -\int \mathcal{P}(\boldsymbol{x}, t) \ln(\mathcal{P}_{ss}/(\mathcal{P}(\boldsymbol{x}, t)) d\boldsymbol{x} \\
&\geq -\int \mathcal{P}(\boldsymbol{x}, t)\Big(\frac{\mathcal{P}_{ss}}{(\mathcal{P}(\boldsymbol{x}, t)} - 1\Big) d\boldsymbol{x} \\
&= \int \mathcal{P}(\boldsymbol{x}, t) d\boldsymbol{x} - \int \mathcal{P}_{ss}(\boldsymbol{x}) d\boldsymbol{x} = 0
\end{aligned}$$

Differentiating with respect to time:

$$\begin{aligned}
\frac{d\mathcal{F}}{dt} &= \mathcal{T}\frac{d}{dt}\int \mathcal{P} \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x} \\
&= \mathcal{T}\int \frac{d\mathcal{P}}{dt} \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x} + \mathcal{T}\int \frac{d\mathcal{P}}{dt} d\boldsymbol{x} \\
&= -\mathcal{T}\int \nabla \bullet \mathcal{J} \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x} + \mathcal{T}\int \mathcal{J} \bullet \nabla \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x}
\end{aligned}$$

By the conservation of probability, our flux around the boundary of the state space should be zero. Thus, by Gauss' theorem we can eliminate the first term on the RHS:

$$\frac{d\mathcal{F}}{dt} = \mathcal{T}\int \mathcal{J} \bullet \nabla \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x} \tag{3.13}$$

From our definition of $\boldsymbol{J}$ obtained via the Fokker-Planck equation:

$$\begin{aligned}
&\boldsymbol{J} = \boldsymbol{F}P - T\nabla(\boldsymbol{D}P) \\
&\implies \frac{\mathcal{J}}{\mathcal{P}} + \mathcal{T}\boldsymbol{D} \bullet \nabla \ln \mathcal{P} = \boldsymbol{F} = \frac{\mathcal{J}_{ss}}{\mathcal{P}_{ss}} + \mathcal{T}\boldsymbol{D} \bullet \nabla \ln \mathcal{P}_{ss} \\
&\iff \mathcal{J} = \mathcal{P}[\mathcal{J}_{ss}/\mathcal{P}_{ss} - \mathcal{T}\boldsymbol{D} \bullet \nabla \ln(\mathcal{P}/\mathcal{P}_{ss})]
\end{aligned}$$

Which, inserting into Eqn. (3.13) gives:

$$\begin{aligned}
\frac{d\mathcal{F}}{dt} &= \mathcal{T}\int \mathcal{P}\Big[\frac{\mathcal{J}_{ss}}{\mathcal{P}_{ss}} - \mathcal{T}\boldsymbol{D} \bullet \nabla \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big)\Big] \bullet \nabla \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x} \\
&= \mathcal{T}\int \mathcal{J}_{ss} \bullet \nabla\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) d\boldsymbol{x} - \mathcal{T}^2 \int \Big[\nabla \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) \bullet \boldsymbol{D} \bullet \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big)\Big] \mathcal{P} d\boldsymbol{x}
\end{aligned}$$

Using the product rule on the first term, noting $d\mathcal{P}_{ss}/dt = 0$, and again using Gauss' theorem along the boundary yields the final expression for the free energy time derivative

$$\frac{d\mathcal{F}}{dt} = -\mathcal{T}^2 \int \Big[\nabla \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big) \bullet \boldsymbol{D} \bullet \ln\Big(\frac{\mathcal{P}}{\mathcal{P}_{ss}}\Big)\Big] \mathcal{P} d\boldsymbol{x} \leq 0 \tag{3.14}$$

Hence, the non-equilibrium free energy serves as a Lyapunov function in the weak fluctuation limit, reaching a minimum at the steady-state $\mathcal{F} = -\mathcal{T} \ln \mathcal{Z}$.

3.4 Relative Entropy as a Lyapunov Function

Unfortunately, the weak fluctuation limit imposes serious restrictions on our system, as the probability function $P(\boldsymbol{x},t)$ may not fully explore the whole state space. Although our previous two measures of Lyapunov stability take as an assumption weak fluctuations, relative entropy or the Kullback-Leibler divergence between the time dependent and steady-state probability distributions:

$$KL(P||P_{ss}) = \int d\boldsymbol{x} P(\boldsymbol{x},t) \ln \frac{P(\boldsymbol{x},t)}{P_{ss}(\boldsymbol{x})} \tag{3.15}$$

The Lyapunov property follows the same derivation as the free energy case above.

3.5 Exact solution for steady-state probability with linear response

Consider the Fokker-Planck equation for our system given by Eqn. (3.2), we linearize the system taking g(x) = x and insert our expression for the force:

$$\frac{\partial P(\boldsymbol{x},t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\Big[\Big(x_i - \sum_k M_{ik}x_k\Big)P(\boldsymbol{x},t)\Big] + T\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j} D_{ij} P(\boldsymbol{x},t) \tag{3.16}$$

To obtain the steady-state solution, we begin by setting the RHS to zero and taking a Fourier transform. Denote $G(\boldsymbol{k})$ to be the multidimensional Fourier transform of $P(\boldsymbol{x})_{ss}$, namely:

$$G(\boldsymbol{k}) = \int P(\boldsymbol{x})_{ss} e^{\sum_i^N ik_ix_i} d\boldsymbol{x} \tag{3.17}$$

For each part of the above expression, we obtain the transforms:

$$\begin{aligned}
\mathcal{F}\Big\{\frac{\partial}{\partial x_i}P(\boldsymbol{x})\Big\}(\boldsymbol{k}) &= ik_iG(\boldsymbol{k}) \\
\mathcal{F}\Big\{x_i\frac{\partial}{\partial x_i}P(\boldsymbol{x})\Big\}(\boldsymbol{k}) &= k_i\frac{\partial G(\boldsymbol{k})}{\partial k_i} \\
\mathcal{F}\Big\{\frac{\partial^2}{\partial x_i \partial x_j}P(\boldsymbol{x})\Big\}(\boldsymbol{k}) &= -k_ik_jG(\boldsymbol{k}) \\
\mathcal{F}\Big\{\sum_k M_{ik}x_k\frac{\partial}{\partial x_i}P(\boldsymbol{x})\Big\}(\boldsymbol{k}) &= k_i\Big[\sum_l M_{il}\frac{\partial G(\boldsymbol{k})}{\partial k_l}\Big]
\end{aligned}$$

Hence, Eqn. (3.16) becomes:

$$0 = \sum_i k_i\Big[\sum_l M_{il}\frac{\partial G(\boldsymbol{k})}{\partial k_l} - \frac{\partial G(\boldsymbol{k})}{\partial k_i}\Big] - \sum_{i,j} Tk_ik_jG(\boldsymbol{k}) \tag{3.18}$$

$$\iff \sum_{i,j} Tk_ik_jG(\boldsymbol{k}) = \sum_i \Big[\sum_l k_lM_{li} - k_i\Big]\frac{\partial G(\boldsymbol{k})}{\partial k_i} \tag{3.19}$$

