Synchronization on Second Order Networks

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Liqiong Zhao

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Duane Nykamp

October, 2012

© Liqiong Zhao 2012
ALL RIGHTS RESERVED

Acknowledgements

I would like to express my sincere gratitude to my thesis advisor Duane Nykamp for his support, encouragement and guidance on my research.
I would like to thank Chin-Yueh Liu and Patrick Campbell for helpful discussions.

Dedication

Dedicated to my parents.

Abstract

Networks are complex and high dimensional. To describe either network activity or connectivity requires many dimensions. It is important to extract key features of the network activity and connectivity, and find lower dimensional descriptions of networks.

We propose a framework of SONETs (second order networks), which defines four second order connectivity statistics based on the relative frequency of two-connection network motifs. We generate SONETs using Dichotomized Gaussian.

We study how the second order statistics influence the network properties such as visualization, degree distribution, average path length, clustering coefficients, assortativity and spectrum. We also investigate how network structure can influence the tendency for a neuronal network to synchronize, independent of the dynamic model of each neuron. The synchrony analysis identifies two of the statistics, convergent connections and chain connections, as highly influencing the synchrony. Simulations verify that synchrony increases with the frequency of chain connections and decreases with the frequency of convergent connections.

These trends persist with simulations of multiple models for the neuron dynamics and for different types of networks. We found that divergent connections, which determine the fraction of shared inputs, do not strongly influence the synchrony. The critical role of chains in influencing synchrony can be explained by their increasing the effective coupling strength. The decrease of synchrony with convergent connections is due to the resulting heterogeneity in firing rates.

In addition, we developed a kinetic theory approach that describes dynamics of interacting neuronal populations. We focused on all-to-all coupled network, in which each neuron has a connection to any other neurons in the network. We derived the third order truncation for kinetic theory equation of neural networks, and approximate the full probability density to make the third order truncation closed.

Contents

List of Figures

Chapter 1

Introduction

Networks are notable for their complicated descriptions. To describe either their activity or connectivity requires many dimensions. In this thesis, we extract key features of the network activity and connectivity, provide a low-dimensional description of networks. We propose a framework of Second Order Network model, which is based on the relative frequency of two-connection network motifs.

This thesis is organized into the following chapters.

- Chapter 2 presents a framework of second order network model, as well as its simulation methods and visualization..
- In Chapter 3, we demonstrates the properties of second order network model, such as degree distribution, avearge shortest path, clustering coefficients and assortativity.
- In Chapter 4, we investigate how the network connectivity statistics affect the syncrhonization of networks.
- In Chapter 5, we developed a kinetic theory approach modeling all-to-all coupled Integrate-and-fire neuron, and investigate the higher order statistics of the firing pattern.

Chapter 2

Second Order Networks

2.1 Background

Complex networks are gaining increasing attention these years. Broad applications such as neural networks, airline networks, Biochemical networks, the Internet, the World Wide Web and social networks make it important to investigate the topology and the mechanisms of dynamics on complex networks. For instance, due to the emergence of large scale social networks as well as various activities on them, the study of the prevalence of certain ideas, technologies or products over the network becomes important.

The traditional study of random networks focuses on graphs where the connections among two nodes are independent from the connections among others. The probability of connection might depend on identity, location or other properties of nodes. In 1959, Paul Erdős and Alfréd Rényi[13] studied a undirected random graph model, in which the connection between any pair of vertices is an independent Bernoulli random variable in which the connection is present with probability p. Thus the expected number of connections in a random network of N vertices is $N(N-1)p/2$. For directed Erdős-Rényi model, the pairwise connectivity is also a independent Bernoulli distribution with a single parameter. The Erdős-Rényi network is characterized by a single constant p, which describes the first order statistic of the random network. The independent random network model is the natural model that results from specifying only the connection

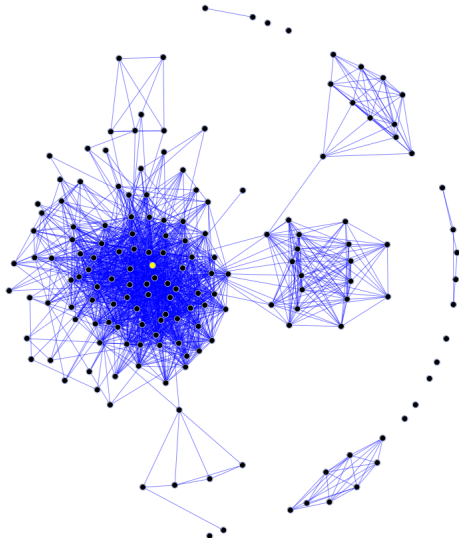


Figure 2.1: An example of a social network diagram

probabilities, as it adds minimal structure beyond what is required to by the connection probabilities [18].

Let $G = (V, E)$ denote a network G, where V is the set of vertices and E is the set of all connections among them. Let W be the adjacency matrix of network G. $W_{ij} = 1$ if there is a connection from vertex j to vertex i and $W_{ij} = 0$ otherwise, where the vertices are labeled from 1 to the total number of vertices N.

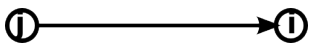


Figure 2.2: The connection from vertex j to i with probability p

However, many complex networks exhibit features that are not described by the Erdős-Rényi network. For local cortical circuits, bidirectional connections are more common than expected in Erdős-Rényi random networks and several three-node connectivity patterns are overrepresented compared to E-R random networks[46]. Perin et al.[37] also find that the frequency of other patterns are different from independent model when they were investigating synaptic connectivity in groups of pyramidal neurons in the neocortex.

The independent random network model is based only on the first order statistics of the network connectivity, i.e.,the marginal probability any given connection exists, while ignoring the higher order statistics, such as covariance of connections. To capture these features of correlations of network connectivity, random network models with higher order statistics are needed.

2.2 Second Order Networks

A notable feature of complex networks is their complicated description. To describe either network activity or connectivity requires a huge number of dimensions. Our objective is to extract key features of the network activity and connectivity to obtain a useful low-dimensional description.

Let's start with E-R networks, which are specified by a single parameter p, and add more structures to them. We would like to find a lower dimensional features of networks that could describe their structures. There are different methods of adding additional structures to random networks and obtaining lower dimensional representations. The standard approach is via the degree distribution(see section 3.1). For instance, given a probability distribution $p(x, y)$ for the degree of the node, where x refers to the number of in-coming edges and y refers to the number of out-going edges, a network can be generated such that the nodes have the degree distribution $p(x, y)$. distribution $p(x, y)$. However, the degree distribution makes the network representation high dimensional. Another approach of low dimensional representation of networks is to add statistics of three-node patterns in random graph[46]. There are sixteen motifs that include connections to all three nodes(Figure 2.3), which also provides a seventeen dimensional representation of networks.

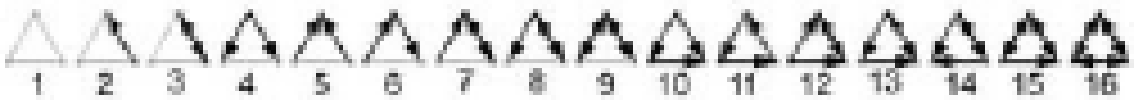


Figure 2.3: Three-node motifs[46]

We would like to privide a lower dimensional description of networks with prescribed first and second order statistics of network connectivity. We assume the network is homogeneous in the sense that the statistics of the connections do not depend on the indices of the nodes involved. We don't allow self conncetions, so $W_{ii} = 0$ for $i = 1 \dots N$. The first order statistic p is used to characterize the probability of connection between any two nodes of the network:

$$\Pr(W_{ij} = 1) = p \tag{2.1}$$

for any nodes i and j. Thus the expected value of W_{ij} is

$$E(W_{ij}) = 1\Pr(W_{ij} = 1) + 0\Pr(W_{ij} = 0) = p \tag{2.2}$$

For a random network with N vertices, the maximum number of connections is $N(N-1)$. If the first order statistic p is given, the expected number of connections in the network is $N(N-1)p$.

In our model, we add second order statistics to the network connections for two-edge motifs. In addition to retaining the first order statistics p, we define four second order statistics for each pair of connections that share at least one common node. If we let any pair of connections be independent, we obtain the Erdős-Rényi network, which is the most straightforward realization of random networks. In this case, $\Pr(W_{ij} = 1, W_{kl} = 1) = p^2$ for any $i \neq j$ and $k \neq l$.

Now we vary the dependence of the pair of connections that share at least one single node. The two edge connectivity pattern can be classified into four groups, i.e., reciprocal motif (figure 2.4), convergent motif (figure 2.5), divergent motif (figure 2.6) and chain motif (figure 2.7). The first case is the reciprocal motif, in which the reciprocal edges W_{ij} and W_{ji} coexist.

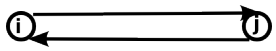


Figure 2.4: Reciprocal motif

Let the deviation from the independence be defined by α_{recip}, and

$$\Pr(W_{ij} = 1, W_{ji} = 1) = p^2(1 + \alpha_{\text{recip}}) \tag{2.3}$$

for any pair of vertices i and j. If $\alpha_{\text{recip}} > 0$, the probability of coexistence of connections W_{ij} and W_{ji} is bigger compared to Erdős-Rényi network with the same first order statistic. Else if $\alpha_{\text{recip}} < 0$, the concurrence of the reciprocal motif is relatively smaller. Also, we could obtain that

$$\begin{aligned}\Pr(W_{ij} = 1, W_{ji} = 0) &= \Pr(W_{ij} = 1) - \Pr(W_{ij} = 1, W_{ji} = 1) \\ &= p - p^2(1 + \alpha_{\text{recip}}) \end{aligned} \tag{2.4}$$

by (2.3) and

$$\begin{aligned}\Pr(W_{ij} = 0, W_{ji} = 0) &= 1 - \Pr(W_{ij} = 1, W_{ji} = 1) - \Pr(W_{ij} = 1, W_{ji} = 0) - \Pr(W_{ij} = 0, W_{ji} = 1) \\ &= (1 - p)^2 + p^2\alpha_{\text{recip}} \end{aligned} \tag{2.5}$$

equation (2.5) indicates that when $\alpha_{\text{recip}} > 0$, the probability that there is no connection among two nodes is relatively larger compared to the case that $\alpha_{\text{recip}} = 0$, in which the probability that both the reciprocal connections do not exist equals $(1 - p)^2$.

The expected value of the reciprocal motif is

$$\begin{aligned}E(W_{ij}W_{ji}) &= 1\Pr(W_{ij} = 1, W_{ji} = 1) + 0(1 - \Pr(W_{ij} = 1, W_{ji} = 1)) \\ &= p^2(1 + \alpha_{\text{recip}}) \end{aligned} \tag{2.6}$$

For a network of N nodes, there are $\binom{N}{2}$ pairs of vertices, thus there are at most $N(N - 2)/2$ reciprocal motifs. Given the second order statistics α_{recip}, the expected number of reciprocal motifs in the network is $N(N - 1)p^2(1 + \alpha_{\text{recip}})/2$.

Let's take a look at an example in which the α_{recip} attains its maximum value. Assume that if $W_{ij} = 1$, its reciprocal edge $W_{ji} = 1$. By equations (2.1) and (2.3), we obtain $p = p^2(1 + \alpha_{\text{recip}})$. Thus the maximum value α_{recip} could obtain is $1/p - 1$.

The second motif is the convergent motif, which exists if there is a common in-coming node for two connections. It can be illustrated in figure (2.5): nodes j and k connect to node i.

The deviation from the independence of the two edges W_{ij} and W_{ik} are defined by α_{conv}, and

$$\Pr(W_{ij} = 1, W_{ik} = 1) = p^2(1 + \alpha_{\text{conv}}) \tag{2.7}$$

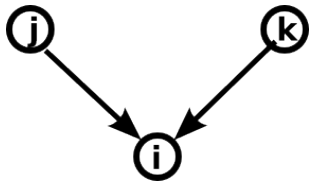


Figure 2.5: Convergent motif

for any triple $i \neq j \neq k$. It demonstrates that larger α_{conv} results in higher probability of the convergent motif. Similar to equation (2.5), we obtain

$$\Pr(W_{ij} = 0, W_{ik} = 0) = (1-p)^2 + p^2\alpha_{\text{conv}} \tag{2.8}$$

which shows that for $\alpha_{\text{conv}} > 0$, the probability that two edges connecting to a single node do not exist is large compared to that of the E-R network.

The expected value of the number of convergent motif with the in-coming node i is

$$\begin{aligned} E(W_{ij}W_{ik}) &= 1\Pr(W_{ij} = 1, W_{ik} = 1) + 0(1 - \Pr(W_{ij} = 1, W_{ik} = 1)) \\ &= p^2(1 + \alpha_{\text{conv}}) \end{aligned} \tag{2.9}$$

There are at most $\binom{N-1}{2}$ convergent motifs for a given vertex as the in-coming node. Thus the largest possible number of convergent motifs in network with size N is $N(N-1)(N-2)/2$. The expected number of convergent motifs in the network is $N(N-1)(N-2)p^2(1+\alpha_{\text{conv}})/2$.

Similarly, the two-edge motif where the connections share the same out-going node is defined as the divergent motif, which is illustrated in (2.6):

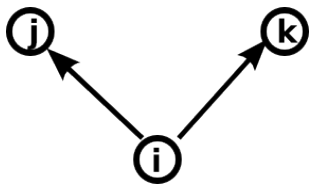


Figure 2.6: Divergent motif

The corresponding second order parameter α_{div} is

$$\Pr(W_{ji} = 1, W_{ki} = 1) = p^2(1 + \alpha_{\text{div}}) \tag{2.10}$$

for any triple $i \neq j \neq k$.

The expected value of the divergent motif is

$$\begin{aligned} E(W_{ji}W_{ki}) &= 1\Pr(W_{ji} = 1, W_{ki} = 1) + 0(1 - \Pr(W_{ji} = 1, W_{ki} = 1)) \\ &= p^2(1 + \alpha_{\text{div}}) \end{aligned} \tag{2.11}$$

Similar to the convergent motifs, there exist at most $N(N-1)(N-2)/2$ divergent motifs for any directed network with size N. Also, the expected number of divergent motifs in the network with given statistic p and α_{div} is $N(N-1)(N-2)p^2(1+\alpha_{\text{div}})/2$.

The fourth motif is the chain motif, which occurs when the in-coming node of one edge is the out-going node of another and the two connections are not reciprocal:

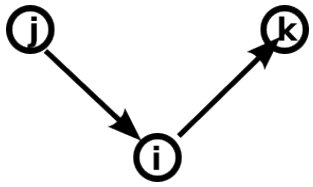


Figure 2.7: Chain motif

$$\Pr(W_{ij}W_{ki} = 1) = p^2(1 + \alpha_{\text{chain}}) \tag{2.12}$$

for any triple $i \neq j \neq k$.

The expected value of the chain motif is

$$\begin{aligned} E(W_{ij}W_{ki}) &= 1\Pr(W_{ij} = 1, W_{ki} = 1) + 0(1 - \Pr(W_{ij} = 1, W_{ki} = 1)) \\ &= p^2(1 + \alpha_{\text{chain}}) \end{aligned} \tag{2.13}$$

and the expected number of chain motifs of a random network with statistics p and α_{chain} is $N(N-1)(N-2)p^2(1+\alpha_{\text{chain}})$.

If the two connections don't share a single node, we assume that their covariance is 0. By (2.6),(2.9),(2.11) and (2.13), we obtain that the correlation of the two edges sharing one node is exactly the corresponding second order statistic α.

$$\alpha_{\text{recip}} = \frac{\text{cov}(W_{ij}W_{ji})}{\text{E}(W_{ij})\text{E}(W_{ji})} \tag{2.14}$$

$$\alpha_{\text{conv}} = \frac{\text{cov}(W_{ij}W_{ik})}{\text{E}(W_{ij})\text{E}(W_{ik})} \tag{2.15}$$

$$\alpha_{\text{div}} = \frac{\text{cov}(W_{ji}W_{ki})}{\text{E}(W_{ji})\text{E}(W_{ki})} \tag{2.16}$$

$$\alpha_{\text{chain}} = \frac{\text{cov}(W_{ij}W_{ki})}{\text{E}(W_{ij})\text{E}(W_{ki})} \tag{2.17}$$

Positive α means the positive correlation of the occurence of connections.

The second order network is a maximum entropy model, which maximize the entropy of a five-dimensional probability distribution(parameterized by p and four α's) for the adjacency matrix that satisfies conditions (2.1),(2.3),(2.7),(2.10),(2.12). As the probability distribution adds minimal structure(i.e.,higher order correlations) beyond these constraints, the SONET reduces to the Erdős-Rényi model when all α's are zero.

2.2.1 Estimating Connectivity Statistics

The above definition of the first order and second order connectivity statistics are based on a probability distribution for the adjacency matrix W. From a given network connectivity matrix W, we could estimate the five statistics directly. Denote the statistics obtained from W by $\hat{p}$, $\widehat{\alpha}_{\text{recip}}$,$\widehat{\alpha}_{\text{conv}}$,$\widehat{\alpha}_{\text{div}}$ and $\widehat{\alpha}_{\text{chain}}$ respectively. $\hat{p}$ denotes the mean first order statistic, which is the ratio of number of edges present in W and all the possible number of connection:

$$\hat{p} = N_{\text{edge}} / \left(N(N-1)\right) \tag{2.18}$$

where $N_{\text{edge}} = \sum_{i \neq j} W_{ij} = \|W\|_1$.

Define $\widehat{\alpha}_{\text{recip}}$ such that $\hat{p}^2(1+\widehat{\alpha}_{\text{recip}})$ describes the fraction of reciprocal motifs present in the network W, which is the ratio of number of reciprocals motifs present in W and all the possible reciprocal motifs that could be present in a network of size N. Thus $\hat{p}^2(1+\widehat{\alpha}_{\text{recip}}) = \frac{N_{\text{recip}}}{N(N-1)/2}$ where N_{recip} is the number of reciprocal motifs present in W.

We can obtain

$$\widehat{\alpha}_{\text{recip}} = \frac{N_{\text{recip}}}{\hat{p}^2 N(N-1)/2} - 1 \tag{2.19}$$

N_{recip} can be calculated directly from W:

$$N_{\text{recip}} = \sum_{i \neq j} W_{ij} W_{ji} = \text{Tr}(W^2)/2 \tag{2.20}$$

where $\text{Tr}(W)$ denote trace.