Using the assumption $M_{ii} = 0, \forall i$. We now make the ansatz that $G(\boldsymbol{k})$ (and thus $P(\boldsymbol{x})_{ss}$) is a multivariate Gaussian:

$$G(\boldsymbol{k}) \propto e^{-\frac{1}{2}\boldsymbol{k}\boldsymbol{\Theta}\boldsymbol{k}^T}$$

Taking the derivative and inserting back into Eqn. (3.19) yeilds:

$$\begin{aligned}
\sum_{i,j} Tk_ik_j &= \sum_i \Big[\sum_l k_lM_{li} - k_i\Big]\sum_r \Theta_{ir}k_r \\
&= \sum_{i,j} k_i\Theta_{i,j}k_j - \sum_l k_lM_{li}\Theta_{i,j}k_j \\
\implies \boldsymbol{k}T\boldsymbol{k}^T &= \boldsymbol{k}(\boldsymbol{\Theta} - \boldsymbol{M}\boldsymbol{\Theta})\boldsymbol{k}^T \\
\implies (\boldsymbol{\Theta} - \boldsymbol{M}\boldsymbol{\Theta}) + (\boldsymbol{\Theta} - \boldsymbol{M}\boldsymbol{\Theta})^T &= 2T\boldsymbol{I} \\
\iff (\boldsymbol{I} - \boldsymbol{M})\boldsymbol{\Theta} + \boldsymbol{\Theta}(\boldsymbol{I} - \boldsymbol{M})^T &= 2T\boldsymbol{I}
\end{aligned}$$

In the derivation we have utilized the symmetry of the covariance matrix. The last line is a Lyapunov equation, with the solution given by:

$$\boldsymbol{\Theta} = 2T \int_0^\infty e^{-(\boldsymbol{I}-\boldsymbol{M})t} e^{-(\boldsymbol{I}-\boldsymbol{M})^T t} \tag{3.20}$$

Therefore, under our ansatz, our steady state probability distribution is:

$$P(\boldsymbol{x})_{ss} = \frac{1}{(2\pi)^{N/2}\sqrt{\det \Theta}} e^{-\frac{1}{2}\boldsymbol{x}\Theta^{-1}\boldsymbol{x}^T} \tag{3.21}$$

Eqn. (3.20) will have a convergent solution if $(\boldsymbol{I}-\boldsymbol{M})+(\boldsymbol{I}-\boldsymbol{M})^T$ is positive(semi)-definite, which in general does not hold.

Chapter 4

Results

Here we present and discuss our results. Due to the nature of the problem, simulations required forming a discrete probability space which required a large trade-off between computational complexity and accuracy of simulations. As a consequence, we were severely restricted in our system size: e.g. computing the landscape and entropy for 10 neurons on a discrete grid of 100 units requires 100^{10} operations. Our major aim is to determine the effect of asymmetry on network dynamics, therefore the bulk of our results are computed for systems with 5 neurons with a preference for high accuracy.

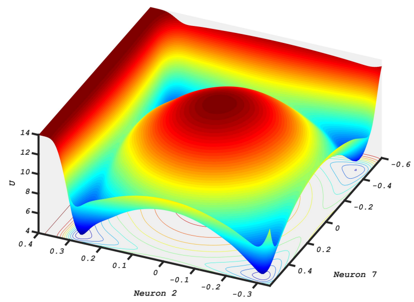
4.1 Probability Landscape

Figure (4.1) illustrates a variety of potential energy landscapes derived from the steady state distribution for various η, N, coupling variance M^2, and T. Each landscape was generated by binning the steady-state distribution of 5000 runs into 100-300 bins depending on system size. Calculation of the N dimensional joint probability requires the solution to N coupled differential equations, which is highly non-trivial without sufficient computing power. For our results we have made use of the Hartree mean field approximation, namely:

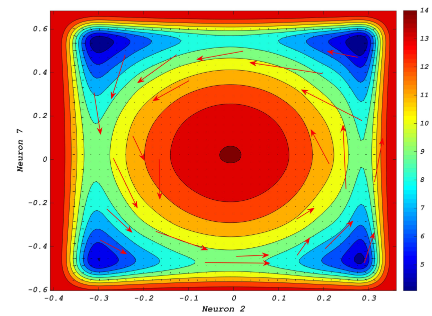
$$Prob[x_1, ..., x_N, t] = \prod_{i}^{N} Prob[x_i, t] \tag{4.1}$$

For most calculations, this reduces the number of operations from L^N where L is the number of points on our discrete space to $L \times N$. However, for determining our probability flux J and entropy production rate (EPR), we must solve the complete coupled system to achieve accurate results. One could choose to apply a Gaussian approximation[77, 75] to estimate probability current which may be applicable in the asymmetric case; though, as we can clearly see from Figures (4.1c, 4.1g, 4.1h), more symmetric interaction matrices must be at minimum described by a sum of Gaussians. Furthermore, many landscapes such as Figure (4.1e) have intricate surfaces that are obfuscated by such an approximation and probability current vectors cannot be easily obtained. On the contours, the trajectory along the gradient and probability flux are indicated by black and red arrows respectively.

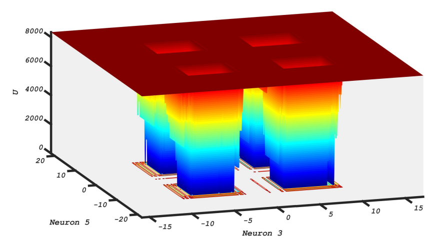
From Figures (4.1c, 4.1g, 4.1h) we clearly see that symmetric interaction matrices lead to stable stationary solutions at values reflected across the zero of the axis. In the case of perfectly symmetric matrices, the flux term tends to zero in the low temperature limit and the dynamics follow the energy function in Eqn. (2.1).



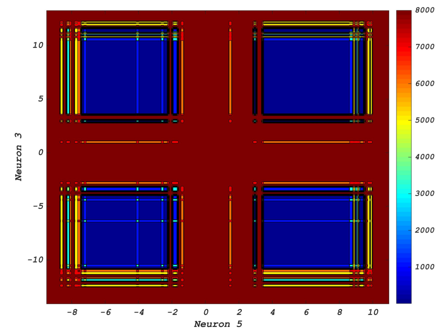
(a) $\eta = -0.9$, N = 20, $\mathrm{M}^2 = 1, T = 0.001$



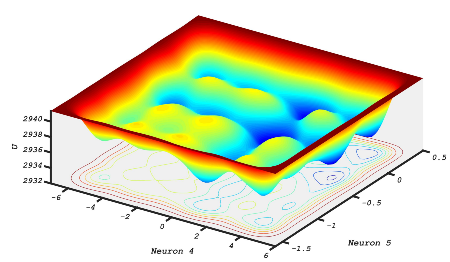
(b) Contour for (a)



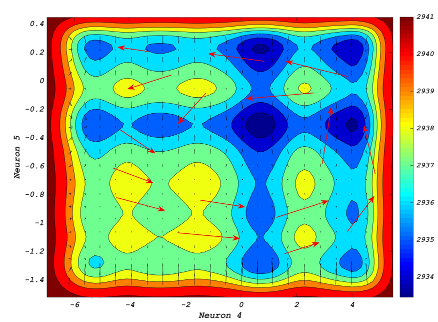
(c) $\eta = 1$, N = 20, $\mathrm{M}^2 = 5, T = 0.001$



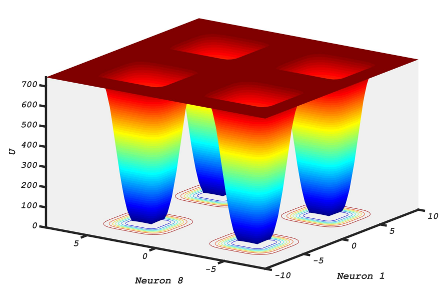
(d) Contour for (c)

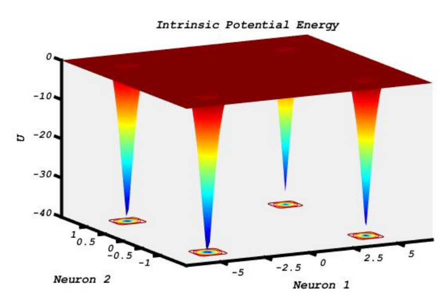


(e) $\eta = 0$, N = 20, J = $\mathrm{M}^2, T = 0.01$



(f) Contour for (e)





(g) $\eta = 0.75$, N = 10, $\mathrm{M}^2 = 10, T = 0.05$

(h) $\eta = 1$, N = 5, $\mathrm{M}^2 = 1, T = 0.001$

Figure 4.1: Probability landscapes for various degrees of asymmetry (η), diffusion coefficients T, coupling variances M^2, and system size N. Landscapes are generated from the steady-state distribution obtained by simulating over 5000 initial conditions. Black arrows show the underlying potential gradient of the surface and red arrows the probability flux.

For the asymmetric case in Figure (4.1a), we find stationary solutions around the edges of a protruded hump, however the landscape itself has become much more smooth allowing trajectories to cycle along the stationary states. Now, the trajectory follows both the gradient of the potential towards the minimum points, but is also pushed by the probability flux around the potential. Likewise, in the completely random case, Figure (4.1e), the trajectory follows the saddle points between minimums around the landscape. By increasing the temperature our landscape becomes flattened, trajectories can freely flow around the minimum due to flux without impediment by barriers; this effect is quantified in terms of entropy production in Figure (4.4f).

4.2 Lyapunov Stability

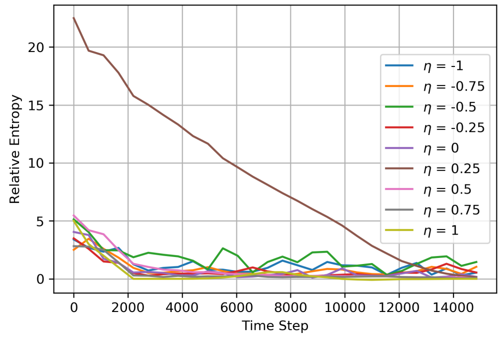
The global stability was assessed by tracking the relative entropy defined in Eqn. (3.15) as shown in Figure (4.2) as the system evolves towards the steady state. At the start of the simulation, the divergence between the initial and steady state distribution is large, though for our range of temperatures we converge to a minimum value. However, in the case of strong coupling, the relative entropy does not converge. Clearly this is due to the failure of the mean-field approximation used to describe the probability distribution, which cannot accurately portray strongly interacting systems.

Although this method allows for finite fluctuations, it has the unfortunate condition, like the free energy, of requiring the steady-state distribution to be known beforehand, which is generally not the case for non-equilibrium systems. Thus, this can really only be a lyapunov function if the steady-state is known, however, this leads to a circular argument: if the steady-state distribution is known, you have already assumed convergence, and if you assume convergence, then what do you need a lyapunov function for? Granted, it is still applicable to cases where the distribution can be obtained or approximated beforehand, or if the system has previously been well studied. The analytical result from Eqn. (3.21) may be able to provide a best estimate for these cases, provided the interaction matrix fits the limitations. The Hamilton-Jacobi lyapunov function derived in Section 3.2 may overcome this limitation, but as previously stated is highly non-trivial to solve and cannot describe systems with large diffusion coefficients.

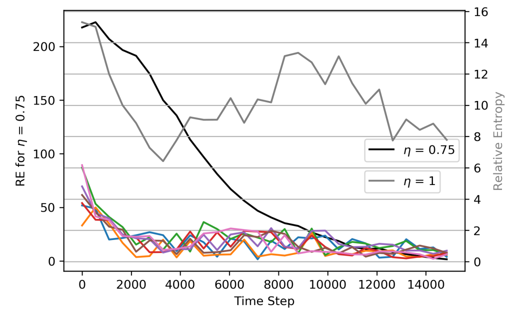
4.3 Entropy Production

The total entropy production, defined in Eqn. (3.8) from considering the evolution of a stochastic trajectory, has been determined for systems with η ranging from -1 to 1. For the symmetric systems, the steady-state is expected to be stationary, and hence the trajectory will center around a constant value. On the other hand, in asymmetric and randomly connected systems, we expect the trajectory to continuously evolve over time, moving from minima to minima over saddle points on the potential landscape. Such a mass-transfer like process will necessarily create a production of entropy, as energy dissipates from the system from battling the gradient of the potential. To better classify the asymmetry of the system, we introduce a new measure the 'Degree of Asymmetry' defined as the ratio of the norms of the anti-symmetric $M_A = (M - M^T)/2$ and symmetric $M_S = (M + M^T)/2$ parts of the interaction matrix: $||M_A||/||M_S||$. Figure (4.3) shows the results of entropy production at the steady state for 18 interaction matrices with various degrees of asymmetry and system properties averaged over 6000 realizations.

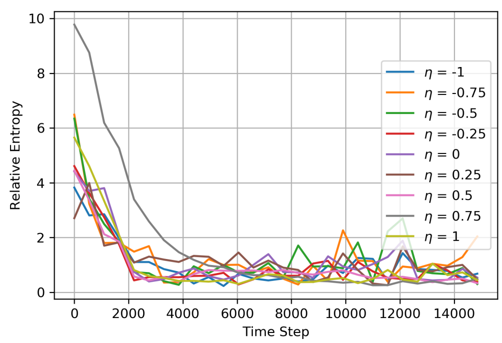
From our simulations we find three main results: 1) for systems with without strong coupling, we find that entropy production decreases as the interactions become more symmetric 2) in general, we also find a corresponding increase in entropy as temperature increases and 3) strongly coupled systems follow the opposite trend, with more symmetric systems admitting higher entropy production.