The average second order statistics can likewise be determined by the total number of corresponding two-connection motifs that are present in the network, divided by the total possible associated motifs. Then we can find the estimated second order statistics:

$$\widehat{\alpha}_{\text{conv}} = \frac{N_{\text{conv}}}{\hat{p}^2 N(N-1)(N-2)/2} - 1 \tag{2.21}$$

$$\widehat{\alpha}_{\text{div}} = \frac{N_{\text{div}}}{\hat{p}^2 N(N-1)(N-2)/2} - 1 \tag{2.22}$$

$$\widehat{\alpha}_{\text{chain}} = \frac{N_{\text{chain}}}{\hat{p}^2 N(N-1)(N-2)} - 1 \tag{2.23}$$

where $N_{\text{conv}}, N_{\text{div}}, N_{\text{chain}}$ refer to the number of corresponding second order motifs present in W. The number of convergent motifs with in-coming node i is $\frac{1}{2}\sum_{j \neq k} W_{ij} W_{ik}$. thus we have

$$N_{\text{conv}} = \sum_{i=1}^{N} \sum_{j \neq k} W_{ij} W_{ik}/2 = \|W^T W\|_1/2 - \|W\|_1/2 \tag{2.24}$$

The number of divergent motifs with out-going node i is $\frac{1}{2}\sum_{j \neq k} W_{ji} W_{ki}$. We obtain

$$N_{\text{div}} = \|W W^T\|_1/2 - \|W\|_1/2 \tag{2.25}$$

Similarly, the number of chain motifs with center at node i is $\sum_{j \neq k} W_{ij} W_{ki} = \sum_{j,k} W_{ij} W_{ki} - \sum_j W_{ij} W_{ji}$. Thus we have

$$N_{\text{chain}} = \|W^2\|_1 - \text{trace}(W^2) \tag{2.26}$$

From (2.18),(2.19),(2.20),(2.21),(2.22),(2.23),(2.24),(2.25) and (2.26), we are able to calculate the statistics of motifs from a given network W.

In our network simulations, we compare the measured $\hat{\alpha}$ to the network topological properties or synchrony even when we know the underlying probability distribution and hence the parameters p and α from (2.1),(2.7),(2.10),(2.12). Because these networks are generated randomly, the actual $\hat{p}$ and $\hat{\alpha}$ will vary slightly from α.

2.2.2 Maximum Entropy Model

Given the mean connectivity p and statistics of the four motifs $\alpha's$, we would like to construct random networks that have the expected mean connectivity p and corresponding second order statistic α's while minimizing any additional structure among the connections. Our approach is to use the maximum entropy method.

For the adjacency matrix W, there are $N(N-1)$ random variables with Bernoulli distributions. Thus the there are $2^{N(N-1)}$ states of W. Let $P(W)$ denote the probability distribution of W. The entropy H is defined by

$$H = -\sum_{W} P(W) \log P(W) \tag{2.27}$$

where $P(W)$ is the probability distribution of the $N(N-1)$ dimensional vector, and the sum is over all possible networks. We have

$$\sum_{W} P(W) = 1 \tag{2.28}$$

By equation (2.1), for each pair (i, j) with $i \neq j$, we have

$$E(W_{ij}) = \sum_{W} W_{ij} P(W) = p \tag{2.29}$$

Thus the probability distribution must satisfy the form of equation (2.29).

By equation (2.6), for nodes i, j with $i \neq j$, we have

$$E(W_{ij}W_{ji}) = \sum_{W} W_{ij}W_{ji} P(W) = p^2(1 + \alpha_{\text{recip}}) \tag{2.30}$$

There are $N(N-1)/2$ constraints of the form of equation (2.30).

By equation (2.9), for distinct triple (i, j, k), we have

$$E(W_{ij}W_{ik}) = \sum_{W} W_{ij}W_{ik} P(W) = p^2(1 + \alpha_{\text{conv}}) \tag{2.31}$$

There are $N(N-1)(N-2)/2$ equations for (2.31), given that the expression is symmetric in j and k.

Similarly, the expectation of divergent motif is $p^2(1 + \alpha_{\text{div}})$. We have $N(N-1)(N-2)/2$ constraints:

$$E(W_{ji}W_{ki}) = \sum_{W} W_{ji}W_{ki} P(W) = p^2(1 + \alpha_{\text{div}}) \tag{2.32}$$

There are $N(N-1)(N-2)$ constraints for chains. For distinct triple (i,j,k),

$$E(W_{ij}W_{ki}) = \sum_W W_{ij}W_{ki}P(W) = p^2(1+\alpha_{\text{chain}}) \tag{2.33}$$

We would like to maximize (2.27) subject to the constraints (2.29)(2.30) (2.31)(2.32) and (2.33). The corresponding Lagrange function is

$$\begin{aligned}
&\Lambda(P(w),\mu,\lambda_{ij},\eta_{ijji},\eta_{ijik},\eta_{jiki},\eta_{ijki}) \\
&= -\sum_W P(W)\log P(W) + \mu(\sum_W P(W) - 1) + \sum_{i\neq j}\lambda_{ij}(\sum_W W_{ij}P(W) - p) \\
&+ \sum_{i\neq j}\eta_{ijji}(\sum_W W_{ij}W_{ji}P(W) - p^2(1+\alpha_{\text{recip}})) \\
&+ \sum_{j<k}\eta_{ijik}(\sum_W W_{ij}W_{ik}P(W) - p^2(1+\alpha_{\text{conv}})) \\
&+ \sum_{j<k}\eta_{jiki}(\sum_W W_{ji}W_{ki}P(w) - p^2(1+\alpha_{\text{div}})) \\
&+ \sum_{i\neq j\neq k}\eta_{ijki}(\sum_W W_{ij}W_{ki}P(w) - p^2(1+\alpha_{\text{chain}}))
\end{aligned} \tag{2.34}$$

where $\mu, \lambda_{ij}, \eta_{ijji}, \eta_{ijik}, \eta_{jiki}$ and η_{ijki} are Lagrange multipliers. λ_{ij} corresponds to the constraint for W_{ij}, thus there are $N(N-1)$ different λ_{ij}. Similarly, the Lagrange multipliers $\eta_{ij\tilde{i}\tilde{j}}$ corresponds to the interactions of W_{ij} and $W_{\tilde{i}\tilde{j}}$.

In order to find the maxima to H subject to all the constraints, we should find the

critical values of the Lagrangian function with respect to all its variables:

$$
\begin{aligned}
0 &= \frac{\partial \Lambda}{\partial P(W)} = -1 - \log P(W) + \mu + \sum_{i \neq j} \lambda_{ij} W_{ij} + \sum_{i \neq j} \eta_{ijji} W_{ij} W_{ji} \\
&+ \sum_{j<k} \eta_{ijik} W_{ij} W_{ik} + \sum_{j<k} \eta_{jiki} W_{ji} W_{ki} + \sum_{i \neq j \neq k} \eta_{ijki} W_{ij} W_{ki} \\
0 &= \frac{\partial \Lambda}{\partial \mu} = \sum_{W} \Pr(W) - 1 \\
0 &= \frac{\partial \Lambda}{\partial \lambda_{ij}} = \sum_{W} W_{ij} P(W) - p \\
0 &= \frac{\partial \Lambda}{\partial \eta_{ijji}} = \sum_{W} W_{ij} W_{ji} P(W) - p^2 (1 + \alpha_{\text{recip}}) \\
0 &= \frac{\partial \Lambda}{\partial \eta_{ijik}} = \sum_{W} W_{ij} W_{ik} P(W) - p^2 (1 + \alpha_{\text{conv}}) \\
0 &= \frac{\partial \Lambda}{\partial \eta_{jiki}} = \sum_{W} W_{ji} W_{ki} P(W) - p^2 (1 + \alpha_{\text{div}}) \\
0 &= \frac{\partial \Lambda}{\partial \eta_{ijki}} = \sum_{W} W_{ij} W_{ki} P(W) - p^2 (1 + \alpha_{\text{chain}})
\end{aligned} \tag{2.35}
$$

Differentiating Λ with respect to $\mu, \lambda_{ij}, \eta_{ijji}, \eta_{ijik}, \eta_{jiki}, \eta_{ijki}$, we obtain equations (2.28),(2.29),(2.31),(2.32) and (2.33). Differentiate Λ with respect to $P(W)$, we obtain

$$
\begin{aligned}
P(W) = \exp(\mu - 1 &+ \sum_{i \neq j} \lambda_{ij} W_{ij} + \sum_{i \neq j} \eta_{ijji} W_{ij} W_{ji} + \sum_{j<k} \eta_{ijik} W_{ij} W_{ik} \\
&+ \sum_{j<k} \eta_{jiki} W_{ji} W_{ki} + \sum_{i \neq j \neq k} \eta_{ijki} W_{ij} W_{ki})
\end{aligned} \tag{2.36}
$$

By the homogeneity of second order networks, we know that λ_{ij} have the same value for each pair (i, j) with $i \neq j$. Let λ_{edge} denote λ_{ij}. Thus $\sum_{i \neq j} \lambda_{ij} W_{ij} = \lambda_{edge} N_{\text{edge}}(W)$ where $N_{\text{edge}}(W)$ is the total number of connections in W.. Also, η_{ijji} have the same value for each distinct pair (i, j). Define $\eta_{\text{recip}} = \eta_{ijji}$, we have $\sum_{i \neq j} \eta_{ijji} W_{ij} W_{ji} = \eta_{\text{recip}} N_{\text{recip}}(W)$. Similarly, by defining $\eta_{\text{conv}} = \eta_{ijik}$, $\eta_{\text{div}} = \eta_{jiki}$ and $\eta_{\text{chain}} = \eta_{jiki}$, we obtain $\sum_{j<k} \eta_{ijik} W_{ij} W_{ik} = \eta_{\text{recip}} N_{\text{conv}}(W)$, $\sum_{j<k} \eta_{jiki} W_{ji} W_{ki} = \eta_{\text{div}} N_{\text{div}}(W)$, and $\sum_{i \neq j \neq k} \eta_{ijki} W_{ij} W_{ki} = \eta_{\text{chain}} N_{\text{chain}}(W)$. $N_{\text{recip}}(W)$,$N_{\text{conv}}(W)$,$N_{\text{div}}(W)$ and $N_{\text{chain}}(W)$ are the number of associated second order motifs present in W.

Equation (2.36) can be rewritten as:

$$\begin{aligned} P(W) =& \exp(\mu - 1 + \lambda_{\text{edge}} N_{\text{edge}}(W) + \eta_{\text{recip}} N_{\text{recip}}(W) + \eta_{\text{conv}} N_{\text{conv}}(W) + \eta_{\text{div}} N_{\text{div}}(W) \\ & + \eta_{\text{chain}} N_{\text{chain}}(W)) \\ =& e^{\mu-1} \exp(\lambda_{\text{edge}} N_{\text{edge}}(W) + \eta_{\text{recip}} N_{\text{recip}}(W) + \eta_{\text{conv}} N_{\text{conv}}(W) + \eta_{\text{div}} N_{\text{div}}(W) \\ & + \eta_{\text{chain}} N_{\text{chain}}(W)) \end{aligned} \tag{2.37}$$

Since $P(W)$ is a probability distribution over all the possible adjacency matrix W, we can write $e^{\mu-1}$ in terms of the other parameters. Let Z denote the partition function:

$$Z = \sum_W \exp(\lambda_{\text{edge}} N_{\text{edge}}(W) + \eta_{\text{recip}} N_{\text{recip}}(W) + \eta_{\text{conv}} N_{\text{div}}(W) + \eta_{\text{div}} N_{\text{div}}(W) + \eta_{\text{chain}} N_{\text{chain}}(W)) \tag{2.38}$$

We can rewrite $P(W)$ as

$$P(W) = \frac{1}{Z} \exp(\lambda_{\text{edge}} N_{\text{edge}}(W) + \eta_{\text{recip}} N_{\text{recip}}(W) + \eta_{\text{conv}} N_{\text{conv}}(W) + \eta_{\text{div}} N_{\text{div}}(W) + \eta_{\text{chain}} N_{\text{chain}}(W)) \tag{2.39}$$

We could obtain the Lagrange multipliers λ_{edge}, η_{recip}, η_{conv}, η_{div} and η_{chain} by pluging the equation (2.39) into the last five equations of (2.35):

$$\begin{aligned} & p = \sum_W \frac{W_{ij}}{Z} \exp(\lambda_{\text{edge}} N_{\text{edge}}(W) + \eta_{\text{recip}} N_{\text{recip}}(W) + \eta_{\text{conv}} N_{\text{conv}}(W) + \eta_{\text{div}} N_{\text{div}}(W) \\ & + \eta_{\text{chain}} N_{\text{chain}}(W)) \\ & p^2(1 + \alpha_{ij\tilde{i}\tilde{j}}) = \exp(\lambda_{\text{edge}} N_{\text{edge}}(W) + \eta_{\text{recip}} N_{\text{recip}}(W) + \eta_{\text{conv}} N_{\text{conv}}(W) + \eta_{\text{div}} N_{\text{div}}(W) \\ & + \eta_{\text{chain}} N_{\text{chain}}(W)) \end{aligned} \tag{2.40}$$

where $\alpha_{ij\tilde{i}\tilde{j}}$ describes the four second order motifs. Equations (2.38) and (2.40) is a closed system of equations. Once the Lagrange multipliers are solved, the probability distribution for each W could be obtained by equation (2.39).

2.3 Simulation Methods

2.3.1 Relation with the Ising Model

Ising model provides a method representing correlated multivariate binary data. For binary variable s_i takes values on $\{-1, 1\}$, and give

Given an adjacency matrix W with N nodes, we can reshape it by a $(N-1)^2 \times 1$ vector V, where the components of v enumerate all the non-diagonal entries of W. For a SONET with given p and $\alpha's$, the expression of probability distribution $P(W)$ could be written as

$$P(v) = \frac{1}{Z} \exp\left(-(Hv + \frac{1}{2} v^T J v)\right) \tag{2.41}$$

where H is a $1 \times (N-1)^2$ vector such that each entry of H equals $-\lambda_{\text{edge}}$, and J is a $(N-1)^2 \times (N-1)^2$ matrix. $J_{\vec{i},\vec{j}} = -\eta_{\text{edge}}$ if the connection $\vec{i}$ and $\vec{j}$ form reciprocal motifs. $J_{\vec{i},\vec{j}}$ are set to be $-\eta_{\text{conv}}$, $-\eta_{\text{div}}$ and $-\eta_{\text{chain}}$, if $\vec{i}, \vec{j}$ form convergent motifs, divergent motif and chain, respectively. Otherwise $J_{\vec{i},\vec{j}} = 0$. Here, $\vec{i}$ and $\vec{j}$ are multiindices that enumerate all the non-diagonal entries of an $N \times N$ matrix. In this case, $Hv = -\eta_{\text{edge}} N_{\text{edge}}(W)$, and $\frac{1}{2} v^T J v = -\eta_{\text{edge}} N_{\text{edge}}(W) - \eta_{\text{conv}} N_{\text{conv}}(W) - \eta_{\text{div}} N_{\text{div}}(W) - \eta_{\text{chain}} N_{\text{chain}}$. Z is the partition function.

.

2.3.2 Dichotomized Gaussian Method

For the vector $V = (V_1, \ldots, V_N)$ of Bernoulli random variables with the same mean p and pairwise covariance $p^2 \alpha_{ij}$ for (V_i, V_j), we could approximate V by a vector $\mathbf{X}$. $\mathbf{X}$ is generated by a multivariate normal distribution with mean γ and covariance matrix Σ. Let $V_i = 1$ for $X_i \geq 0$ and $V_i = 0$ if $X_i < 0$. γ and Σ are chosen so that $1 \leq i, \leq j \leq N$,

$$\Pr(X_i \geq 0) = p \tag{2.42}$$

$$\Pr(X_i \geq 0, X_j \geq 0) - \Pr(X_i \geq 0) \mathrm{P}(X_j \geq 0) = p^2 \alpha_{ij} \tag{2.43}$$

which is equivalent to

$$\Phi(-\gamma) = p \tag{2.44}$$

$$\Phi_2(-\gamma, -\gamma, \rho_{ij}) = p^2(1 + \alpha_{ij}) \tag{2.45}$$

where $\Phi(x)$ is the cumulative distribution function of normal distribution $N(0,1)$, and $\Phi_2(x, y, \rho_{ij})$ is the cumulative distribution function of bivariate normal distribution with mean 0 and covariance matrix $\begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}$. By solving equation (2.45), we get γ and ρ_{ij}. Then the covariance matrix Σ has $\Sigma_{ij} = \rho_{ij}$.

In the maximum entropy model of second order networks of given first order statistic p and second order statistics $\alpha's$, the mean of the $N(N-1)$ binary random variables is p. The covariance matrix Σ has $\Sigma_{ij,\tilde{i}\tilde{j}} = \rho_{ij,\tilde{i}\tilde{j}}$, which is determined by p and $\alpha_{ij,\tilde{i}\tilde{j}}$. By the dichotomized method, the mean and covariance matrix of the multivariate normal distribution can be solved numerically. Then we have a more efficient way of generate the second order networks. The algorithm generating the second order networks is described in the Appendix.

2.4 Visualization of SONETs

By the simulation method introduced in (2.3.2), we could obtain samples of SONETs with prescribed first and second order statistics. Now let's study how the second order statistics influence the structure of networks.

First let's take a look at the Erdős-Rényi networks where all second order statistics are 0. Figure 2.8 shows the visualization of E-R model.

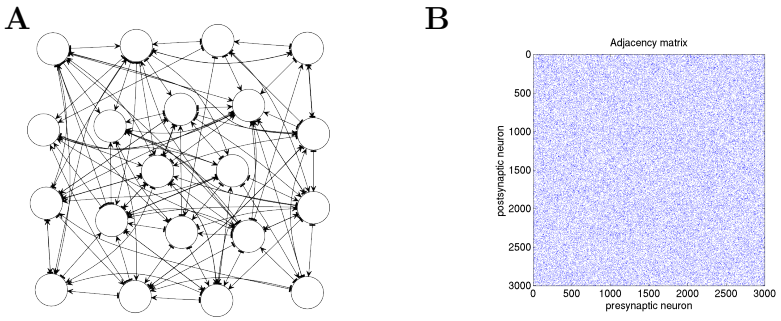


Figure 2.8: Visualization of Erdős-Rényi networks. (**A**)Sample E-R network generated by parameters $N = 20$, $p = 0.3$. (**B**) Sparsity pattern of sample E-R network generated by parameters $N = 3000$, $p = 0.01$.

Then we add reciprocal connections(i.e., increasing α_{recip}), and illustrate the effect of high α_{recip} by 2.9. Figure 2.9 **A** shows that there are more reciprocal motifs compared with the E-R model that has the same first order statistic.