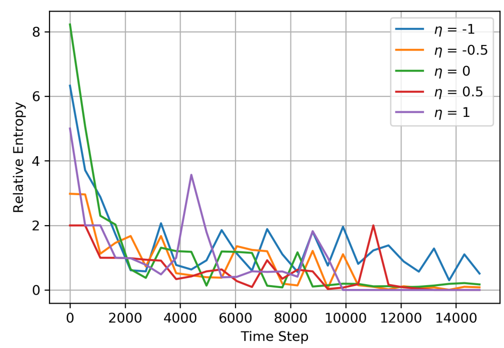
(a) $\mathrm{M}^2 = 1, T = 0.001$



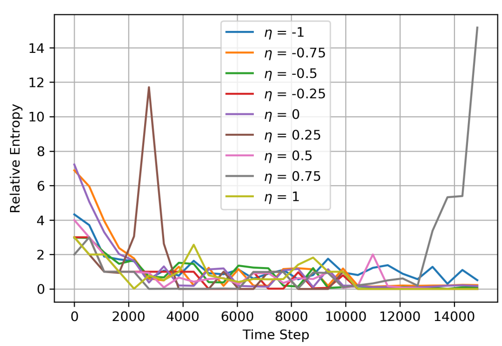
(b) $\mathrm{M}^2 = 1, T = 0.05$



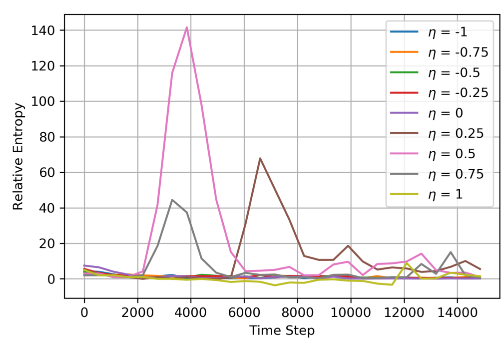
(c) $\mathrm{M}^2 = 1, T = 1$



(d) $\mathrm{M}^2 = 1, T = 10$



(e) $\mathrm{M}^2 = 20, T = 0.001$

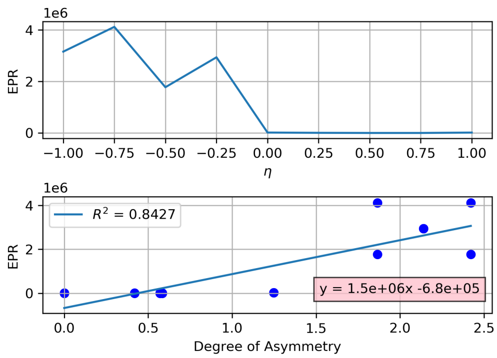


(f) $\mathrm{M}^2 = 10, T = 1$

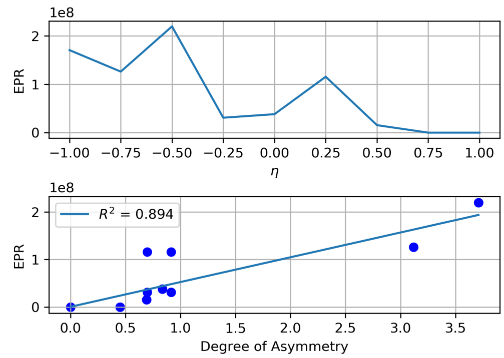
Figure 4.2: Relative entropy (Eqn. (3.15)) with respect to timestep for five (5) neuron systems with various parameters. Runs were calculated over 6000 realizations of different initial conditions for each interaction matrix.

The first result was expected. For this model, symmetrically interacting networks more closely resemble equilibrium systems, where the trajectories can be determined via the gradient of some potential into local minima. Such near-equilibrium systems can be classified by Prigogine's minimum entropy production principle [51]. The driving force for asymmetric and random neural networks is influenced by both the gradient and flux terms; more asymmetric interactions have a higher magnitude of flux[75]. As the temperature of the system increases, the energy barriers become easier to surpass and trajectories can freely explore the state space. Clearly, less inhibited exploration of the state space involves crossing free energy barriers more often, which, by standard thermodynamics, produces an increase in entropy. In the case of an extremely large diffusion coefficient, the entropy production becomes nearly equivalent for all systems, regardless of interaction symmetry (see Figure (4.4f)).

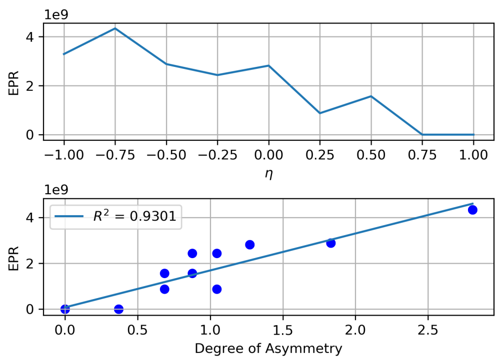
The third result may at first seem to be contradictory to our general statement that asymmetry and temperature drive the production of entropy, though it is completely in line with the notion of entropy production of stochastic systems. Higher degrees of cross-coupling produce a higher driving force for the trajectory, when a small perturbation occurs in the system, i.e. with a non-zero diffusion coefficient, there is a large force exerted that returns the neurons back to the stationary states when the perturbation occurs. Thus the probability flux is calculated to extremely high as one progresses away from the stationary points of the landscape from $\boldsymbol{J} = \boldsymbol{F}(\boldsymbol{x})P(\boldsymbol{x},t) - \nabla \cdot (T\boldsymbol{D}P(\boldsymbol{x},t))$. The large probability flux in turn corresponds to a large production of entropy.