Figure 2.10 demonstrates the network structure when we adding convergent motifs to SONETs while retaining the first order statistic. Figure 2.10 **A** illustrates that increasing α_{conv} leads to increasing convergent motifs, as well as in-coming hubs(i.e., nodes that

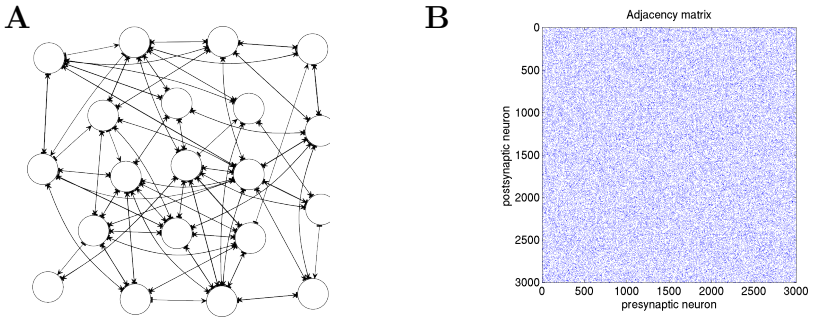


Figure 2.9: Visualization of SONETs with positive α_{recip}. (**A**) Sample SONET generated by parameters $N = 20$, $p = 0.3$, $\alpha_{\text{recip}} = 2.0$, $\alpha_{\text{conv}} = 0$, $\alpha_{\text{div}} = 0$, $\alpha_{\text{chain}} = 0$. (**B**) Sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 80$, $\alpha_{\text{conv}} = 0$, $\alpha_{\text{div}} = 0$, $\alpha_{\text{chain}} = 0$.

have many in-cominning connections). Figure 2.10 **B** and C show the positive correlation of entries in the same row. Thus large α_{conv} result in nodes that either have more in-coming connections or have very few in-coming connections. In this case, high α_{conv} will make the network more heterogeneous.

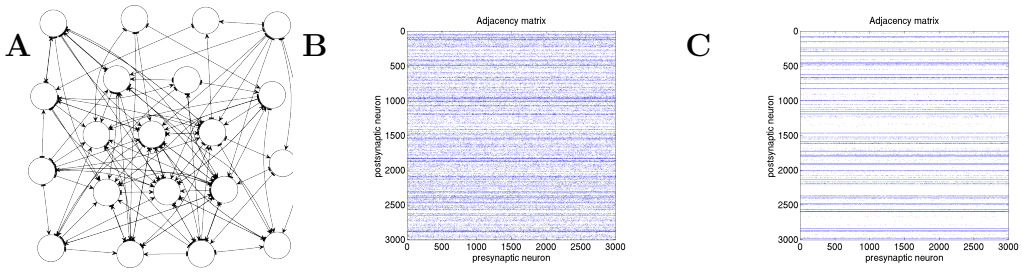


Figure 2.10: Visualization of SONETs with positive α_{conv}. (**A**) Sample SONET generated by parameters $N = 20$, $p = 0.3$, $\alpha_{\text{recip}} = 0.1$, $\alpha_{\text{conv}} = 0.5$, $\alpha_{\text{div}} = 0$, $\alpha_{\text{chain}} = 0$. (**B**) Sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 0$,$\alpha_{\text{conv}} = 10$,$\alpha_{\text{div}} = 0$,$\alpha_{\text{chain}} = 0$. (**C**) Sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 0$, $\alpha_{\text{conv}} = 60$, $\alpha_{\text{div}} = 0$, $\alpha_{\text{chain}} = 0$.

Figure 2.11 shows the sparsity pattern when we adding divergent motifs to SONETs while retaining the first order statistic. Figure 2.11 **A** illustrates that increasing α_{div} leads to increasing divergent motifs present in networks. Also, it results in some out-coming hubs(i.e., nodes that have many out-going connections). Figure 2.11 **B** and C show the positive correlation of entries in the same column. Thus large α_{div} result in nodes that

either have more out-coming connections or have very few out-going connections.

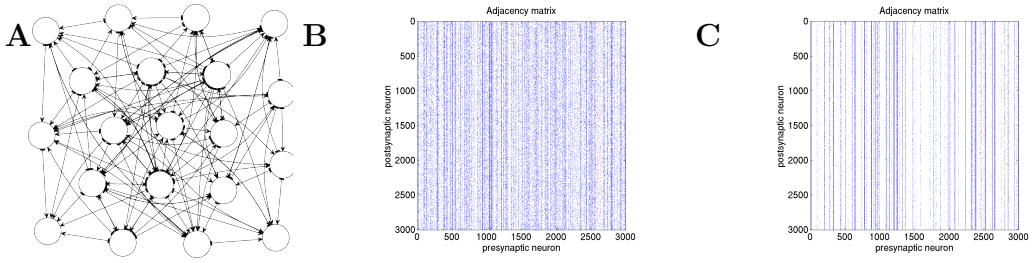


Figure 2.11: Visualization of SONETs with positive α_{div}. (**A**) Sample SONET generated by parameters $N = 20$, $p = 0.3$, $\alpha_{\text{recip}} = -0.1$, $\alpha_{\text{conv}} = 0$, $\alpha_{\text{div}} = 0.5$, $\alpha_{\text{chain}} = 0$. (**B**) sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 0$,$\alpha_{\text{conv}} = 0$,$\alpha_{\text{div}} = 10$, $\alpha_{\text{chain}} = 0$. (**C**) Sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 0$, $\alpha_{\text{conv}} = 0$, $\alpha_{\text{div}} = 60$, $\alpha_{\text{chain}} = 0$.

Figure 2.12 shows the network structure with varied α_{chain}. Figure 2.12 AC shows that when we increase α_{chain}, the in-coming hubs are more likely to be out-going hubs. BD also demonstrates the same property.

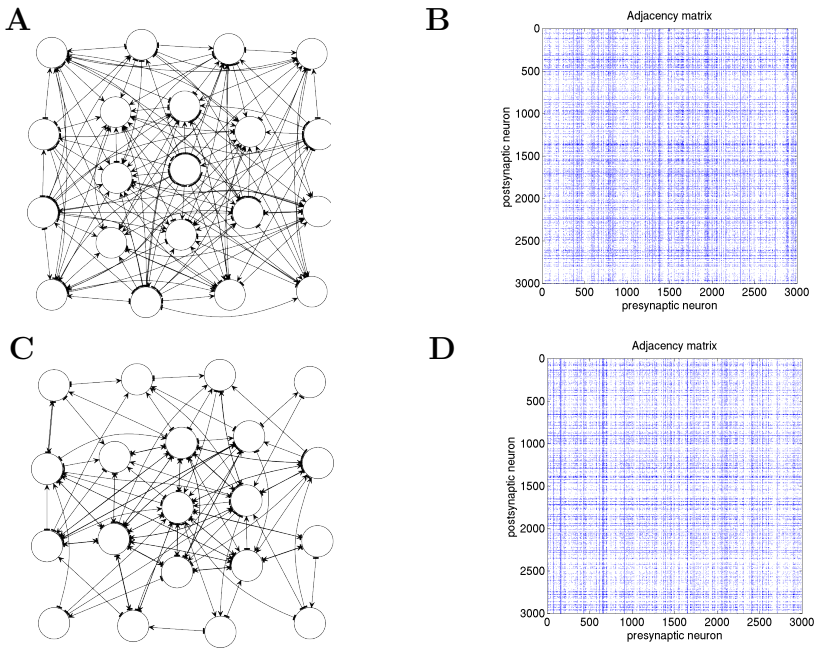


Figure 2.12: Visualization of SONETs with positive α_{conv} and α_{div}, as well as negative α_{chain}. (**A**) Sample SONET generated by parameters $N = 20$, $p = 0.3$, $\alpha_{\text{recip}} = -0.9$, $\alpha_{\text{conv}} = 0.3$, $\alpha_{\text{div}} = 0.4$, $\alpha_{\text{chain}} = -0.3$. (**B**) Sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 0$, $\alpha_{\text{conv}} = 10$, $\alpha_{\text{div}} = 10$, $\alpha_{\text{chain}} = 0$.Visualization of SONETs with positive α_{conv}. (**C**) Sample SONET generated by parameters $N = 20$, $p = 0.3$, $\alpha_{\text{recip}} = 1.0$, $\alpha_{\text{conv}} = 0.3$, $\alpha_{\text{div}} = 0.3$, $\alpha_{\text{chain}} = 0.3$. (**D**) Sparsity pattern of sample SONET generated by parameters $N = 3000$, $p = 0.01$, $\alpha_{\text{recip}} = 20$, $\alpha_{\text{conv}} = 10$, $\alpha_{\text{div}} = 10$, $\alpha_{\text{chain}} = 5$.

Chapter 3

Properties of SONETs

3.1 Degree distribution

3.1.1 Relating Second Order Connectivity Statistics to Degree Distribution

A common method to describe network structure is through the degree distribution[31, 6]. In this section, we derive the relationship of network connectivity and degree distribution.

Given a adjacency matrix W, the in-coming degree of node i is defined as the total number of edges whose in-coming node is i: $d^i_{\text{in}} = \sum_j W_{ij}$. Similarly, the out-going degree of node i is the total number of edges whose out-going node is i: $d^i_{\text{out}} = \sum_j W_{ji}$. Then we could define the degree distribution of random network. For a homogeneous random network, let p_{in} be the probability distribution of the in-coming degree of any node and p_{out} be the out-going degree. Because of homogeneity, the random variables d^i_{in} share the same probability distribution for every single node in the network. For simplicity, we define d_{in} as the random variable of the in-coming degree. Similarly, d_{out} is defined as the random variable for out-going degree. Both d_{out} and d_{in} take values on the set $\{0, 1, 2, \ldots, N-1\}$ where N is the total number of vertices in the network. Let $d = (d_{\text{in}}, d_{\text{out}})$ denote the joint degree and p be the probability distribution of d: For any given node, p_{ij} is the probability that a node has in-coming degree i and out-going degree j. Then d takes values on the grids $\{0, 1, \ldots, N-1\} \times \{0, 1, \ldots, N-1\}$. p_{in} and

p_{out} are the marginal distributions of p.

$$p_{\text{in}}(i) = \sum_{j=0}^{N} p(i,j)$$
$$p_{\text{out}}(j) = \sum_{i=0}^{N} p(i,j)$$

For mathematical convenience, we relax the assumption and allow d_{in} and d_{out} take any value in $[0, N]$. Now we have the probability density functions of incoming degree, outgoing degree, as well as joint degree, which are denoted by $\tilde{p}_{\text{in}}$, $\tilde{p}_{\text{out}}$ and $\tilde{p}$, respectively. Then we have

$$\tilde{p}_{\text{in}}(x) = \int \tilde{p}(x,y)dy \tag{3.1}$$
$$\tilde{p}_{\text{out}}(y) = \int \tilde{p}(x,y)dx \tag{3.2}$$

For a homogeneous random network with in-coming degree distribution $\tilde{p}_{\text{in}}(x)$, the expected value of in-coming degree of any node is $\int x\tilde{p}_{\text{in}}(x)dx$. For a SONET with first order statistic p and second order statistics $\alpha's$, the expected value of in-coming degree of any node is $(N-1)p$. Thus we have

$$\int x\tilde{p}_{\text{in}}(x)dx = (N-1)p \tag{3.3}$$

For any node with in-coming degree x, there are $x(x-1)/2$ convergent motifs with the in-coming node i. Thus the expected value of the number of convergent motifs with given in-coming degree distribution $\tilde{p}_{\text{in}}(x)$ is $\int \frac{x(x-1)}{2}\tilde{p}_{\text{in}}(x)dx$, and that of a SONET with given p and $\alpha's$ is $(N-1)(N-2)p^2(1+\alpha_{\text{conv}})/2$. Therefore, we obtain

$$\int \frac{x(x-1)}{2}\tilde{p}_{\text{in}}(x)dx = (N-1)(N-2)p^2(1+\alpha_{\text{conv}})/2 \tag{3.4}$$

By equation (3.4), we have

$$\alpha_{\text{conv}} = \frac{\int x(x-1)\tilde{p}_{\text{in}}(x)dx}{(N-1)(N-2)p^2} - 1 \tag{3.5}$$

Equation (3.3) demonstrates that $\int x\tilde{p}_{\text{in}}(x)dx$ is of order Np. To neglect this term in comparision to $var(d_{\text{in}})$, we need $var(d_{\text{in}}) \gg E(d_{\text{in}})$.

Thus for large N, equation (3.5) can be approximated by

$$\begin{aligned}\alpha_{\text{conv}} &\approx \frac{\int (x^2 - x)\tilde{p}_{\text{in}}(x)dx - (\int x\tilde{p}_{\text{in}}(x)dx)^2}{(N-1)p(N-1)p} \\ &= \frac{\text{var}(d_{\text{in}}) - E(d_{\text{in}})}{E(d_{\text{in}})E(d_{\text{in}})}\end{aligned} \tag{3.6}$$

Equation(3.6) can be approximated by

$$\alpha_{\text{conv}} \approx \frac{\text{var}(d_{\text{in}})}{E(d_{\text{in}})E(d_{\text{in}})} \tag{3.7}$$

as long as $\text{var}(d_{\text{in}}) \gg E(d_{\text{in}})$, i.e., d_{in} has larger variability than Poisson. Hence equation(3.7) breaks down for E-R network even for large N.

Equation (3.7) indicates that α_{conv} is a measure of spread of the in-coming degree for nodes in the network. For fixed first order statistic p of second order random networks, the larger α_{conv}, the more the in-coming degree of nodes spreads out from $(N-1)p$.

Figure (3.1) shows the histogram of the incoming degree of SONETs samples that share the same first and second order statistics except α_{conv}. In figure 3.1 **A**, in which $\alpha_{\text{conv}} = 0$, the in-coming degree of each node is larger than 235 and smaller than 363. When we increase α_{conv} while retaining the other statistics, there are an increasing number of incoming hubs and nodes with very few incoming connections. In figure 3.1 **E**, there are 2456 nodes have no incoming connections, and almost 200 nodes that almost all nodes connect to.

Similarly, for the out-going degree of second order networks, we have

$$\int y\tilde{p}_{\text{out}}(y)dx = (N-1)p \tag{3.8}$$

$$\int \frac{y(y-1)}{2}\tilde{p}_{\text{out}}(y)dy = (N-1)(N-2)p^2(1+\alpha_{\text{div}})/2 \tag{3.9}$$

For large N, we have

$$\alpha_{\text{div}} \approx \frac{\text{var}(d_{\text{out}}) - E(d_{\text{out}})}{E(d_{\text{out}})E(d_{\text{out}})} \tag{3.10}$$

Equation(3.10) can be approximated by

$$\alpha_{\text{div}} \approx \frac{\text{var}(d_{\text{out}})}{E(d_{\text{out}})E(d_{\text{out}})} \tag{3.11}$$

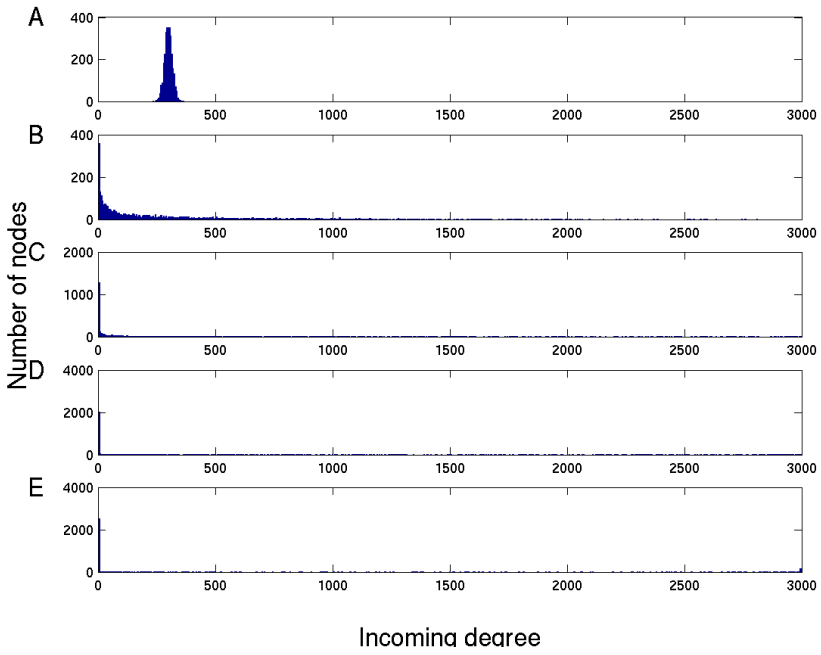


Figure 3.1: Histogram of incoming degree of SONET samples generate from statistics $N = 3000$, $p = 0.1$, $\alpha_{\text{div}} = 0$,$\alpha_{\text{chain}} = 0$ and varied α_{conv}. α_{conv} for **A**,**B**,**C**,**D**,**E** are 0,2,4,6,8, respectively. The measured convergent motif statistic $\widehat{\alpha}_{\text{conv}}$ for the five subfigures are 0.0001, 1.9749,4.1491,6.0008 and 7.9281.

if $\text{var}(d_{\text{out}}) \gg E(d_{\text{out}})$. Equation (3.11) indicates that α_{div} is the variance of the outgoing degree distribution, normalized by the mean degree squred. For fixed first order statistic p of SONET, the larger α_{div}, the more spread of outgoing degree from $(N-1)p$. Figure (3.2) shows the histogram of the outgoing degree of SONETs samples that share the same first and second order statistics except α_{div}.

Figure 3.2 **A** illustrates that the outgoing degree are less spread. **E** shows that the nodes are almost classified into two categories: outgoing hubs and nodes with very few projections.

Now let's take a look at the joint degree distribution of SONET. Assume node i has in-coming degree x and out-going degree y, xy is the sum of number of chain motifs with center at i and number of reciprocal motifs with one vertex at i. Given the probability distirubition $p(x, y)$ of random variable $(d_{\text{in}}, d_{\text{out}})$, $\iint xyp(x, y)dxdy$ is the expected sum

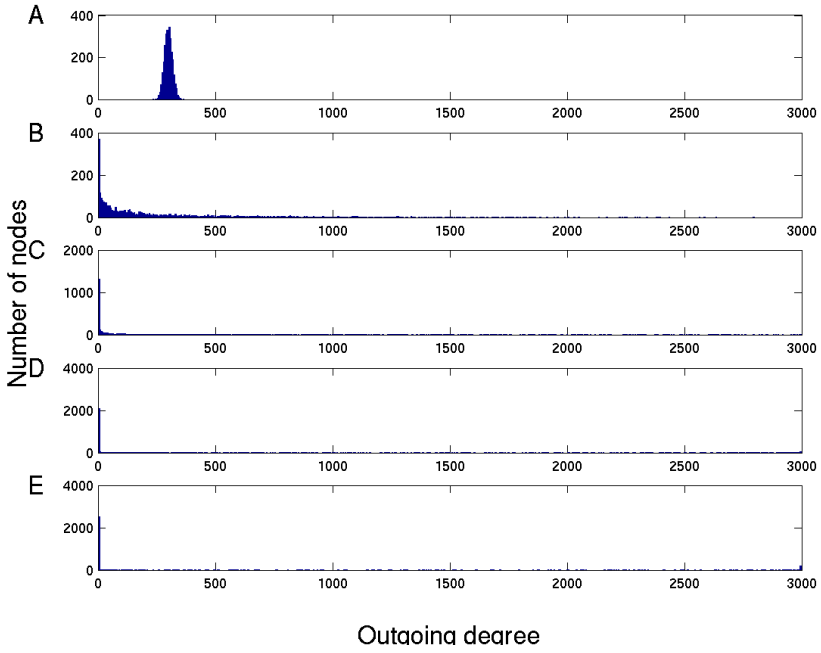


Figure 3.2: Histogram of outgoing degree of SONET samples generate from statistics $N = 3000$, $p = 0.1$, $\alpha_{\text{conv}} = 0$,$\alpha_{\text{chain}} = 0$ and varied α_{div}. α_{div} for **A**,**B**,**C**,**D**,**E** are 0,2,4,6,8, respectively. The measured divergent motif statistic $\widehat{\alpha}_{\text{conv}}$ for the five subfigures are 0.0002, 1.9312,4.0777,6.1365 and 7.5440.

of the number of chains with center at i and the number of reciprocal motifs. For a SONET of given p and $\alpha's$, we obtain

$$\iint xy\tilde{p}(x,y)dxdy = (N-1)(N-2)p^2(1+\alpha_{\text{chain}}) + (N-1)p^2\alpha_{\text{recip}} \tag{3.12}$$

Thus we have

$$\alpha_{\text{chain}} = \frac{\int\int xy\tilde{p}(x,y)dxdy - (N-1)(N-2)p^2 - 2(N-1)p^2\alpha_{\text{recip}}}{(N-1)(N-2)p^2} \tag{3.13}$$

Assume $\alpha_{\text{recip}} = O(1)$, $(N-1)p^2\alpha_{\text{recip}}$ is of order N^2p. Combining with equation(3.3)(3.8), we have

$$\alpha_{\text{chain}} \approx \frac{\text{cov}(d_{\text{in}}, d_{\text{out}})}{E(d_{\text{in}})E(d_{\text{out}})} \tag{3.14}$$

The first and second order statistics of second order networks are related with variance of the in-degree/out-degree distribution, as well as the covariance of the two distributions.

From equations (3.7), (3.11) and (3.14), we have the inequality for second order networks:

$$\alpha_{\text{chain}}^2 \leq \alpha_{\text{conv}}\alpha_{\text{div}} \tag{3.15}$$

as long as $\text{var}(d_{\text{in}}) \gg E(d_{\text{in}})$ and $\text{var}(d_{\text{out}}) \gg E(d_{\text{out}})$. Equations(3.14) illustrates that α_{chain} is the correlation of the in-coming degree and out-going degree. $\alpha_{\text{chain}} = 0$ leads to independent incoming degree and outgoing degree. When α_{chain} is positive, the in-coming degree and out-going degree are positively correlated. In this case, incoming hubs are more likely to be outgoing degree, and nodes with few incoming connections are more likely to project to few nodes. Negative α_{chain} results in negative correlations of the two marginal degree distribution, which means that the incoming hubs are less likely to be outgoing hubs. The results are be illustrated by figure 3.3.

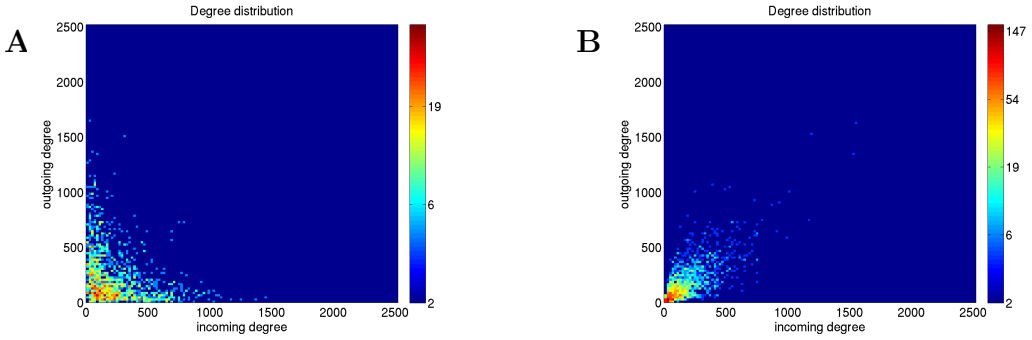


Figure 3.3: Effect of α_{chain} on the incoming degree and outgoing degree distribution of single node. Figure 3.3 **A** shows the joint degree histogram of SONET sample of 3000 nodes with measured statistics $\hat{p} = 0.102$, $\widehat{\alpha}_{\text{recip}} = 1.0333$, $\widehat{\alpha}_{\text{conv}} = 1.0964$, $\widehat{\alpha}_{\text{div}} = 1.0582$, $\alpha_{\text{chain}} = -0.2059$. **B** shows the joint degree distribution of SONET sample of 3000 nodes with measured statistics $\hat{p} = 0.1000$, $\widehat{\alpha}_{\text{recip}} = 1.1117$, $\widehat{\alpha}_{\text{conv}} = 1.1098$, $\widehat{\alpha}_{\text{div}} = 1.1493$, $\widehat{\alpha}_{\text{chain}} = 0.8769$.

3.1.2 In-coming and out-going degree distribution

In this section, we derive an analytic expression for the incoming degree distribution of SONETs. The outgoing degree distribution is analogous.

For node i, the $N-1$ random variables $V_i = (W_{i1}, \ldots, W_{i,i-1}, W_{i,i+1}, \ldots, W_{iN})$ represent the connections onto i. The V_i are sampled from a correlated Bernoulli distribution. They could be obtained by dichotomizing a multivariate normal distribution $(\tilde{W}_{i1}, \ldots, \tilde{W}_{i,i-1}, \tilde{W}_{i,i+1}, \ldots, \tilde{W}_{i,N})$ to 0 and 1.

Let $A = (A_1, A_2, \ldots, A_N)$ be a vector of Bernoulli random variables such that they have the same mean and pairwise covariance. The random vector can be obtained by dichotomizing a multivariate normal distribution X with mean γ and covariance matrix Σ where $\Sigma_{ii} = 1$ and $\Sigma_{ij} = \rho$ for $i \neq j$. We set A_i to be 1 if X_i is nonnegative and set it to be 0 otherwise. X can be written as

$$\begin{aligned} X_i &= \gamma + \sqrt{\rho}U + \sqrt{1-\rho}V_i \\ U &\sim \mathcal{N}(0,1) \\ V_i &\sim \mathcal{N}(0,1) \end{aligned} \tag{3.16}$$

To make the dichotomized Gaussian match the statistics of correlated bernoulli with mean p and correlation α, we should solve equation(2.45), which is equivalent to solving the system of equations(3.17).

$$\begin{aligned} &\int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\gamma)^2\right) dx = p \\ &\int_0^\infty \int_0^\infty \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}(x_1-\gamma, x_2-\gamma) \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x2 \end{pmatrix}\right) dx_1 dx_2 = p^2(1+\alpha) \end{aligned} \tag{3.17}$$

Let $K = \sum_{i=1}^N A_i$. For fixed integer $0 \leq k \leq N$, there are $\binom{N}{k}$ subsets of $(A_1, \ldots, A_N)$ that have exactly k positive values. Thus

$$\begin{aligned} \Pr(K=k) &= \binom{N}{k} \Pr(A_1 = \cdots = A_k = 1, A_{k+1} = \cdots = A_N = 0) \\ &= \binom{N}{k} \Pr(X_1 \geq 0, \ldots, X_i \geq 0, X_{i+1} \leq 0, \ldots, X_N \leq 0) \end{aligned} \tag{3.18}$$

Let $R = K/N$ and $p(r)$ denote that probability distribution of R. Let $k = Nr$, we have

$$\begin{aligned}
p(r) &= \Pr(R = r) \\
&= \Pr(K = Nr) \\
&= \binom{N}{Nr} \Pr(A_1 = 1, \ldots, A_{Nr} = 1, A_{Nr+1} = 0, \ldots, A_N = 0) \\
&= \binom{N}{Nr} \int_{-\infty}^{\infty} \phi(s) (\Pr(A_1 = 1|s))^{Nr} (\Pr(A_1 = 0|s))^{N(1-r)} ds \\
&= \binom{N}{Nr} \int_{-\infty}^{\infty} \phi(s) (\Pr(X_1 \geq 0|s))^{Nr} (\Pr(X_1 < 0|s))^{N(1-r)} ds \\
&= \binom{N}{Nr} \int_{-\infty}^{\infty} \phi(s) f(s)^{Nr} (1 - f(s))^{N(1-r)} ds
\end{aligned} \tag{3.19}$$

where $\phi(s)$ is the probability density function of standard normal distribution and $f(s)$ is the probability distribution function of $A_1 = 1$ conditioned on s. By equation(3.16), we have

$$\begin{aligned}
f(s) &= \Pr(A_1 = 1|s) = \Pr(X_1 \geq 0|s) \\
&= \Pr(\gamma + \sqrt{\rho} s + \sqrt{1-\rho} v > 0|s) \\
&= \Pr\left(v > -\frac{\gamma + \sqrt{\rho} s}{\sqrt{1-\rho}}\right) \\
&= \Phi\left(\frac{\gamma + \sqrt{\rho} s}{\sqrt{1-\rho}}\right)
\end{aligned} \tag{3.20}$$

where Φ is the cummulative distribution function of the standard normal distribution.

Then we have

$$p(r) = \binom{N}{Nr} \int_{-\infty}^{\infty} \phi(s) \exp Nh(s) ds \tag{3.21}$$

where

$$h(s) = r \log f + (1 - r) \log(1 - f) \tag{3.22}$$

By the saddle-point approximation,

$$p(r) \approx \binom{N}{Nr} \phi(s_0) e^{Nh(s_0)} \sqrt{\frac{2\pi}{-Nh''(s_0)}} \tag{3.23}$$

where h has a unique maximum at s_0.

$$h'(s) = \left(\frac{r}{f(s)} - \frac{1-r}{1-f(s)}\right)\phi\left(\frac{\gamma+\sqrt{\rho}s}{\sqrt{1-\rho}}\right)\sqrt{\frac{\rho}{1-\rho}} \tag{3.24}$$

The global maxima $s_0 = f^{-1}(r)$ is obtained. Let $\eta = \frac{\gamma+\sqrt{\rho}s_0}{\sqrt{1-\rho}}$. Then $\eta = \Phi^{-1}(r)$. Also, we have

$$\begin{aligned} &h(s_0) = 0 \\ &h''(s_0) = -\frac{1}{r(1-r)} \cdot \frac{\rho}{1-\rho} \cdot \phi^2(\eta) \end{aligned} \tag{3.25}$$

By equations (3.23), (3.25), we obtain

$$\begin{aligned} p(r) &= \binom{N}{Nr}\sqrt{\frac{2\pi r(1-r)(1-\rho)}{N\rho}}\frac{\phi(s_0)}{\phi(\eta)} \\ &= \binom{N}{Nr}\sqrt{\frac{2\pi r(1-r)(1-\rho)}{N\rho}}\exp\left(-\frac{1}{2}\left(\frac{\sqrt{1-\rho}\eta-\gamma}{\sqrt{\rho}}\right)^2+\frac{\eta^2}{2}\right) \\ &= \binom{N}{Nr}\sqrt{\frac{2\pi r(1-r)(1-\rho)}{N\rho}}\exp\left(\frac{\gamma^2}{2-4\rho}\right)\exp\left(-\frac{1}{2}\frac{\left(\eta-\frac{\sqrt{1-\rho}\gamma}{1-2\rho}\right)^2}{\frac{\rho}{1-2\rho}}\right) \end{aligned} \tag{3.26}$$

For large N, $p(r)$ could be approximated by

$$p(r) \approx \sqrt{\frac{1-\rho}{\rho}}\exp\left(\frac{\gamma^2}{2-4\rho}\right)\exp\left(-\frac{1}{2}\frac{\left(\eta-\frac{\sqrt{1-\rho}\gamma}{1-2\rho}\right)^2}{\frac{\rho}{1-2\rho}}\right) \tag{3.27}$$

where $\eta = \Phi^{-1}(r)$. When $\rho = \frac{1}{2}$, we have

$$p(r) = \sqrt{\frac{1-\rho}{\rho}}\exp(\sqrt{2}\eta r - r^2)$$

Let's take a look at the example where the correlation of the pairwise bernoulli random variables A_i is 0, $\rho = 0$.

$$p(r) = Pr\left(\sum_{i=1}^{N} A_i = Nr\right) = \binom{N}{Nr}p^{Nr}(1-p)^{N(1-r)} \tag{3.28}$$

When the correlation of the pairwise Bernoulli random variables is 1, $\sum_{i=1}^{N} V_i$ is either 0 or N. The probability density function is

$$p(r) = p\delta(r-1) + (1-p)\delta(r) \tag{3.29}$$

When ρ is close to 0, $p(r)$ is more like a binomial distribution. When ρ approaches 1, it goes to a double δ function.

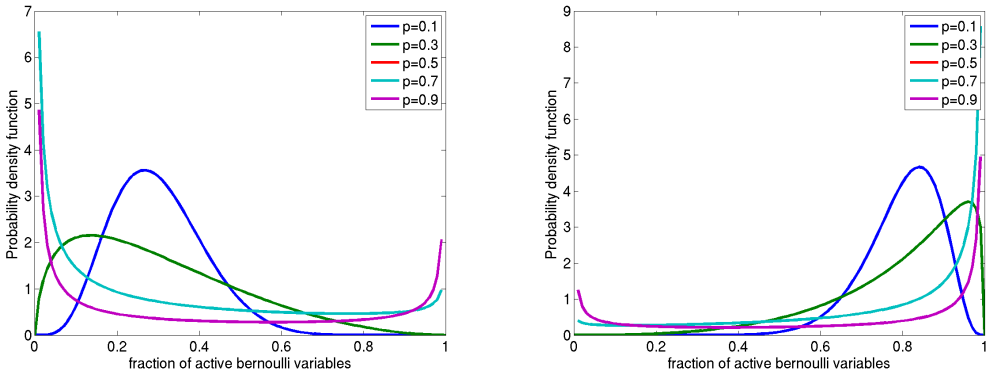


Figure 3.4: Fraction of active Bernoulli random variable of the dichotomized Gaussian for the mean fraction 0.3(left figure) and 0.8(right figure)

When N is large, we can derive the incoming degree distribution of SONET by the dichotomized Gaussian method. For any node i, $d_{\text{in}}^{i} = \sum_{i \neq j} W_{ij}$. By equation(2.15), W_{ij} are correlated Bernoulli random variables with correlation α_{conv}. By solving system (3.17), we obtain the mean and correlation of the underlying multivariate normal distribution. The mean of the dichotomized Gaussian depends only on the statistic p, thus we denote it by γ_p. The pairwise correlation among the Gaussian is a function in terms of p and α_{conv}, and is denoted by ρ_{conv}. By equation(3.27), we have the analytic expression for probability distribution of the incoming degree:

$$p_{\text{in}}(x) \approx \sqrt{\frac{1-\rho_{\text{conv}}}{\rho_{\text{conv}}}} \exp\left(\frac{\gamma_p^2}{2-4\rho_{\text{conv}}}\right) \exp\left(-\frac{1}{2}\frac{\left(\Phi^{-1}(x/N) - \frac{\sqrt{1-\rho_{\text{conv}}}\gamma_p}{1-2\rho_{\text{conv}}}\right)^2}{\frac{\rho_{\text{conv}}}{1-2\rho_{\text{conv}}}}\right) \tag{3.30}$$

for $x \in [0, N]$, where Φ is the cummulative distribution function of the standard normal distribution.

We obtain the analytic expression for the outgoing degree distribution of the SONET:

$$p_{\text{out}}(y) \approx \sqrt{\frac{1-\rho_{\text{div}}}{\rho_{\text{div}}}} \exp\left(\frac{\gamma_p^2}{2-4\rho_{\text{div}}}\right) \exp\left(-\frac{1}{2}\frac{\left(\Phi^{-1}(y/N) - \frac{\sqrt{1-\rho_{\text{div}}}\gamma_p}{1-2\rho_{\text{div}}}\right)^2}{\frac{\rho_{\text{div}}}{1-2\rho_{\text{div}}}}\right) \tag{3.31}$$

for $y \in [0, N]$, where ρ_{div} is the correlation of the underlying multivariate normal distribution obtained by equation(3.17).

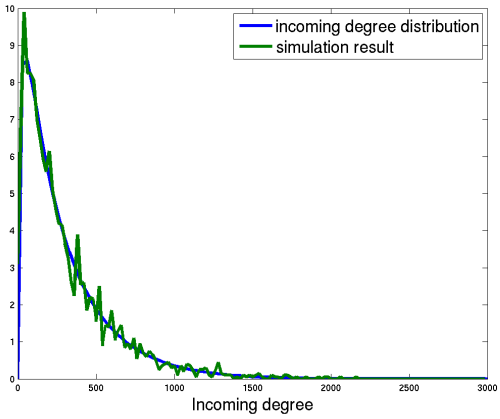


Figure 3.5: Incoming degree distribution of SONET. The blue line is the plot of incoming degree distribution (3.30) with parameters $N = 3000$, $p = 0.1$, $\alpha_{\text{conv}} = 0.9$. The green line is the incoming degree fraction obtained from a SONET sample with measured statistic: $N = 3000$, $\hat{p} = 0.1008$, $\widehat{\alpha}_{\text{recip}} = 2.0129$, $\widehat{\alpha}_{\text{conv}} = 0.9517$, $\widehat{\alpha}_{\text{div}} = 0.3025$,$\widehat{\alpha}_{\text{chain}} = 0.4169$.

3.1.3 Asymptotic joint degree distribution

In Appendix (A.2), we derive the analytic expression of the joint degree distribution $p(d_{\text{in}}, d_{\text{out}})$ for SONET with given statistics. By matching the statistics of the underlying multivariate normal distribution to the corresponding statistics of SONET, we obtain γ_p, ρ_{conv}, ρ_{div} and ρ_{chain}. The asymptotic joint degree distribution of the SONET of N nodes can be described by (3.32).

$$\Pr(x, y) = \sqrt{\frac{(1-\rho_{\text{conv}})(1-\rho_{\text{div}})}{\rho_{\text{conv}}\rho_{\text{div}} - \rho_{\text{chain}}^2}} \exp\left(-\frac{1}{2}((D\eta-\gamma)^T(CC^T)^{-1}(D\eta-\gamma) - \eta^T\eta)\right) \tag{3.32}$$

where $CC^T = \begin{pmatrix} \rho_{\text{conv}} & \rho_{\text{chain}} \\ \rho_{\text{chain}} & \rho_{\text{div}} \end{pmatrix}$, $D = \begin{pmatrix} \sqrt{1-\rho_{\text{conv}}} & 0 \\ 0 & \sqrt{1-\rho_{\text{div}}} \end{pmatrix}$, and $\eta = \begin{pmatrix} \Phi^{-1}(x/N) \\ \Phi^{-1}(y/N) \end{pmatrix}$.