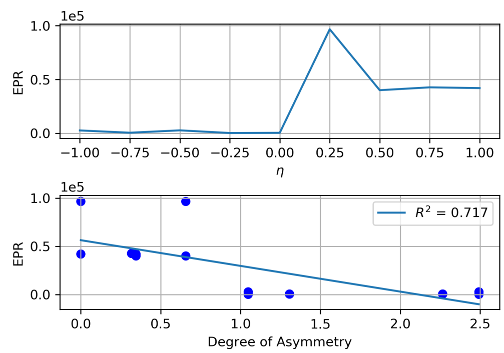
(a) $\mathrm{M}^2 = 1, T = 0.001$



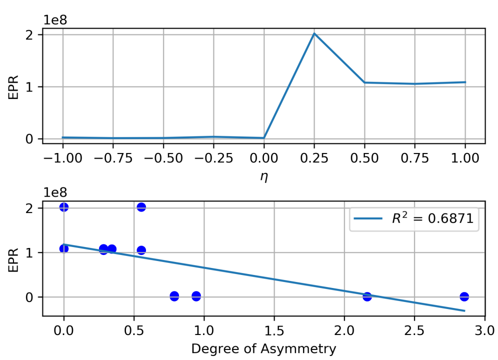
(b) $\mathrm{M}^2 = 1, T = 0.05$



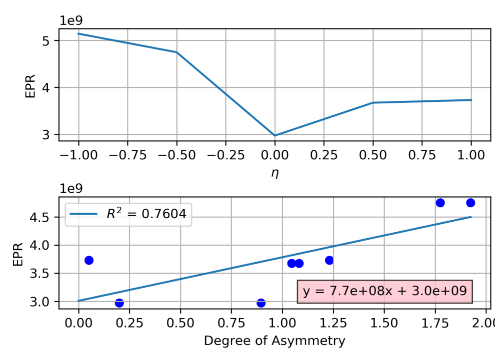
(c) $\mathrm{M}^2 = 1, T = 1$



(d) $\mathrm{M}^2 = 10, T = 1$

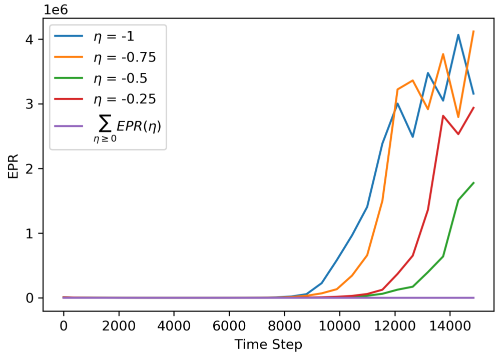


(e) $\mathrm{M}^2 = 20, T = 0.001$

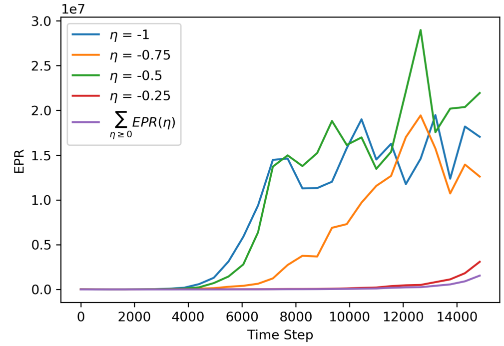


(f) $\mathrm{M}^2 = 1, T = 10$

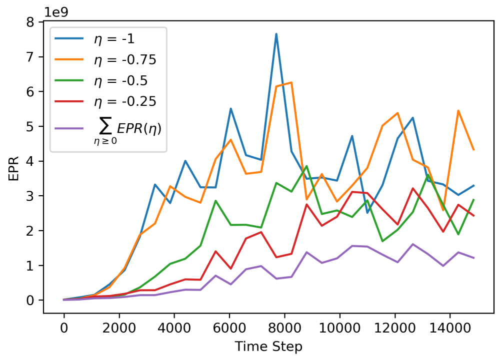
Figure 4.3: The steady-state entropy production rate along various degrees of asymmetry for a variety of neural networks. For set of system parameters, 18 interaction matrices were formed with 6000 realizations each. The degree of asymmetry is defined by the ratio of the norms for the anti-symmetric and symmetric parts of the matrix, $||M_A||/||M_S||$.



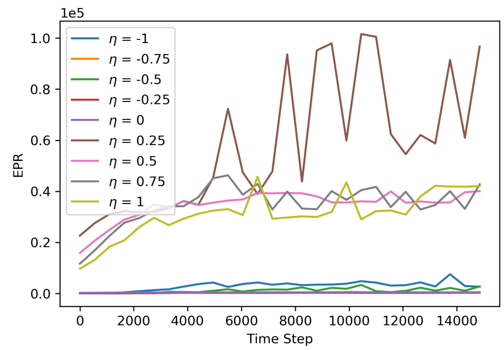
(a) $\mathrm{M}^2 = 1, T = 0.001$



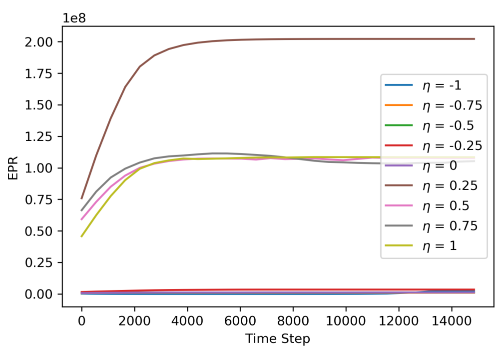
(b) $\mathrm{M}^2 = 1, T = 0.05$



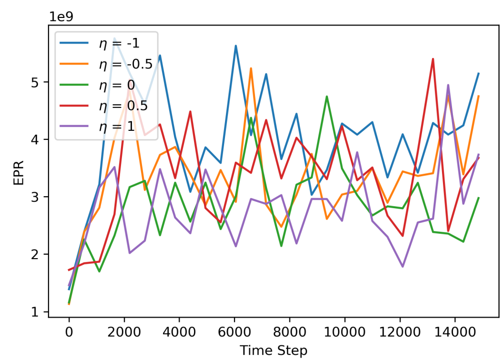
(c) $\mathrm{M}^2 = 1, T = 1$



(d) $\mathrm{M}^2 = 10, T = 1$



(e) $\mathrm{M}^2 = 20, T = 0.001$

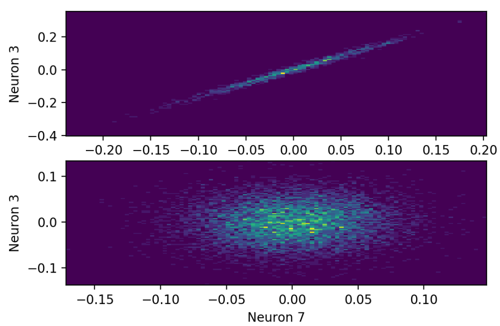


(f) $\mathrm{M}^2 = 1, T = 10$

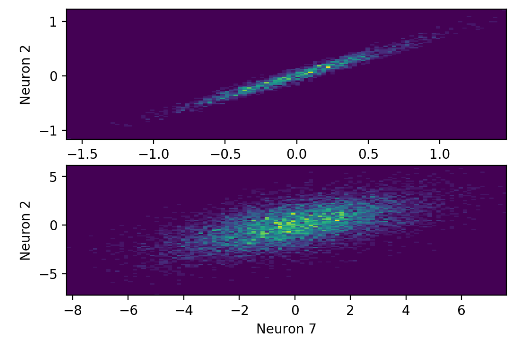
Figure 4.4: Entropy production rate as a function of timestep for size five (5) networks with various coupling strengths and temperatures, averaged over 6000 realizations.