3.2 Topological Properties

3.2.1 Average length path

The average path length describes the efficiency of a network. It is the average number of steps along the shortest path between all possible pairs of nodes of in the network. For a network of N nodes, the average path length is defined as

$$l = \frac{1}{N(N-1)} \sum_{i,j} d^{ij} \tag{3.33}$$

where d^{ij} is the smallest number of edges from node j to i.

However, the geodesic distance d^{ij} is infinity if there is no path from j to i. We just consider the largest component of each network so that there exist paths between any pair of nodes in the component. If two nodes belongs to different components, their geodesic distance is ignored. Another way to avoid the problem is to use the "harmonic distance" among any pair of nodes, which is just the reciprocal of the average of geodesic distance reciprocals:

$$l^{-1} = \frac{1}{N(N-1)} \sum_{i,j} \frac{1}{d^{ij}} \tag{3.34}$$

To investigate how the second order statistics of SONETs affect their average shortest path, we generate 186 SONETs with $N = 3000$ and $p = 0.1$. We sample the second order statistics in the range $\alpha_{\text{recip}} \in [-1, 4]$, $\alpha_{\text{conv}} \in [0, 1]$, $\alpha_{\text{div}} \in [0, 1]$, and $\alpha_{\text{chain}} \in [-1, 1]$.

In fig(3.6), the average shortest path is plotted against the measured second order statistics $\hat{\alpha}$. Fig(3.6)**A** and **B** shows that there is no simple relation between the average shortest path and α_{conv} or α_{div}. Fig(3.6) **C** illustrates that the average shortest path decreases as α_{chain} increases. Intuitively this is reasonable. When α_{chain} is high, an incoming hub is more likely to be an outgoing hub. The path between two nodes becomes shorter by traversing via the hubs.

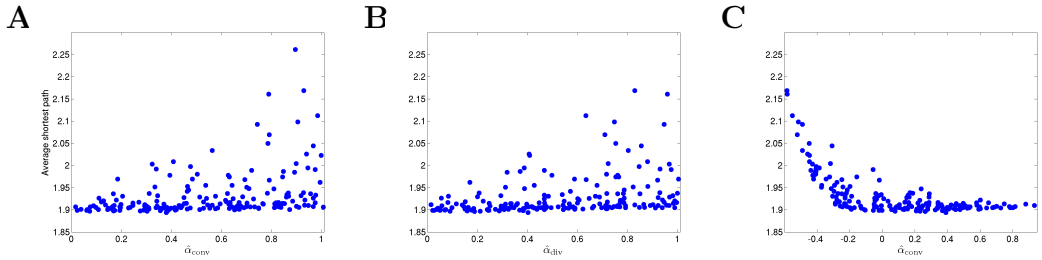


Figure 3.6: Average shortest path as a function of $\widehat{\alpha}_{\text{conv}}$, $\widehat{\alpha}_{\text{div}}$ and $\widehat{\alpha}_{\text{chain}}$ for the 186 sampled SONETs. For each panel, each dot corresponds to one network. Estimates of second order connectivity statistics $\hat{\alpha}$ are determined from each network connectivity matrix by equations (2.21)(2.22)(2.23).

3.2.2 Clustering Coefficients and Transitivity

The clustering coefficient is one of the most important measures in networks. It is a measure of how likely the neighbors of a vertex to form triangles. In many networks, if the connectivity between i, j and j, k exist, the probability that i and k are connected is higher than that of the independent random network.

Figure(3.7) shows the two types of triangles in directed graph. The Type I triangle consists three chain motifs, and the Type II triangle is composed of one convergent, one divergent and one chain motif.

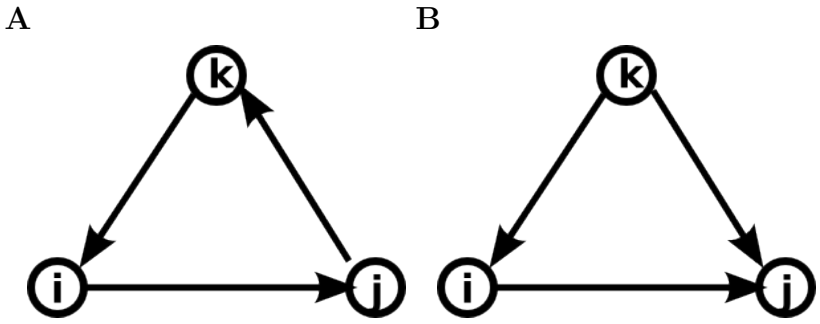


Figure 3.7: **A**: Type I triangle. **B**: Type II triangle

For directed networks, there are four types of clustering coefficients $C_l, l = 1 \ldots 4$. C_1 is the fraction of presence of Type I Triangle given the presence of a chain motif. C_2, C_3, C_4 are the fraction of presence of Type II Triangle given the presence of

a convergent motif, divergent motif and chain motif, respectively. The definition are approximation are in Appendix (A.3).

To investigate the effect of second order connectivity statistics on the clustering coefficients, we calculate the coefficients of 186 sampled SONETs. In figure(3.8), $\hat{C}_1$ is plotted against the measured second order statistics α calculated by equation(2.21)(2.22)(2.23). Figure(3.8)**C** shows that C_1 is strongly related to α_{chain}. C_1 increases as we increase α_{chain}.

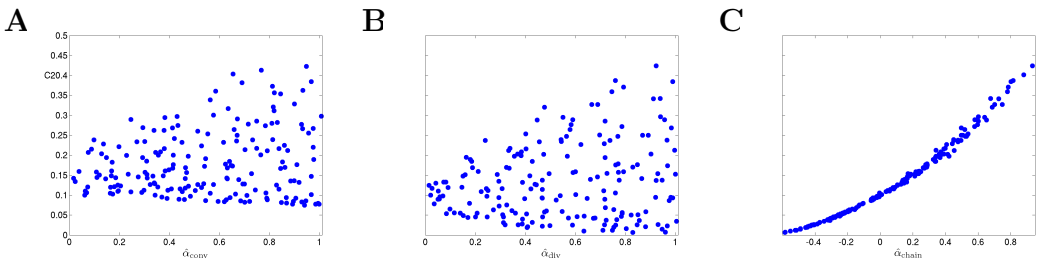


Figure 3.8: Scatter plots of $\hat{C}_1$ measured by equation(A.47)(A.51) as a function of individual connectivity statistics for the 186 sampled SONETs. Figure **A**,**B**,**C** denote $\hat{C}_1$ in terms of $\widehat{\alpha}_{\text{conv}}$, $\widehat{\alpha}_{\text{div}}$ and $\widehat{\alpha}_{\text{chain}}$, respectively. For each panel, each dot corresponds to one network. Estimates of second order connectivity statistics $\hat{\alpha}$ are determined from each network connectivity matrix. Figure(3.8)**C** shows that C_1 is closely related with $\widehat{\alpha}_{\text{chain}}$. C_1 increases as α_{chain} increases.

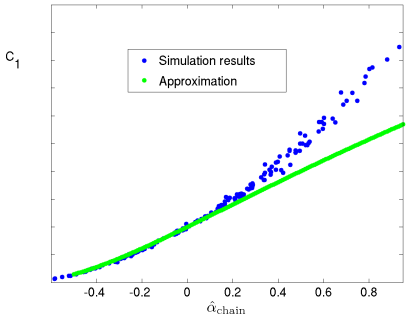


Figure 3.9: Scatter plots of $\hat{C}_1$ with respect to α_{chain}. The blue dots represent $\hat{C}_1$ obtained from 186 sampled SONETs, and the green dots are approximation of C_1 with respect to α_{chain} obtained from equation(A.43)

.

Figure(3.10)A,B,C are scatter plots of C_2 against single variable $\widehat{\alpha}_{\text{conv}}$, $\widehat{\alpha}_{\text{div}}$ and

$\widehat{\alpha}_{\text{chain}}$. No simple relationship between C_2 and any single connectivity statistics is obvious. Figure(3.10)D,E,F shows the plot of C_2 against pairs of connectivity statistics. C_2 seems to be a function of $(\alpha_{\text{div}}, \alpha_{\text{chain}})$. C_2 varies smoothly in the $(\alpha_{\text{div}}, \alpha_{\text{chain}})$ projection although α_{recip} and α_{conv} varies widely across the whole plot. C_2 increases as α_{div} or α_{chain} increases.

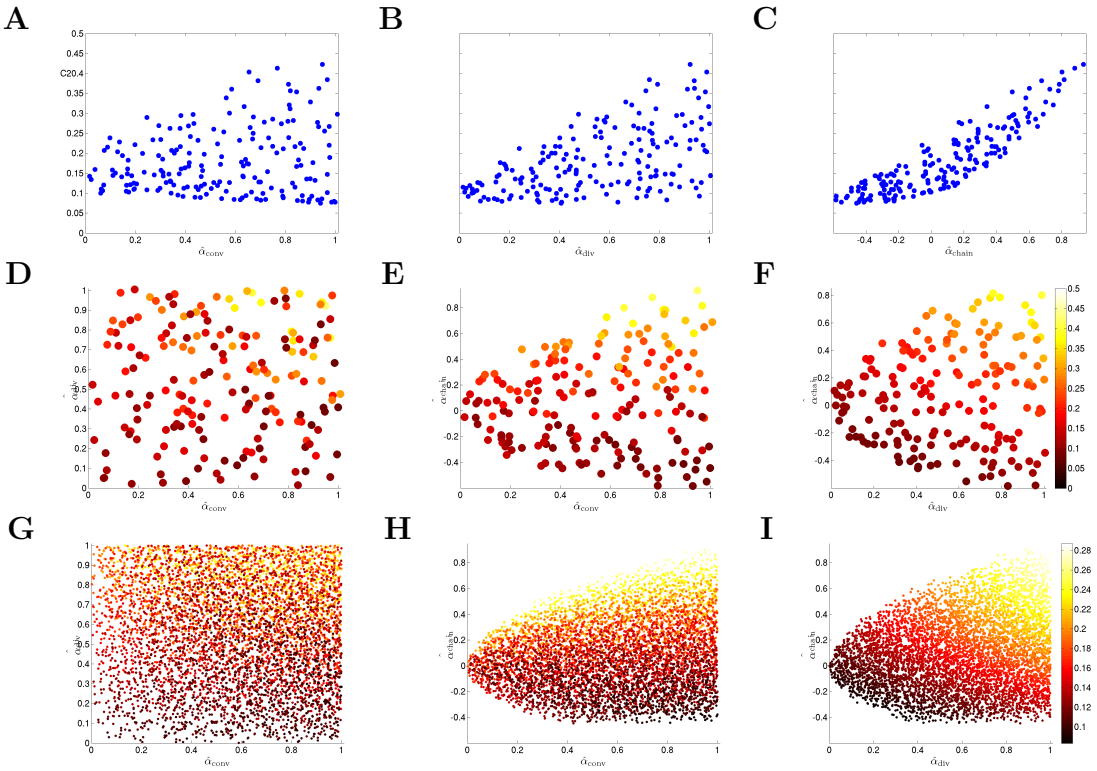


Figure 3.10: Scatter plots of the second type of clustering coefficients. Figure(3.10)**ABC** are $\hat{C}_2$ measured by equation(A.48)(A.51) as a function of individual connectivity statistics for the 186 sampled SONETs. Figure(3.10)**D,E,F** are scatter plots of $\hat{C}_2$ as a function of pairs of connectivity statistics. Figure(3.10)D are replotted as a function of connectivity statistic pair $(\alpha_{\text{conv}}, \alpha_{\text{div}})$. E,F are replotted as function of $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ and $(\alpha_{\text{div}}, \alpha_{\text{chain}})$ respectively. Each dot represent a network, and color indicates $\hat{C}_2$ measured from the adjacency matrix. Figure(3.10)**G,H,I** are scatter plots of the approximation of C_2 with respect to connectivity statistics pair.

Figure(3.11)A,B,C are scatter plots of C_3 against single variable $\widehat{\alpha}_{\text{conv}}$, $\widehat{\alpha}_{\text{div}}$ and $\widehat{\alpha}_{\text{chain}}$. Similarly, there is no simple relationship between it and any single connectivity

statistics. Figure(3.11)D,E,F shows the plot of C_3 against pairs of connectivity statistics. C_3 seems to be a function of $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ and is not dependent on α_{div} or α_{recip}.

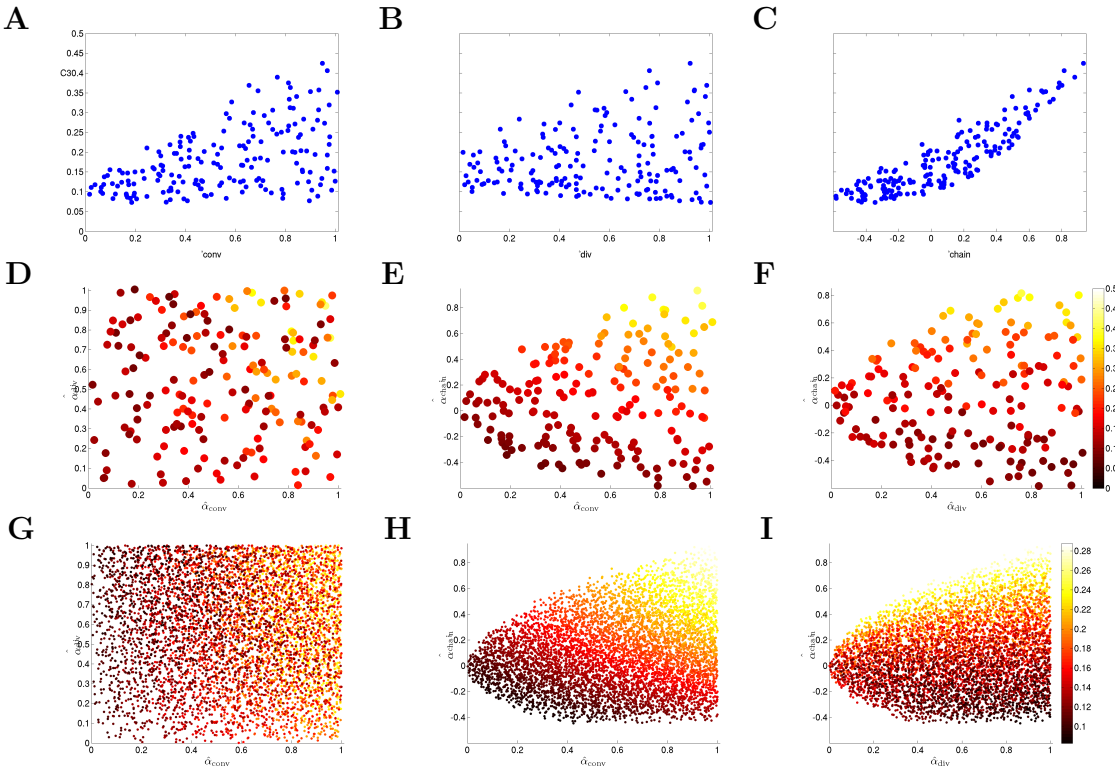


Figure 3.11: Scatter plots of the third type of clustering coefficients. Figure(3.11)**A**,**B**,**C** are $\hat{C}_3$ measured by equation(A.49)(A.51) as a function of individual connectivity statistics for the 186 sampled SONETs. Figure(3.11)**D,E,F** are scatter plots of $\hat{C}_3$ as a function of pairs of connectivity statistics. $\hat{C}_3$ are measured over the 186 SONETs. Figure(3.11)D are replotted as a function of connectivity statistic pair $(\alpha_{\text{conv}}, \alpha_{\text{div}})$. E,F are replotted as function of $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ and $(\alpha_{\text{div}}, \alpha_{\text{chain}})$ respectively. Each dot represents a network, and color indicates $\hat{C}_2$ measured from the adjacency matrix. Figure(3.11)**G,H,I** are scatter plots of the approximation of C_3. with respect to connectivity statistics pair

3.2.3 Assortativity

Assortativity is a measure of network nodes' preference of attaching to others that are similar. The similarity is usually examed in terms of node degree. In this section, we

exam the incoming degree assortativity and outgoing degree assoratativity, and investigate how the second order statistics influence the assortativity of directed network.

The assortativity coefficient is defined as the pearson correlation coefficient of degree between pairs of connected nodes. To test the effect of second order connectivity statistics on the assortativity coefficient, we compute the incoming and outgoing degree assortativity coefficient of the 186 sampled SONETs, and plot the coefficient in terms of the statistics. Figure(3.13) C shows that the incoming degree assortativity coefficient is decreasing as α_{chain} increases. Figure(3.14)C demonstrates the similar property.

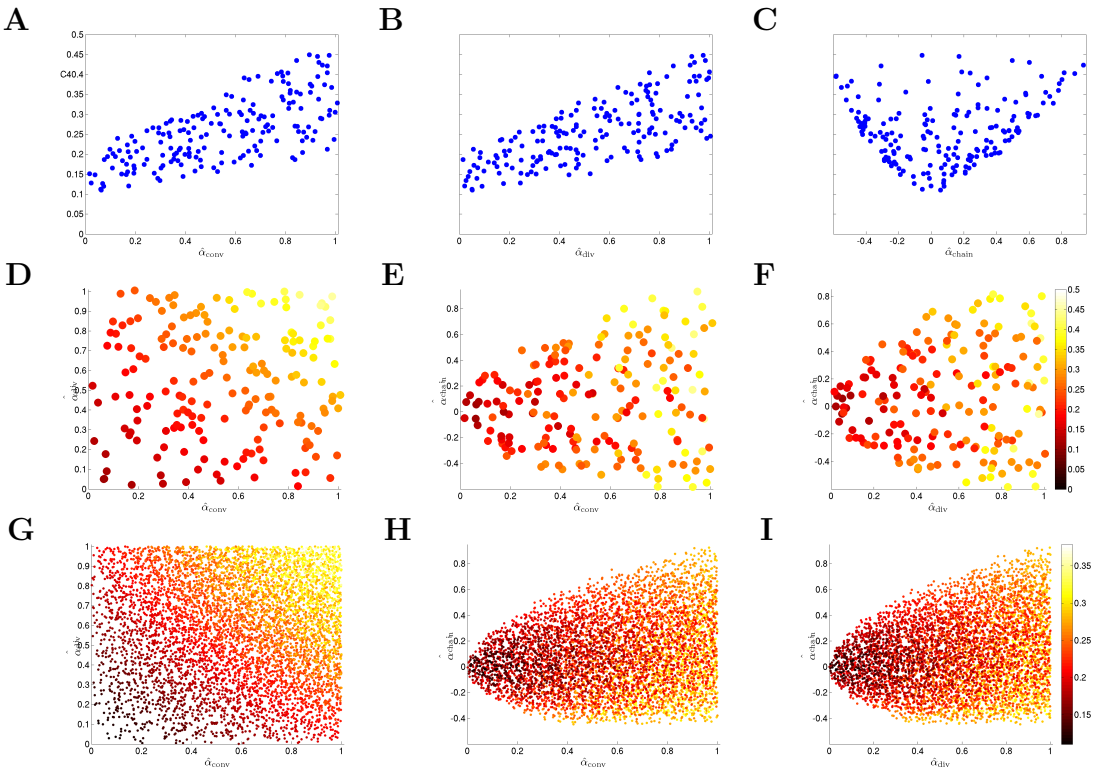


Figure 3.12: Scatter plots of the fourth type of clustering coefficients. Figure(3.12)**A**,**B**,**C** are $\hat{C}_4$ as a function of individual connectivity statistics for the 186 sampled SONETs. Figure(3.12)**D,E,F** are scatter plots of $\hat{C}_3$ as a function of pairs of connectivity statistics. Figure(3.12)D are replotted as a function of connectivity statistic pair $(\alpha_{\text{conv}}, \alpha_{\text{div}})$. E,F are replotted as function of $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ and $(\alpha_{\text{div}}, \alpha_{\text{chain}})$ respectively. Each dot represent a network, and color indicates $\hat{C}_2$ measured from the adjacency matrix. Figure(3.12)**G,H,I** are scatter plots of the approximation of C_3. with respect to connectivity statistics pair

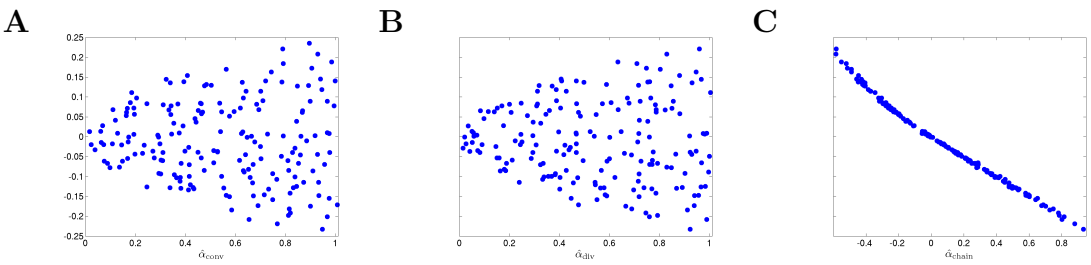


Figure 3.13: Scatter plot of incoming degree assortativity coefficient against network single connectivity statistics. Figure(3.13)**C** demonstrates that the incoming degree assortativity coefficient is a function of α_{chain}.

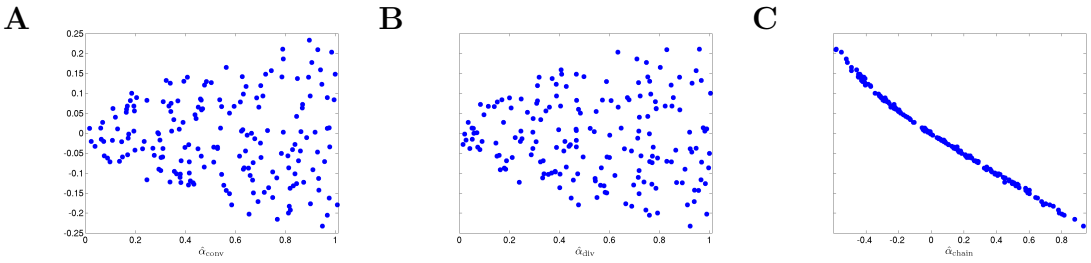


Figure 3.14: Scatter plot of outgoing degree assortativity coefficient against network single connectivity statistics. Figure(3.14)**C** demonstrates that the incoming assortativity is a function of α_{chain}.

Chapter 4

Synchrony

The neuron system is highly organized to allow information to flow efficiently [49][46], but is robust to pathological behaviors such as seizures. Neuronal synchrony is thought to play an important role in memory formation [3] while pathological amounts of synchrony are thought to be indicative of schizophrenia[51], Parkinson's disease [4], Alzheimer's disease[20], and epilepsy [30][52]. In many diseases, strong evidence suggests reorganization of the neuronal connections causing or caused by the disease[8][35] [54] may play a role in generating these pathological behaviors.

The synchronization of spiking activity neuronal networks is due to a complex interplay among many factors; individual neuron dynamics, the types of synaptic response, external inputs to the network, as well as the network topology all play a role in determining the level of synchronization. The analysis of small networks of neurons has provided significant insight into the ways in which neuron dynamics and synaptic response can influence the tendency of the network to synchronize [14][47]. The focus of the present section is to examine what types of network structures can decrease or increase the likelihood of a network to synchronize.

The study of the synchronization properties of networks has lead to notions of *synchronizability*[2], essentially a ranking of networks based on the range of neuron and synapse models that would synchronize when interconnected via each network. Clearly without specifying a neuron and synapse model, one cannot determine if a network will synchronize. Nonetheless, a network with high synchronizability would be expected to

synchronize for a large range of model choices and parameters than would a network with low synchronizability.

The analysis underlying the notion of synchronizability is based on the master stability function of approach of Pecora and Carroll [36]. The result is that the synchronizability of a given network is determined by the eigenvalues of a matrix closely related to the network adjacency matrix, the matrix specifying the connectivity graph of the network.

To connect the synchronizability analysis more directly with network structure, we employ the framework of *second order networks*(2). These second order networks, when combined with the synchronizability analysis, will give insight into the types of local network structures that play a large role in determining the influence of the network topology on synchrony.

4.1 Analysis of synchrony

To examine the influence of the second order connectivity statistics on network synchrony, we will analyze the stability of the two extreme cases of synchrony, the perfectly synchronous state and the completely asynchronous state. Through linearizations around these states, we derive conditions between the connectivity statistics and synchrony that are valid near these states. We begin by analyzing complete synchrony.

4.1.1 Stability of synchrony determined using master stability function.

To determine the stability of the perfectly synchronous case, we linearize the equations for network dynamics around synchrony using the master stability function approach by Pecora and Carrol [36]. This approach is based on determining synchrony in a system of identical, noise-free, coupled oscillators:

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{F}(\mathbf{x}_i) + S\sum_j W_{ij}(\mathbf{H}(\mathbf{x}_j) - \mathbf{H}(\mathbf{x}_i)), \tag{4.1}$$

where the state of oscillator i at time t is described by the vector $\mathbf{x}_i(t)$. The dynamics of this system are determined by the scalar coupling strength S and two vector-valued

functions of the same dimension as the interval variable vector $\mathbf{x}_i$: the internal dynamics function $\mathbf{F}$ and the coupling function $\mathbf{H}$. We rewrite equation(4.1) in terms of the Laplacian matrix $L = D - W$, where D is the diagonal of row sums of the connectivity matrix W,

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \mathbf{F}(\mathbf{x}_i) - S\sum_j L_{ij}\mathbf{H}(\mathbf{x}_j). \tag{4.2}$$

Using the master stability function approach, one can determine a critical stability region in the complex plane that is based solely on the neuron dynamics (given by $\mathbf{F}$ and $\mathbf{H}$) and is independent of network connectivity. The influence of the network on the stability enters through the eigenvalues μ_i of L, as the eigenvalues scaled by the coupling strength $S\mu_i$ must lie in the critical region for the synchronous state to be stable [2].

If the neuron model is not known, it is not possible to know the critical region in the complex plane, reflecting the fact that the topology alone cannot determine whether a network will synchronize. However, a tighter spread of the eigenvalues will make it easier to scale them by S to fit within the critical region. Therefore, changes in network structure that decrease the spread of the eigenvalues will favor *synchronizability*. One measure of this eigenvalue spread for asymmetric matrices, as suggested by Nishikawa and Motter [32], is the variance σ_μ^2 of the eigenvalues normalized by the mean degree squared,

$$\sigma_\mu^2 = \frac{1}{d^2(N-1)}\sum_{i=2}^{N}|\mu_i - \bar{\mu}|^2, \tag{4.3}$$

where $\bar{\mu} = \frac{1}{N-1}\sum_{i=2}^{N}\mu_i$ is the mean of the eigenvalues and d is the mean degree. This variance ignores the zero eigenvalue $\mu_1 = 0$ present in every Laplacian matrix.

Modifications for chemical synapses. The master stability function requires the subtraction of $\mathbf{H}(\mathbf{x}_i)$ in (4.1) to ensure that the synchronous state $\mathbf{x}_i(t) = \hat{\mathbf{x}}(t)$ exists and satisfies the equation $\frac{\mathrm{d}\hat{\mathbf{x}}}{\mathrm{d}t} = \mathbf{F}(\hat{\mathbf{x}})$. The $\mathbf{H}(\mathbf{x}_i)$ term is suitable for neuron models with electrical synapses but is not an appropriate model for chemical synapses because it implies that a connection from neuron j onto neuron i leads to an equivalent negative

effect from neuron i onto itself. To model chemical synapses, we must remove the offending $\mathbf{H}(\mathbf{x}_i)$ term, yielding:

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \mathbf{F}(\mathbf{x}_i) + S\sum_j W_{ij}\mathbf{H}(\mathbf{x}_j). \tag{4.4}$$

The rows of the connectivity matrix W may not sum to the same value, in contrast to the Laplacian matrix L of (4.2) which must have rows that sum to zero. Hence, if we substitute the synchronous state $\mathbf{x}_i = \hat{\mathbf{x}}$, we would get different, inconsistent equations that $\hat{\mathbf{x}}$ must satisfy:

$$\frac{\mathrm{d}\hat{\mathbf{x}}}{\mathrm{d}t} = \mathbf{F}(\hat{\mathbf{x}}) - d^i_{\mathrm{in}} S\mathbf{H}(\hat{\mathbf{x}}), \tag{4.5}$$

where the ith row sum of W is d^i_{in}, the incoming degree of neuron i. If all neurons do not have the same incoming degree, then we cannot solve for $\hat{\mathbf{x}}$ independent of i and the synchronous state does not even exist. Therefore, for chemical synaptic networks (4.4), rather than analyzing the stability of the synchronous state, we seek to determine how far the network is from having a synchronous solution. In the spirit of the master stability function analysis, we desire a measure of synchronizability analogous to (4.3) based only on network properties. We propose that network synchrony will decrease with the variance of the in-degrees, normalized by their mean squared

$$\sigma_d^2 = \frac{1}{d^2 N}\sum_i |d^i_{\mathrm{in}} - d|^2. \tag{4.6}$$

Connecting stability measures to network topology. Fortunately, we can merge the two normalized variances σ_μ^2 and σ_d^2 into the same measure. It turns out that, for many large networks, the eigenvalues of the Laplacian matrix $L = D - W$ are dominated by the diagonal entries D, which are the in-degrees d^i_{in}. If the connectivity matrix were normal ($WW^T = W^TW$), then this fact would follow from the Wielandt-Hoffman theorem [17], as this theorem shows that (with a suitable reordering of the eigenvalues μ_i)

$$\sum_i |\mu_i - d^i_{\mathrm{in}}|^2 < \sum_{i,j} W_{ij}^2 = \sum_{i,j} W_{ij} = Nd.$$

Therefore, the relative average squared deviation is

$$\frac{\sum_i |\mu_i - d^i_{\text{in}}|^2}{\sum_i |d^i_{\text{in}}|^2} < \frac{d}{\frac{1}{N}\sum_i |d^i_{\text{in}}|^2},$$

which will tend to be small. The worst case would be when all in-degrees d^i_{in} are the same (and equal to d), in which case the bound would be $1/d$. Variation in the degrees makes the bound smaller and the eigenvalues μ_i even closer to the in-degrees d^i_{in}.

Although our matrices W are not necessarily normal, we have observed that the eigenvalues μ_i are indeed close to the in-degrees d^i_{in}. An example is shown in figure 4.1. Thus, we infer that our two measures based on (4.1) and (4.4) are approximately equal, $\sigma^2_\mu \approx \sigma^2_d$, and will simply use σ^2_μ.

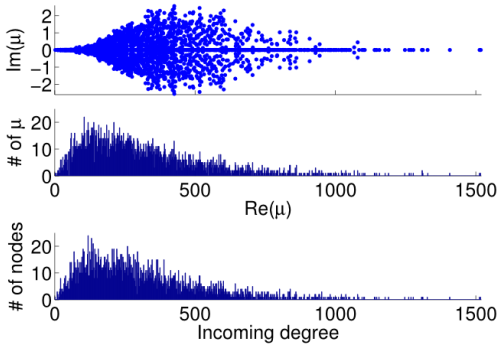


Figure 4.1: Comparison between the eigenvalues μ_i of the Laplacian and the network in-degree d^i_{in} for a sample SONET. A scatter plot of the eigenvalues is shown in the top plot. Note that the real axis scale is much larger than the imaginary axis, so the imaginary parts of the eigenvalues are relatively small. The middle panel shows a histogram of the real parts of these eigenvalues. The histogram of the in-degrees shown at bottom closely matches that of the eigenvalues. Network statistics are: $p = 0.1$, $N = 3000$, $\alpha_{\text{recip}} = 0$, $\alpha_{\text{conv}} = \alpha_{\text{div}} = \alpha_{\text{chain}} = 0.5$.

Given (3.6), we can go one step further, and connect the synchronizability measure σ^2_μ directly with the network statistic α_{conv}. Comparing (3.6) to (4.6), we see that

$$\alpha_{\text{conv}} \approx \sigma^2_d - \frac{1}{d} \approx \sigma^2_\mu - \frac{1}{d}, \tag{4.7}$$

where we have used the mean degree d of the network as a proxy for its expected value $E(d)$. Hence, we expect a higher frequency of the convergent connection motif to decrease synchrony.

4.1.2 Stability of the asynchronous state

The emergence of synchrony. An approach to perform a similar stability analysis for the completely asynchronous state of heterogeneous, noisy oscillators has been recently developed by Restrepo et al.[39, 40]. The asynchronous state is where the oscillators are as spread out evenly as though they were uncoupled (in a sense determined by the oscillator dynamics). The analysis begins with a coupled oscillator model similar to the one used in the master stability function approach (4.1):

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \mathbf{F}(\mathbf{x}_i, \eta_i) + Sk(\mathbf{x}_i) \sum_j W_{ij}(\mathbf{H}(\mathbf{x}_j) - \langle \mathbf{H}(\mathbf{x}) \rangle) + \xi_i(t),$$

where $\xi_i(t)$ is a noise term and the model is specified by the functions $\mathbf{F}$, $\mathbf{H}$ and k. One notable difference in this model is that the oscillators can be heterogeneous since their dynamics depend on the parameter vectors η_i. Two important conditions of the analysis are that the parameters cannot be correlated with the network structure and that each neuron must receive many inputs.

For the asynchronous state to be an approximate solution of this system, the factor multiplying W must be approximately zero when the neurons are asynchronous. To satisfy this condition, the equation has an extra term $\langle \mathbf{H}(\mathbf{x}) \rangle$, which is the average of $\mathbf{H}(\mathbf{x})$ over the oscillators when they are spread over the asynchronous state. Since $\langle \mathbf{H}(\mathbf{x}) \rangle$ is just a constant vector, it is less problematic than the $\mathbf{H}(\mathbf{x}_i)$ of (4.1).

Restrepo et al. analyze the coupling strength where synchrony just begins to emerge in the network (i.e., where the asynchronous state becomes unstable). Their key result is that this critical coupling strength depends on the network structure only through the largest eigenvalue $\lambda_{\max}$ of the connectivity matrix W. Clearly, this coupling strength also depends on the oscillator models and the manner in which they are coupled, but these influences are a separate factor multiplied by $\lambda_{\max}$. Given two networks with different values of $\lambda_{\max}$ but with neurons whose dynamics are governed by the same oscillator model, synchrony will emerge at a lower coupling strength for the network with the larger $\lambda_{\max}$. This separation of the oscillator dynamics from the network structure is in the same spirit as the master stability function analysis and allows us to incorporate $\lambda_{\max}$ into our notion of the synchronizability of the network.

Linking emergence of synchrony to network topology. Restrepo et al.[41] also derive approximate expressions for the largest eigenvalue $\lambda_{\max}$. In SONETs, the degree of two neighboring neurons is essentially uncorrelated. For such networks, $\lambda_{\max}$ is a simple function of the degree distribution $\lambda_{\max} \approx E(d^i_{\text{in}} d^i_{\text{out}})/E(d)$. Given the relationship (3.14) between α_{chain} and the covariance of the degree distribution, the maximum eigenvalue can be written in terms of α_{chain}:

$$\lambda_{\max} \approx \frac{\text{cov}(d^i_{\text{in}}, d^i_{\text{out}}) + E(d)^2}{E(d)} = (\alpha_{\text{chain}} + 1)E(d). \tag{4.8}$$

Hence, $\lambda_{\max}$ will be nearly zero when $\alpha_{\text{chain}} = -1$ (its minimum value) and will increase linearly with α_{chain} with slope equal to the mean degree.

4.1.3 Neuron models

We simulated many neuron models to explore the extent to which our results were model independent.

Kuramoto model. The Kuramoto model [22], originally proposed by Yoshiki Kuramoto, is a network of continuously coupled oscillators with phase θ_i (modulo 2π),

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega + \frac{S}{pN} \sum_{j \neq i} W_{ij} \sin(\theta_j - \theta_i) + \sigma \xi_i(t) \tag{4.9}$$

where ω is the natural frequency and $\xi_i(t)$ is white noise, $\langle \xi_i(t) \rangle = 0, \langle \xi_i(t) \xi_j(t') = \delta_{ij} \delta(t - t') \rangle$, which we scale by σ. We specify the connectivity strength S relative to the number of oscillators N and the connection probability p of the connectivity matrix W.

PRC model. The phase response curve networks are pulse coupled networks, meaning that the effect of the pre-synaptic neuron on the post-synaptic neuron's phase only occurs when the pre-synaptic neuron crosses the zero phase (modulo 2π), to simulate the effect of synaptic release at the time of an action potential. At the time of the action potential, the phases θ_i of the post-synaptic neurons are advanced according to their phase response curves $f_i(\theta_i)$. If a neuron fires, it is unresponsive to any other inputs it receives at that

same instant, but its phase remains at $\theta = 0$. The PRC model

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega_i + \frac{S}{pN} f_i(\theta_i) \sum_{j \neq i} W_{ij} \sum_k \delta(t - T_j^k) + \sigma \xi_i(t)$$
$$f_i(\theta) = c(a_i)\theta^{a_i}(2\pi - \theta) \tag{4.10}$$

uses the same parameters as the Kuramoto model, except in some simulations we allow the natural frequency ω_i to depend on neuron i. T_j^k is the time of kth spike of neuron j. The shape of the phase response curve f_i is determined by choice of a_i. c_{a_i} is chosen to make the maximum value of phase response curve f_i be 1. We make the average of a_i be 2 to make PRCs that would promote synchrony in a network coupled through excitatory synapses. Figure 4.2 shows some phase response curves.