4.4 A Priori Steady State Distribution

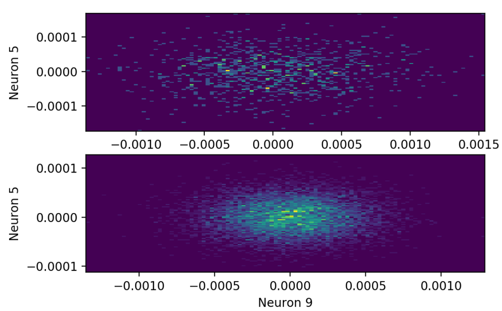
According to our analytical solution in Eqn. (3.21), we should be able to predict the steady state distribution for an interaction matrix $\boldsymbol{M}$ if $(\boldsymbol{I}-\boldsymbol{M})+(\boldsymbol{I}-\boldsymbol{M})^T$ is positive(semi)-definite and our ansatz holds. For various systems of size 10 generated by matrices that fit the above criteria we have performed 16000 simulations and obtained the stead-state distribution and compared to the analytical results in Figure (4.5). Analytical results were obtained by integrating over 1 million iterations with a timestep of 0.001, sufficient for convergence of the covariance matrix. The covariance matrix was used to draw 16000 multivariate Gaussian random variables, and were plotted as a 2 dimensional histogram under the steady state distribution. Generally, we find poor convergence between the two results with the exception of Figures (4.5c and 4.5e). As we have seen in Section 4.1, the topology of symmetric matrices generally requires a multiple multivariate Gaussian fit, rather than our singular ansatz. For non-symmetric matrices, the analytic result may prove useful to be a 'best-guess' stationary distribution for use in a Lyapunov function.



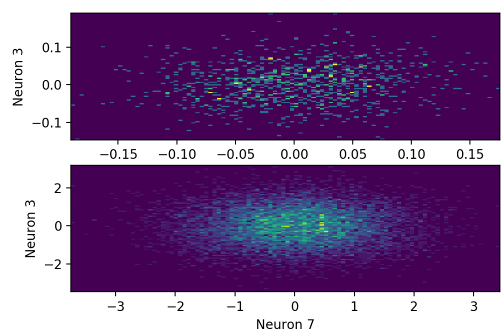
(a) η = 1, T = 0.001



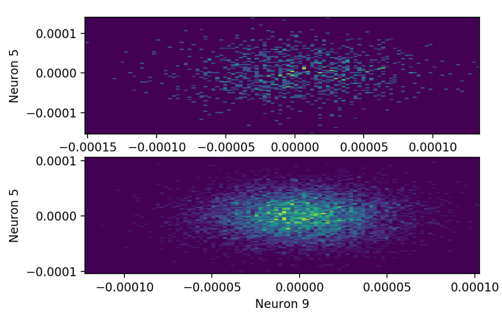
(b) η = 1, T = 1



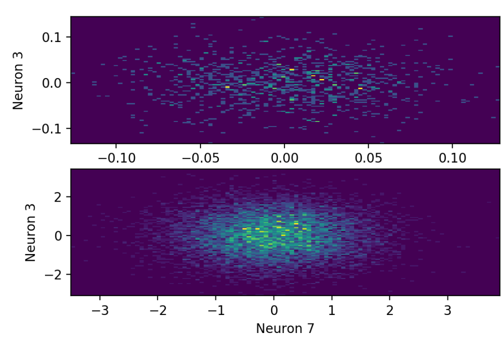
(c) η = 0, T = 0.001



(d) η = 0, T = 1



(e) η = -1, T = 0.001



(f) η = -1, T = 1

Figure 4.5: The steady-state marginal distribution from simulations (top) and analytical distribution (bottom) for asymmetry parameter η = 1 (a,b), =0 (c,d) and =-1 (e,f). Histograms were constructed via 16000 steady state solutions and multivariate Gaussian samples for the simulated and analytical cases respectively. N = 10, M^2 = 1.

Chapter 5

Conclusion

Throughout our exploration, we have uncovered a number of exciting results for our neural network model. Potential energy landscapes and their corresponding fluxes derived from steady-state probabilities have been visualized to provide a qualitative picture into the dynamics of the system. It has been shown that networks with non-symmetric interactions admit trajectories to navigate between minimums in the landscape. These types of dynamics have been used to explain phenomena from memory retrieval[75], stem-cell fate[7], progress through the cell cycle[77], and the process of evolution[10]. Whether or not this proves to be a rigorously valid description has yet to be determined, though the formalism itself does lead to qualitative insights regarding energy flow in open, dissipative, non-equilibrium systems. As we have seen, we can also derive quantitative data in the form of entropy production, which provides a measure to the non-equilibrium behaviour of the system. Higher degrees of asymmetry corresponds to greater entropy production, excepting the case of strong correlations. Further exploration may consider the period and amplitude of the limit cycle or the effect of barrier height on transition between minima. Finally, an analytical expression for the steady-state distribution has been derived for linear responses. Though subject to some stringent conditions, it may prove useful to act as a 'Bayesian estimator' for lyapunov functions of unknown systems.

Bibliography

[1] Robert E Page Jr and Sandra D Mitchell. Self organization and adaptation in insect societies. In *PSA: Proceedings of the biennial meeting of the philosophy of science association*, volume 1990, pages 289–298. Philosophy of Science Association, 1990.

[2] Robert McCredie May. *Stability and complexity in model ecosystems*, volume 6. Princeton university press, 2001.

[3] Keun-Young Kim and Jin Wang. Potential energy landscape and robustness of a gene regulatory network: toggle switch. *PLoS computational biology*, 3(3):e60, 2007.

[4] Feng Luo, Yunfeng Yang, Jianxin Zhong, Haichun Gao, Latifur Khan, Dorothea K Thompson, and Jizhong Zhou. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC bioinformatics*, 8(1):299, 2007.

[5] John Foster. From simplistic to complex systems in economics. *Cambridge Journal of Economics*, 29(6):873–892, 2005.

[6] David Easley, Jon Kleinberg, et al. *Networks, crowds, and markets*, volume 8. Cambridge University Press Cambridge, 2010.

[7] Jin Wang, Li Xu, Erkang Wang, and Sui Huang. The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophysical journal*, 99(1):29–39, 2010.

[8] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.

[9] Steven Strogatz. *Sync: The emerging science of spontaneous order.* Penguin UK, 2004.

[10] Feng Zhang, Li Xu, Kun Zhang, Erkang Wang, and Jin Wang. The potential and flux landscape theory of evolution. *The Journal of chemical physics*, 137(6):065102, 2012.

[11] John J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, 81(10):3088–3092, 1984.

[12] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[13] Yan V Fyodorov and Boris A Khoruzhenko. Nonlinear analogue of the may- wigner instability transition. *Proceedings of the National Academy of Sciences*, 113(25):6827–6832, 2016.

[14] Haim Sompolinsky, Andrea Crisanti, and Hans-Jurgen Sommers. Chaos in random neural networks. *Physical review letters*, 61(3):259, 1988.

[15] FR Waugh, CM Marcus, and RM Westervelt. Fixed-point attractors in analog neural computation. *Physical Review Letters*, 64(16):1986, 1990.

[16] Eric Goles-Chacc, Françoise Fogelman-Soulié, and Didier Pellegrin. Decreasing energy functions as a tool for studying threshold networks. *Discrete Applied Mathematics*, 12(3):261–277, 1985.

[17] Kartik Anand, Jonathan Khedair, and Reimer Kühn. Structural model for fluctuations in financial markets. *Physical Review E*, 97(5):052312, 2018.

[18] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

[19] Dong-Liang Lee. New stability conditions for hopfield networks in partial simultaneous update mode. *IEEE Transactions on neural networks*, 10(4):975–978, 1999.

[20] Jian-Hua Li, Anthony N Michel, and Wolfgang Porod. Qualitative analysis and synthesis of a class of neural networks. *IEEE Transactions on Circuits and Systems*, 35(8):976–986, 1988.

[21] Anthony N Michel and DL Gray. Analysis and synthesis of neural networks with lower block triangular interconnecting structure. *IEEE Transactions on Circuits and Systems*, 37(10):1267–1283, 1990.

[22] Anthony N Michel, Jay A Farrell, and Wolfgang Porod. Qualitative analysis of neural networks. *IEEE Transactions on Circuits and Systems*, 36(2):229–243, 1989.

[23] Hua Yang and Tharam S Dillon. Exponential stability and oscillation of hopfield graded response neural network. *IEEE Transactions on Neural Networks*, 5(5):719–729, 1994.

[24] Yuguang Fang and Thomas G Kincaid. Stability analysis of dynamical neural networks. *IEEE Transactions on Neural Networks*, 7(4):996–1006, 1996.

[25] Abdesselam Bouzerdoum and Tim R Pattison. Neural network for quadratic optimization with bound constraints. *IEEE transactions on neural networks*, 4(2):293–304, 1993.

[26] Jigen Peng, Z-B Xu, Hong Qiao, and Bo Zhang. A critical analysis on global convergence of hopfield-type neural networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(4):804–814, 2005.

[27] YJ Cao and QH Wu. A note on stability of analog neural networks with time delays. *IEEE Transactions on Neural Networks*, 7(6):1533–1535, 1996.

[28] Hongtao Lu. On stability of nonlinear continuous-time neural networks with delays. *Neural Networks*, 13(10):1135–1143, 2000.

[29] Mauro Forti, Stefano Manetti, and Mauro Marini. Necessary and sufficient condition for absolute stability of neural networks. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 41(7):491–494, 1994.

[30] Tianguang Chu, Zongda Zhang, and Zhaolin Wang. A decomposition approach to analysis of competitive–cooperative neural networks with delay. *Physics Letters A*, 312(5-6):339–347, 2003.

[31] Eugenius Kaszkurewicz and Amit Bhaya. On a class of globally stable neural circuits. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 41(2):171–174, 1994.

[32] AM Albano, A Passamante, T Hediger, and Mary Eileen Farrell. Using neural nets to look for chaos. *Physica D: Nonlinear Phenomena*, 58(1-4):1–9, 1992.

[33] JB Elsner. Predicting time series using a neural network as a method of distinguishing chaos from noise. *Journal of Physics A: Mathematical and General*, 25(4):843, 1992.

[34] Steve Renals and Richard Rohwer. A study of network dynamics. *Journal of Statistical Physics*, 58(5-6):825–848, 1990.

[35] KL Babcock and RM Westervelt. Dynamics of simple electronic neural networks. *Physica D: Nonlinear Phenomena*, 28(3):305–316, 1987.

[36] YG Zheng and LJ Bao. Slow–fast dynamics of tri-neuron hopfield neural network with two timescales. *Communications in nonlinear science and numerical simulation*, 19(5):1591–1599, 2014.

[37] Yan Huang and Xiao-Song Yang. Hyperchaos and bifurcation in a new class of four-dimensional hopfield neural networks. *Neurocomputing*, 69(13-15):1787–1795, 2006.

[38] Peter G Wolynes, Jose N Onuchic, D Thirumalai, et al. Navigating the folding routes. *SCIENCE-NEW YORK THEN WASHINGTON-*, pages 1619–1619, 1995.

[39] Aleksandr O Lykhin, Danil S Kaliakin, Gwen E dePolo, Alexander A Kuzubov, and Sergey A Varganov. Nonadiabatic transition state theory: Application to intersystem crossings in the active sites of metal-sulfur proteins. *International Journal of Quantum Chemistry*, 116(10):750–761, 2016.

[40] Jin Wang, Li Xu, and Erkang Wang. Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proceedings of the National Academy of Sciences*, 105(34):12271–12276, 2008.

[41] HK Janssen. Field-theoretic method applied to critical dynamics. In *Dynamical critical phenomena and related topics*, pages 25–47. Springer, 1979.

[42] Joseph Xu Zhou, MDS Aliyu, Erik Aurell, and Sui Huang. Quasi-potential landscape in complex multi-stable systems. *Journal of the Royal Society Interface*, 9(77):3539–3553, 2012.

[43] H Haken. Exact stationary solution of a fokker-planck equation for multimode laser action including phase locking. *Zeitschrift für Physik A Hadrons and nuclei*, 219(3):246–268, 1969.

[44] R Graham and H Haken. Generalized thermodynamic potential for markoff systems in detailed balance and far from thermal equilibrium. *Zeitschrift für Physik A Hadrons and nuclei*, 243(3):289–302, 1971.

[45] R Graham and H Haken. Fluctuations and stability of stationary non-equilibrium systems in detailed balance. *Zeitschrift für Physik A Hadrons and nuclei*, 245(2):141–153, 1971.

[46] R Graham and T Tél. On the weak-noise limit of fokker-planck models. *Journal of statistical physics*, 35(5-6):729–748, 1984.

[47] Hu Gang. Master equation without strict detailed balance. *Physica A: Statistical Mechanics and its Applications*, 132(2-3):586–592, 1985.

[48] Bo Han and Jin Wang. Quantifying robustness and dissipation cost of yeast cell cycle network: The funneled energy landscape perspectives. *Biophysical journal*, 92(11):3755–3763, 2007.

[49] Ping Ao. Potential in stochastic differential equations: novel construction. *Journal of physics A: mathematical and general*, 37(3):L25, 2004.

[50] Josiah Willard Gibbs. *The scientific papers of J. Willard Gibbs*, volume 1. Longmans, Green and Company, 1906.

[51] Martin J Klein and Paul HE Meijer. Principle of minimum entropy production. *Physical Review*, 96(2):250, 1954.

[52] Ilya Prigogine. Etude thermodynamique des phénomènes irréversibles. 1947.

[53] H Ziegler. An introduction to thermomechanics. 1983 amsterdam. *The Netherlands: North-Holland*.

[54] Leonid M Martyushev and Vladimir D Seleznev. Maximum entropy production principle in physics, chemistry and biology. *Physics reports*, 426(1):1–45, 2006.

[55] David Enskog. Kinetische theorie der vorgänge in mässig verdünnten gasen. 1917.