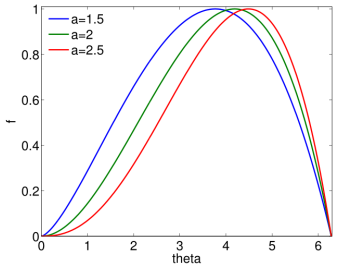


Figure 4.2: Phase response curves with $a = 1.5$ (blue), $a = 2$ (green), and $a = 2.5$ (red).

4.1.4 Spike rate normalization using an integral controller

When neurons receive different numbers of synaptic inputs, the disparity in level of synaptic drive can cause neurons to fire at different rates. In some simulations, we desired to compensate for this effect to make the neurons have similar firing rates. Rather than attempt to numerically solve for the input currents to achieve a desired firing rate, we developed an auto-tuning network by implementing an integral controller that adjusts the input current for each neuron until it fires at the desired firing rate. The controller calculates the input current for each neuron at each firing from the error between the interspike interval and that of the desired rate: $e(j) = \mathrm{isi}(j) - (\text{desired rate})^{-1}$, where $\mathrm{isi}(j)$ is the neuron's jth interspike interval. The current applied to the neuron is $I(j) = I(0) + K_i \sum_{l=0}^{j} e(l)$, where K_i is the integral feedback constant [33]. For our purposes,

we set $K_i = 1$, which produced long transients, but results in stable network behaviors. Simulation were run to steady state with the integral controller on and then the current values were held constant (virtually identical results were obtained when controllers were left on through the entire simulation).

4.1.5 Measuring synchrony with the order parameter

We use a simple quantification of network synchrony, the Kuramoto order parameter, to measure synchrony in the network simulations. The measure is based on defining a phase $\theta_j(t)$ of each neuron, which ranges from 0 to 2π to indicate the relative state of the neuron through its spiking cycle. For models that are not directly based on the phase, we define the phase as the time since the previous spike, normalized by the average period over the previous 5 spikes.

To calculate the order parameter, we represent the phase of each neuron by a vector on the unit circle at angle $\theta_j(t)$. The population vector is the average of these unit vectors. The order parameter is the length of the population vector [48]

$$r = \left| \frac{1}{N} \sum_{j=1}^{N} e^{i\theta_j} \right|, \tag{4.11}$$

where $i = \sqrt{-1}$ and $|\cdot|$ represents the absolute value of the complex vector. The order parameter ranges between 0 and 1, with $r = 1$ representing perfect synchrony. The minimum value $r = 0$ occurs when the network is asynchronous and the phases are evenly distributed around the unit circle. However, $r = 0$ whenever the population is balanced around the circle, which could also occur when two or more equal clusters are evenly spread around the circle.

4.2 Results

4.2.1 Overview

We examine the influence of network structure on the synchrony of neuronal networks. We generate networks using the framework of second order networks (SONETs) which

allows us to systematically vary second order connectivity statistics, which are the frequency of reciprocal, convergent, divergent, and chain connections. SONETs extend the commonly used Erdős-Rényi random network model [13], retaining the feature of adding minimal structure beyond what is required to match the connectivity statistics.

The analyses of perfect synchrony and asynchrony (see section 4.1) predict that two second order statistics of connectivity should play key roles in determining synchrony. First, increasing the relative frequency of the convergent connection motif (increasing α_{conv}) should decrease synchrony, as it pushes the network state further from complete synchrony. Second, increasing the relative frequency of the chain connection motif (increasing α_{chain}) should increase synchrony, as it decreases the stability of the asynchronous state.

We demonstrate in simulations that the frequency of chains and convergent motifs can influence network synchrony even when a network is not at or near the extremes of perfect synchrony and asynchrony. We present the results in four parts. First, we generate SONETs with a range of second order connectivity statistics and test the predictions of the analysis through simulations of a simple neuron model. Second, to demonstrate that these results can be explained by the analysis, we confirm that critical eigenvalue quantities from the analysis do have the predicted relationships with the network statistics. Third, to assure that our findings are not model dependent, we test networks with several different single neuron models. Fourth, we test the results with a larger range of networks to evaluate the influence of higher order network statistics on the predictions based on second order statistics.

4.2.2 Synchrony determined by second order connectivity statistics

To test the effects of the second order connectivity statistics on synchrony, we generate 186 SONETs with $N = 3000$ neurons and 10% connectivity ($p = 0.1$). In addition to using the standard Erdős-Rényi network (all α's zero), we randomly sample the second order connectivity statistics in the ranges $\alpha_{\text{recip}} \in [-1, 4]$, $\alpha_{\text{conv}} \in [0, 1]$, $\alpha_{\text{div}} \in [0, 1]$, and $\alpha_{\text{chain}} \in [-1, 1]$. These ranges are chosen to include the α's used to reproduce the motif spectrum obtained from neuronal networks published by [46]: $(\alpha_{\text{recip}}, \alpha_{\text{conv}}, \alpha_{\text{div}}, \alpha_{\text{chain}}) \approx (3, 0.4, 0.3, 0.2)$. As hinted by (3.6), the valid range of α_{chain}

depends on the α_{conv} and α_{div}. We maximize its range by sampling some networks where α_{chain} is set to its maximum and minimum value conditioned on α_{conv} and α_{div}.

We measure synchrony in network simulations of 3000 pulse coupled PRC neuron models (4.10) in these 186 SONETs. Three example networks are shown in figure 4.3. For each network, we calculate synchrony measured by the Kuramoto order parameter (4.11) as a function of time, and calculate its average once the network reaches steady state.

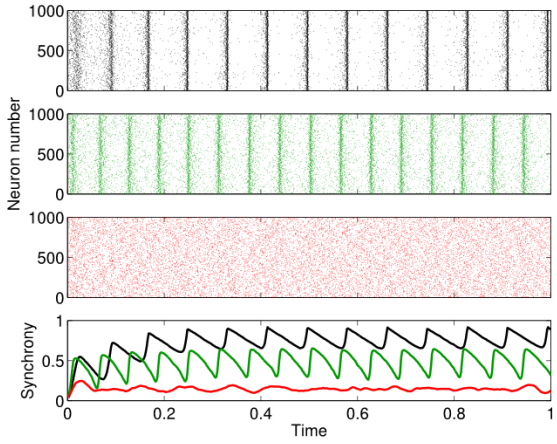


Figure 4.3: Sample output from the beginning of three network simulations with the PRC model (4.10). Top three panels are rastergrams showing the spikes of 1000 neurons (out of 3000 in each network) for the first second of time. The bottom panel shows synchrony measured by the Kuramoto order parameter (4.11) calculated as a function of time, where colors correspond to the associated rastergrams. The networks were chosen to illustrate the range of observed synchrony. The top network (in black) is the Erdős-Rényi random network ($\alpha_{\text{recip}} = \alpha_{\text{conv}} = \alpha_{\text{div}} = \alpha_{\text{chain}} = 0$) which reached high synchrony (steady state $r = 0.8$). The middle network (in green) that reached moderate synchrony ($r = 0.5$) was generated from the second order connectivity statistics $\alpha_{\text{recip}} = -0.2$, $\alpha_{\text{conv}} = 0.7$, $\alpha_{\text{div}} = 0.6$, $\alpha_{\text{chain}} = 0.6$. The bottom network (in red) that stayed fairly asynchronous ($r = 0.1$) was generated from $\alpha_{\text{recip}} = 0.1$, $\alpha_{\text{conv}} = 0.9$, $\alpha_{\text{div}} = 0.9$, $\alpha_{\text{chain}} = -0.6$. For all networks, average connectivity was $p = 0.1$. PRC model parameters were $S = 6$, $\sigma = 3$, $a_i = 2$ and $\omega_i = 60$ for all neurons i.

In figure 4.4 the steady state synchrony is plotted against the measured second order statistics α calculated by (3.14). No simple relationship between any single connectivity statistic and synchrony is obvious, though it appears the Erdős-Rényi network (black) is

among the most synchronous. There appears to be a bimodal distribution of synchrony, suggesting that synchrony rapidly jumps up when some threshold is crossed.

When synchrony is plotted against pairs of connectivity statistics, as shown in figure 4.5, the pattern of dependence on connectivity emerges. The synchrony seems to be a function of $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ alone, as the level of synchrony varies more or less smoothly in the $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ projection (lower left of figure 4.5) despite the fact that α_{recip} and α_{div} are varying widely across the whole plot. For a given value of α_{conv}, there appears to be a threshold of α_{chain} above which synchrony jumps up, supporting the trend of higher synchrony with α_{chain} predicted by our analysis. Similarly, synchrony jumps down at a threshold value of α_{conv} for a given value of α_{chain}.

Note the dependence of the valid range of α_{chain} on the other statistics (bottom row of figure 4.5). The relationship between $\alpha_{\text{conv}}/\alpha_{\text{div}}/\alpha_{\text{chain}}$ and the variances/covariance of the in- and out-degree distribution shown in (3.14) implies that the valid range of α_{chain} increases with α_{conv} and α_{div} (left two panels of last row of figure 4.5). The range of α_{chain} is also affected by α_{recip} (lower-right panel of figure 4.5). The dependence of the range of α_{chain} and the other statistics creates the appearance of spurious correlations between synchrony and $\alpha_{\text{recip}}/\alpha_{\text{div}}$ in the top two rows of figure 4.5.

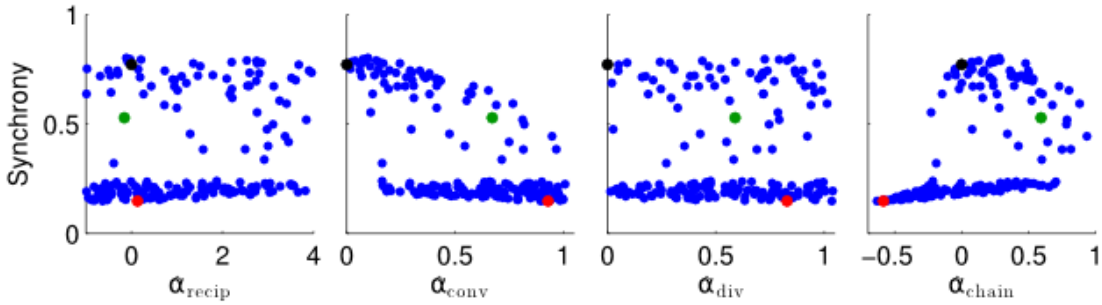


Figure 4.4: Scatter plots of steady state synchrony measured by the order parameter (4.11) as a function of individual connectivity statistics for the 186 sampled SONETs. For each panel, each dot corresponds to one network. Black, green, and red dots correspond to the networks simulated in figure 4.3. Estimates of second order connectivity statistics $\hat{\alpha}$ are determined from each network connectivity matrix W via (2.19)(2.21)(2.22)(2.23). PRC model parameters were $S = 6$, $\sigma = 3$, $a_i = 2$ and $\omega_i = 60$ for all neurons i.

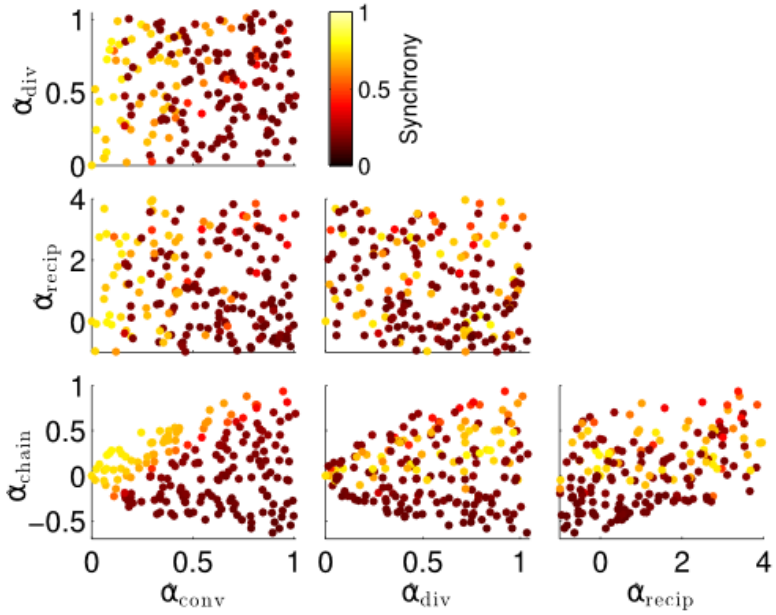


Figure 4.5: Synchrony plotted as a function of pairs of connectivity statistics. The simulation results of the PRC model over the 186 SONETs (figure 4.4) are replotted as a function of each pair of second connectivity statistics α estimated from the connectivity matrix using (3.14). Each dot corresponds to one network, and color indicates steady state synchrony measured by the order parameter (4.11). Synchrony varies smoothly in the graph with respect to α_{chain} and α_{conv} but not in the other graphs.

4.2.3 Eigenvalue analysis

The influence of the connectivity statistics on synchrony was predicted based on analysis of eigenvalues of the network connectivity matrix W and Laplacian matrix L (see section 4.1). An analysis of the asynchronous state predicts synchrony will increase with the largest eigenvalue $\lambda_{\max}$ of the connectivity matrix, which should increase linearly with α_{chain} (see (4.8)). An analysis of the synchronous state predicts synchrony will decrease as the normalized variance σ_μ^2 (4.3) of the Laplacian eigenvalues increases, which should be nearly equal to α_{conv} (see (4.7)).

In figure 4.6, sample spectrum from a few connectivity matrices demonstrate these trends: $\lambda_{\max}$ increases with α_{chain}, and only α_{conv} modulates the variance σ_μ^2 of the eigenvalues of the Laplacian L. These observations are confirmed by the plots of $\lambda_{\max}$

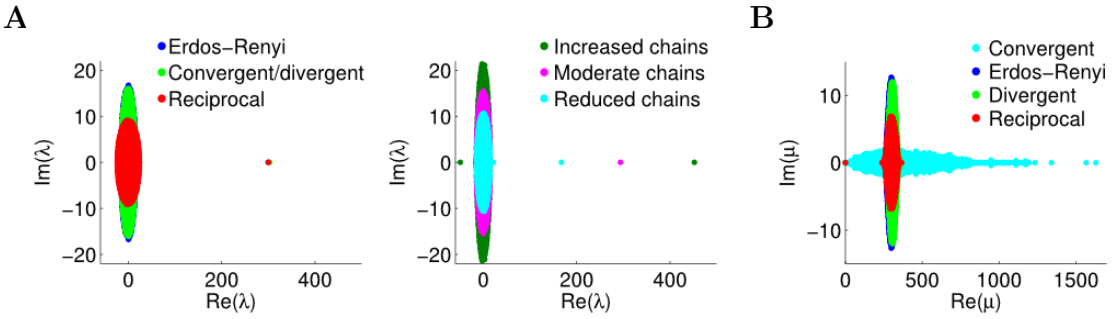


Figure 4.6: Spectra of connectivity matrices and Laplacian matrices for sample networks. All α's not mentioned were zero. **A**. Left: Eigenvalues for the Erdős-Rényi network (blue) are mainly occluded by the spectra of the other matrices. Eigenvalues for convergent and divergent networks were identical. Non-zero parameters: $\alpha_{\text{conv}} = 0.5$ (convergent), $\alpha_{\text{div}} = 0.5$ (divergent), $\alpha_{\text{recip}} = 4$ (reciprocal). Right: Largest eigenvalue $\lambda_{\max}$ increases linearly with α_{chain}. All three networks had substantial convergence and divergence ($\alpha_{\text{conv}} = \alpha_{\text{div}} = 0.5$) to allow large α_{chain}. Parameters: $\alpha_{\text{chain}} = 0.5$ (increased chains), $\alpha_{\text{chain}} = 0$ (moderate chains), $\alpha_{\text{chain}} = -0.4$ (reduced chains). **B**. Eigenvalues of the Laplacian for the networks from the left panel of A. The spread of the eigenvalues remains relatively unchanged with different frequencies of chains (not shown).

and σ_μ^2 versus each connectivity statistic for the 186 generated networks (figure 4.7). The right top panel shows how $\lambda_{\max}$ is nearly completely determined by α_{chain} according to the relationship (4.8) plotted in red. The second panel in bottom row of figure 4.7 shows that the normalized variance σ_μ^2 of the eigenvalues is nearly identical to the α_{conv} and follows the relationship (4.7) plotted in red. The other relationships between the eigenvalues and connectivity statistics simply reflect how the range of sampled values of α_{chain} depend on the other statistics, as seen in figure 4.5.

The left three panels of the top row do show some relationship between $\lambda_{\max}$ and the other three connectivity statistics. However, these relationships are completely explained by the relationship between the sampled values of α_{chain} and the other connectivity statistics α_{recip}, α_{conv}, and α_{div}. If one were to plot the distribution between α_{chain} and the other statistics for the 186 sampled networks, these plots would be nearly identical to the first three panels of figure 4.7. The bottom row of figure 4.7 shows that the normalized variance σ_μ^2 of the eigenvalues is nearly identical to the α_{conv} and follows the relationship (4.7) plotted in red.

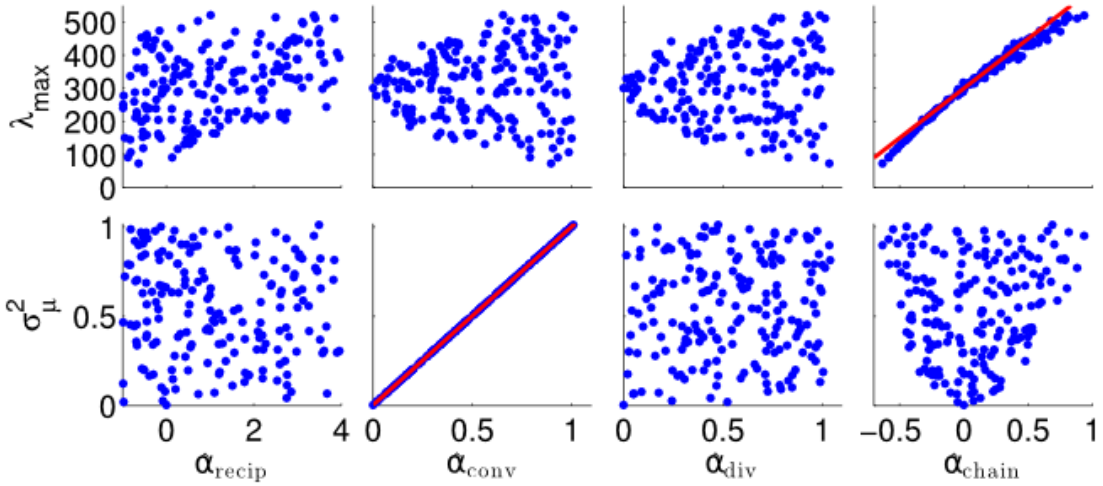


Figure 4.7: The distribution of key eigenvalue quantities as function of connectivity statistics for the 186 sampled SONETs. In each panel, each dot corresponds to one of the networks. Red line in top right panel is predicted dependence (4.8) of largest eigenvalue $\lambda_{\max}$ on α_{chain}. Red line in the second panel of the second row is relationship (4.7) between the normalized variance σ^2_μ of the eigenvalues of the Laplacian and α_{conv}.