[56] Emil John Hellund and EA Uehling. Transport phenomena in mixtures of gases. *Physical Review*, 56(8):818, 1939.

[57] Joel H Ferziger, Hans G Kaper, and Hans G Kaper. *Mathematical theory of transport processes in gases*. North-Holland, 1972.

[58] JM Ziman. The general variational principle of transport theory. *Canadian Journal of Physics*, 34(12A):1256–1273, 1956.

[59] Hisashi Ozawa, Atsumu Ohmura, Ralph D Lorenz, and Toni Pujol. The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle. *Reviews of Geophysics*, 41(4), 2003.

[60] Edwin T Jaynes. Information theory and statistical mechanics. *Physical review*, 106(4):620, 1957.

[61] Edwin T Jaynes. Information theory and statistical mechanics. ii. *Physical review*, 108(2):171, 1957.

[62] Roderick Dewar. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *Journal of Physics A: Mathematical and General*, 36(3):631, 2003.

[63] Roderick C Dewar. Maximum entropy production and the fluctuation theorem. *Journal of Physics A: Mathematical and General*, 38(21):L371, 2005.

[64] Wassim Haddad, Qing Hui, and James Bailey. Human brain networks: Spiking neuron models, multistability, synchronization, thermodynamics, maximum entropy production, and anesthetic cascade mechanisms. *Entropy*, 16(7):3939–4003, 2014.

[65] Fang-Chin Yeh, Aonan Tang, Jon Hobbs, Pawel Hottowy, Wladyslaw Dabrowski, Alexander Sher, Alan Litke, and John Beggs. Maximum entropy approaches to living neural networks. *Entropy*, 12(1):89–106, 2010.

[66] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

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

[68] Hannes Risken. Fokker-planck equation. In *The Fokker-Planck Equation*, pages 63–95. Springer, 1996.

[69] Aleksandr Mikhailovich Lyapunov. The general problem of the stability of motion. *International journal of control*, 55(3):531–534, 1992.

[70] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on progress in physics*, 75(12):126001, 2012.

[71] Ken Sekimoto. *Stochastic energetics*, volume 799. Springer, 2010.

[72] Udo Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Physical review letters*, 95(4):040602, 2005.

[73] Hong Qian. Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. *Physical Review E*, 65(1):016102, 2001.

[74] Leonid M Martyushev and Vladimir D Seleznev. The restrictions of the maximum entropy production principle. *Physica A: Statistical Mechanics and its Applications*, 410:17–21, 2014.

[75] Han Yan, Lei Zhao, Liang Hu, Xidi Wang, Erkang Wang, and Jin Wang. Nonequilibrium landscape theory of neural networks. *Proceedings of the National Academy of Sciences*, 110(45):E4185–E4194, 2013.

[76] Wei Wu and Jin Wang. Potential and flux field landscape theory. i. global stability and dynamics of spatially dependent non-equilibrium systems. *The Journal of chemical physics*, 139(12):09B620_1, 2013.

[77] Chunhe Li and Jin Wang. Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. *Proceedings of the National Academy of Sciences*, 111(39):14130–14135, 2014.

Office use only: Date/Time stamp

KING'S College LONDON

King's College London

Department of Mathematics

Plagiarism Statement for Coursework

The following statement must be signed by all students submitting dissertations, projects or essays as a part of their formal assessments for any degree.

Student Name: Jacob Delveaux

Module Code: 7CCMNE07

Title of Project/Dissertation: Global Stability and Entropy Production for a Stochastic Neural Network Model

PLAGIARISM -

You are reminded that all work submitted as part of the requirements for any examination of the College or University of London must be expressed in your own words and incorporate your own ideas and judgements.

Plagiarism, that is the presentation of another person's thoughts or words as though they were your own, must be avoided, with particular care in coursework and essays and reports written in your own time.

Direct quotations from the published or unpublished work of others must always be clearly identified as such by being placed inside quotation marks, and a full reference to their source must be provided in the proper form. Failure to observe these rules may result in an allegation of cheating.

This includes repetition of your own work, if the fact that the work has been or is to be presented elsewhere (especially if it has already been presented for assessment) is not clearly stated.

DECLARATION:

I have read and understood the above statement and the College's statement on Academic Honesty and Integrity. I declare that the content of this submission is my own work.

I understand that plagiarism is a serious examination offence, an allegation of which can result in action being taken under the College's Misconduct regulations.

Signature:

Date: 16/04/19

Please submit one copy of the plagiarism declaration with your project/essay

Word count
Words in text: 4366
Words in headers: 82
Words outside text (captions, etc.): 350
Number of headers: 25
Number of floats/tables/figures: 5
Number of math inlines: 90
Number of math displayed: 37
Subcounts:

text+headers+captions (#headers/#floats/#inlines/#displayed)

143+13+0 (3/0/0/0) _top_
369+1+0 (1/0/0/0) Chapter: Introduction}\label{chap:Introduction
0+1+0 (1/0/0/0) Chapter: Background}\label{chap:Review
254+2+0 (1/0/14/1) Section: Model Description}\label{sec:model
40+3+0 (1/0/0/0) Section: Review of Literature}\label{sec:LitReview
121+1+0 (1/0/2/1) Subsection: Stability}\label{sec:StabilityReview
124+2+0 (1/0/0/0) Subsection: Chaotic Behavior}\label{sec:ChaosReview
322+2+0 (1/0/3/0) Subsection: Probability Landscape}\label{sec:ProbLandscapeReview
251+5+0 (1/0/0/0) Subsection: Entropy Production in Non-Equilibrium Systems}\label{sec:EnropyProdReview
83+1+0 (1/0/0/0) Chapter: Methods}\label{chap:Methods&Results
323+5+0 (1/0/20/3) Section: Probability Landscape and Flux Formalism}\label{sec:formalism
172+7+0 (1/0/3/6) Section: Lyapunov Stability in the Low Fluctuation Limit}\label{sec:lyapunov1
60+3+0 (1/0/0/0) Section: Non-Equilibrium Stochastic Thermodynamics}\label{sec:NonEqmThermo
195+3+0 (1/0/5/7) Subsection: Entropy Production Rate}\label{sec:EPR
146+6+0 (1/0/9/8) Subsection: Free Energy as a Lyapunov Function}\label{sec:free_energy
64+6+0 (1/0/1/1) Section: Relative Entropy as a Lyapunov Function}\label{sec:relative_entropy
145+8+0 (1/0/6/9) Section: Exact solution for steady-state probability with linear response}\label{sec:analytical
94+1+0 (1/0/1/0) Chapter: Results}\label{chap:results&disc
363+2+90 (1/1/11/1) Section: Probability Landscape}\label{results:prob_land
226+2+0 (1/0/0/0) Section: Lyapunov Stability
482+2+194 (1/3/6/0) Section: Entropy Production
172+5+66 (1/1/9/0) Section: A Priori Steady State Distribution
217+1+0 (1/0/0/0) Chapter: Conclusion}\label{chap:conclusion