The tight relationships between α_{chain} and $\lambda_{\max}$ and between α_{conv} and σ^2_μ confirm that the influence of these two connectivity statistics on synchrony is indeed through their influence on the key spectral properties of the connectivity matrix W and Laplacian matrix L.

4.2.4 Dependence of synchrony on single neuron model

The synchrony analysis, which was based around the completely synchronous and asynchronous states, predicts that the effect of topology on synchrony should be independent of the neuron model. Since these analyses were based on linearizations around those states, there is no guarantee those predictions will apply to other networks states, such as intermediate levels of synchrony. Nonlinearities in neuron models could lead to model-dependent deviations from those predictions.

We test for model-dependence of the results using the three neuron models described in section 4.1. The simplest model is the Kuramoto oscillator (4.9). Second, we run

additional simulations with pulse coupled PRC oscillator models (4.10) under two conditions: one with noise as in figure 4.5 and the other with heterogeneity in the model parameters. We run all models on the above 186 SONETs and observe how synchrony is affected by changing the coupling strength.

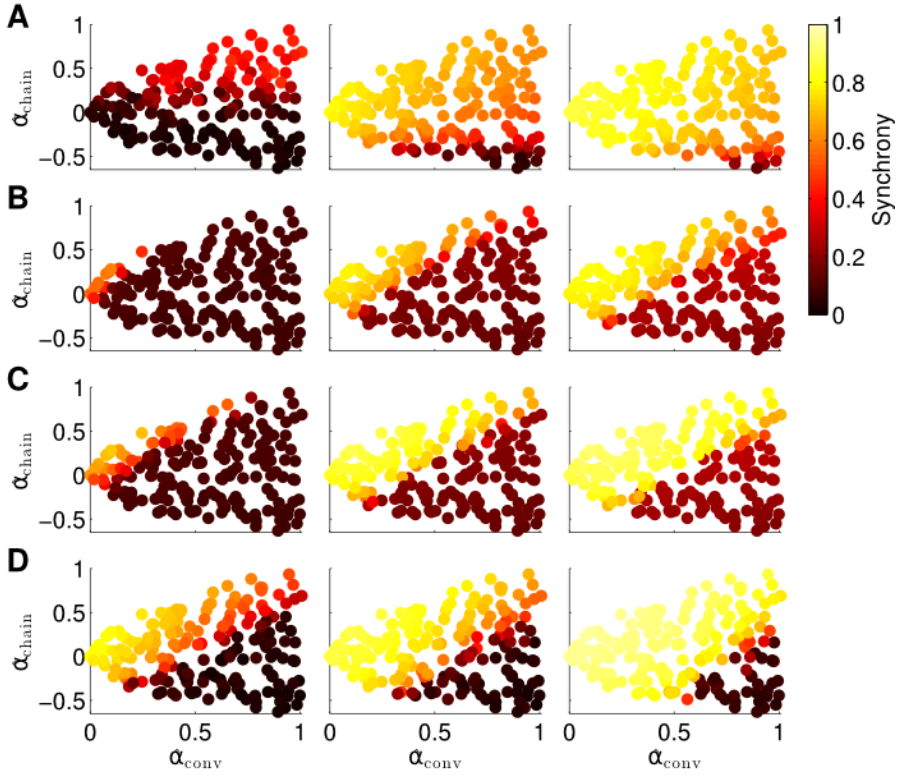


Figure 4.8: Dependence of synchrony on model and coupling strength. Each row corresponds to a different neuron model and coupling strength increases with column. Panels as in figure 4.5. **A**. Kuramoto model (4.9) with coupling strengths $S = 1$, 2, and 3. **B**. PRC model (4.10) with homogeneous parameters as in figure 4.4. Coupling strengths were $S = 3$, 6, and 9, so that middle panel is the same as lower left of figure 4.5. **C**. Noise free ($\sigma = 0$) PRC model (4.10) with heterogeneous parameters: a_i was drawn from a Gaussian of mean 2 and standard deviation 0.2; ω_i was drawn from a Gaussian of mean 60 and standard deviation 6. Coupling strengths were $S = 3$, 6, and 9.

The results shown in figure 4.8 demonstrate that for each model, the influence of the network structure on synchrony is indeed determined only by α_{conv} and α_{chain}, as in each case, synchrony appears to be a function of those parameters. The relative influence of

α_{conv} and α_{chain} does vary across models. For the Kuramoto model, synchrony is more strongly influenced by α_{chain} than α_{conv} (figure 4.8A). For the noisy homogeneous PRC network, the result from figure 4.5 is maintained at different connectivity strengths, as both convergence and chains strongly modulated the synchrony (figure 4.8B). The noise-free heterogeneous PRC network have a similar dependence on the connectivity parameters (figure 4.8C); replacing noise with variability across the population in intrinsic frequency and PRC shape only alters the connectivity strength at which synchrony emerges.

In all cases, increasing α_{chain} or decreasing α_{conv} tends to increase synchrony. The increase in synchrony appears to be concentrated at an effective threshold line through the $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ plane. Whether a particular network will synchronize or not depends on many factors, but relative changes in α_{conv} or α_{chain} affect synchrony in the same way. Since the effects of changing these two connectivity statistics combine in a model-dependent manner, we obtain a two-dimensional index $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ for the synchronizability of a network.

4.3 Synchrony with other networks

All the above simulations were based on SONETs of $N = 3000$ neurons where we fixed the first order connectivity statistics ($p = 0.1$) and varied the second order connectivity statistics α. We did not separately manipulate higher order connectivity statistics, as the definition of SONETs does not add additional structure into the connectivity beyond that determined by the second order statistics. To explore how well α_{conv} and α_{chain} determine synchrony under a wider range of networks, we simulate networks with different first order statistics and networks with different higher order statistics.

To test the effect of changing first order statistics, we decrease the connection probability by a factor of 10 to $p = 0.01$. We generate 130 SONETs over the same ranges of α as the original $p = 0.1$ networks and simulate the PRC model of figure 4.8B As shown in the left panel of figure 4.9A, the results are similar to those with $p = 0.1$. We also increase the network size to $N = 10000$ and set $p = 0.03$ so that the average degree $p(N-1)$ is approximately the same as in the original networks. Simulations from 130 SONETs demonstrate that α_{conv} and α_{chain} determine synchrony in the same manner as

before (figure 4.9A, right).

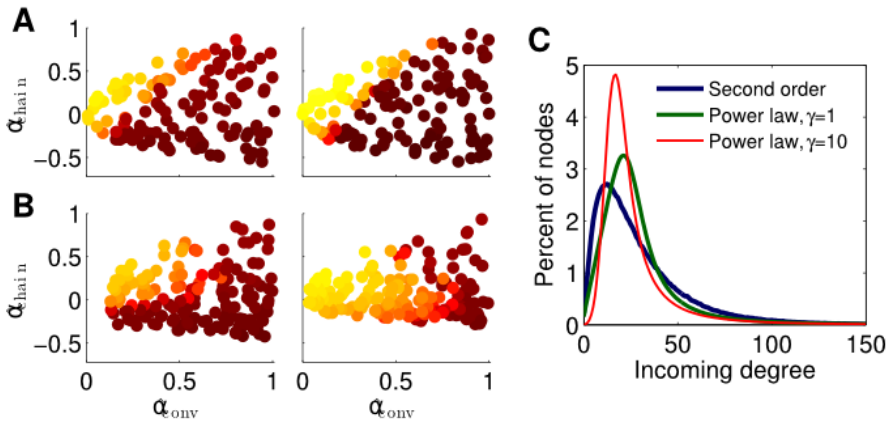


Figure 4.9: Dependence of synchrony on first order connectivity statistics and higher order connectivity statistics. All simulations use the PRC model and coupling strengths $S = 6$ of the middle panel of figure 4.8B. Note that changing p or N changes the effective coupling $S/(pN)$ in the PRC model (4.10). Pseudocolor scale indicates synchrony as in figure 4.8. **A**. SONETs with different first order statistics. Left: sparse networks with $N = 3000$ neurons and connection probability $p = 0.01$. Right: large networks with $N = 10000$ neurons and connection probability $p = 0.03$. **B**. Scale-free networks with $N = 3000$ neurons and connection probability $p = 0.01$. Rising exponents of degree distribution (A.55) $\gamma_{\mathrm{in/out}}$ were set to 1 (left) and 10 (right). For left panel, the minimum attainable value of α_{conv} and α_{div} was 0.1. Maximum degrees: $b_{\mathrm{in}} = b_{\mathrm{out}} = 300$. **C**. Incoming degree distributions for networks with $p = 0.01$ and $\alpha_{\mathrm{conv}} = 0.7$. The degree distributions are smoothed versions of the distribution of expected degree (A.55) as actual connections are generated randomly.

To manipulate higher order connectivity statistics of the network, we generate networks with power law degree distributions, i.e., with scale-free connectivity. The algorithm for generating the power law networks is described in the Appendix. The power law degree distribution has different higher order statistics from the degree distribution of SONETs, for fixed first (p) and second (α) order statistics. By comparing power law networks with SONETs that have the same first and second order statistics, we can evaluate the effect of the higher order statistics.

We exploit relationships (3.6) (3.10) (3.14) between the α's and the degree distribution to generate networks across a range of second order connectivity statistics. We vary

the power law exponents of the in-degree and out-degree distributions (A.55) independently to allow α_{conv} and α_{div} to range in $[0, 1]$. To maximize the range of α_{chain}, we let the correlation between in-degree and out-degree range in $[-1, 1]$.

We generate power law networks with $N = 3000$ neurons and average connectivity $p = 0.01$. We first generate 160 such power law networks where we set the rising exponents of the degree distributions (A.55) to $\gamma_{\text{in}} = \gamma_{\text{out}} = 1$ so that the degree distributions initially increase linearly as shown by the green trace of figure 4.9C. Synchrony for the PRC model and these power law networks is shown in figure 4.9B, left. Even for these scale free networks, synchrony does appear to be a function of just α_{conv} and α_{chain}, as the synchrony varies smoothly in the $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ projection as before. For the most part, the same qualitative trends persist that were predicted by the synchrony analysis and were observed in the SONETs; synchrony tends to increase with α_{chain} and decrease with α_{conv}. It is clear, though, that the higher order connectivity statistics are modifying the synchrony. The synchrony pattern as a function of second order connectivity statistics in the left panels of figures 4.9A and B are not identical despite the model and first order connectivity statistics being the same.

If we increase the steepness of the rise of the degree distribution by setting $\gamma_{\text{in}} = \gamma_{\text{out}} = 10$ (red trace of figure 4.9C), the dependence of synchrony on the α changes dramatically (figure 4.9D). Synchrony is still a function of just $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$, but for larger α_{conv}, synchrony no longer increases monotonically with α_{chain}. Although the eigenvalue $\lambda_{\max}$ still has the same linear relationship (4.8) with α_{chain}, this eigenvalue no longer accurately predicts the emergence of synchrony according to the analysis of the stability of the asynchronous state.

We found large values of γ_{in} also lead to a non-monotonic dependence of synchrony on α_{chain} for other degree distribution parameters. We hypothesize that this breakdown of the asynchronous state analysis is caused by the presence of many neurons with small incoming degree. One assumption of the stability analysis of the asynchronous state is that all neurons receive many inputs. When γ_{in} is large, the number of neurons with small incoming degree (red trace of figure 4.9C) is large enough to invalidate the analysis. The results indicate that the moderate number of neurons with even smaller incoming degree in the SONETs (blue trace of figure 4.9C) does not invalidate the conclusions of

the analysis.

Heterogeneity due to convergence

Since increasing α_{conv} increases the variance of the in-degree distribution, neurons in networks with high α_{conv} will have large variability in the number of synaptic connections that they receive. This heterogeneity in the input strength will lead to heterogeneity in the firing rates across neurons in the network, thus decreasing synchrony [9][11]. We investigate whether or not this heterogeneity is the only mechanism by which increasing α_{conv} decreases synchrony.

We repeat the simulation of the noisy PRC model on the SONETs (figure 4.8B), where we virtually eliminate the variance in firing rates across neurons by adjusting the intrinsic frequency ω_i of each neuron to make its average firing rate be close to 10 Hz. Rather than attempting to solve for the 3000 ω_i, we use an integral controller to creating an auto-tuning network that adjusts the ω_i (see section 4.1, where we refer to ω more generically as an input current I). Using this approach, we can reduce the standard deviation in the firing rate across neurons in each network to below 0.1 (compared to firing rate standard deviations than ranged from 0.1 to upward of 10 for the high α_{conv} networks of figure 4.8B).

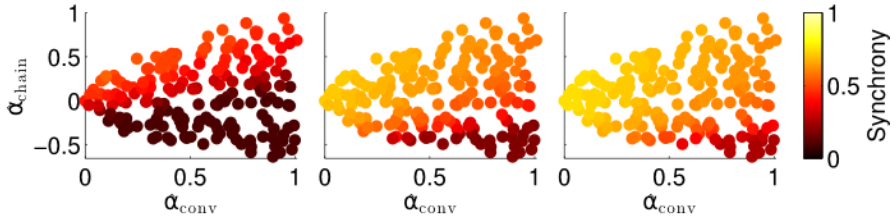


Figure 4.10: Dependence of synchrony on α_{conv} and α_{chain} when firing rate rate heterogeneity is eliminated. Simulations of (4.8)B were repeated where the ω_i were adjusted by an integral controller to fix the firing rate of each neuron at 10 Hz.

As shown in figure 4.10, the effect of α_{conv} is greatly reduced. Synchrony decreases only slightly with increased α_{conv}. The threshold value of α_{chain} where synchrony increases rapidly with α_{chain} depends only weakly on α_{conv}. The maximum value of synchrony for any α_{chain} does still decrease with α_{conv}. This residual effect of α_{conv} cannot

be explained by its effect on the heterogeneity of firing rates, as we have eliminated that heterogeneity.

4.4 Discussion

4.4.1 Linking synchrony to network motifs

In this section we present a framework to relate network connectivity to synchrony in homogeneous neuronal networks. We define network structure through the frequency of second order motifs, which form the basis for four second order statistics of connectivity. These second order connectivity statistics define conditional probabilities of connections between neurons, that is, how a connection from one neuron to a second neuron affects the the probability of (1) a synapse back onto the first neuron (reciprocal), (2) the second neuron receiving inputs from a third neuron (convergent), (3) the first neuron sending out connections to a third neuron (divergent), or (4) either the first neuron receiving inputs or the second neuron sending out connections (chain), as illustrated in Figure (2.4)(2.5)(2.6)(2.7).

The main result of this paper is that the primary influence of second order network connectivity on synchrony is through the chain and convergent connection motifs. We obtain a two-dimensional index $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ of synchronizability that captures the effect of the network structure on the synchrony of the neuronal network. Synchrony tends to increase with the relative frequency of chains and decrease with the relative frequency of convergence, where the majority of this change in synchrony occurs abruptly along a line in $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ parameter space that depends on connectivity strength and the neuron model. There is little dependence of synchrony on divergence or reciprocal connections. The result was robust to changes in first order connectivity statistics or network size.

The value of the synchronizability index $(\alpha_{\text{conv}}, \alpha_{\text{chain}})$ stems from the fact that the influence of these network structures on synchrony was similar for a wide variety of neuron models. While there were model specific differences in the simulations, the qualitative dependency of synchrony on the proportion of chains and convergent connections in the networks was the same for most single neuron models and network models. Since the influence of chains and convergence on synchrony was predicted from linearizations

around synchrony and asynchrony, it is surprising how well those statistics determined the level of synchrony over the whole range of synchrony. We expect that these linear predictions will not hold for very complicated neuron models, especially with strong coupling. The predictions did break down for networks where many neurons received few inputs, as we discovered parameter regimes where synchrony decreased with α_{chain}.

One result of the two-dimensional synchronizability index is that one cannot arrange networks in order of increasing synchronizability, and hence it may be impossible to predict if some network changes will increase or decrease synchrony without knowing the neuron model. If one network has both a higher α_{chain} and α_{conv} than a second network, it maybe the either network may elicit higher synchrony depending on the neuron model. For example, the Erdős-Rényi network with all α's zero usually synchronized better than other networks, even those with higher α_{conv} and α_{chain}. However, this trend was not model-independent, as we observed some exceptions for low connectivity strength with Kuramoto model (figure 4.8A) (or when we eliminated the firing rate heterogeneity caused by positive α_{conv}).

4.4.2 The decrease of synchrony with convergence

The decrease in synchrony with convergence is not surprising, given that the convergence is correlated with the variance in the in-degree distribution (3.6) and therefore results in increased heterogeneity in firing rates. The effect of this in-degree heterogeneity can be reduced making the synaptic strengths inversely proportional to the in-degree as done by [28] or by adjusting the current to each cell so that the firing rates are the same, as we have done in our auto-tuned networks. Reducing the heterogeneity greatly reduced the effect of α_{conv} on synchrony but some residual effect of α_{conv} on synchrony remained indicating higher order effects of α_{conv} on synchrony that cannot be accounted for by firing rate heterogeneity alone.

At first glance, the decrease of synchrony with α_{conv} seems to stand in contradiction of recent results by [43], where they find that the build up of synchrony in feedforward networks is primarily due to the level of convergence in the network. They show that when neurons receive many inputs (have high d_{in}), any small correlation among their inputs is amplified. This pooling of many inputs averages out independent fluctuations but

not the correlations. Thus, increasing convergence increases correlations and synchrony. The apparent contradiction is resolved by observing that Rosenbaum et al. change the level of convergence by modifying the first order statistics of connectivity rather than the second order statistic α_{conv}. They do not introduce the heterogeneity of increased in-degree variance.

4.4.3 Independence of synchrony from common input

One surprising result is that the synchronizability of a network appears to be unaffected by the relative frequency of divergent connection motifs as captured by α_{div}. Intuitively, one would imagine that the common input represented by the divergent connection motif would tend to synchronize neurons, as the common input leads to correlations in their inputs [45]. In feedforward networks, such correlations are amplified and can lead to synchrony in deeper layers of the network [12, 42, 50, 53, 25].

To see the relationship between α_{div} and shared input, define β to be the fraction of shared input, i.e., the ratio between the expected number of shared connections between a pair of neurons and the expected number of connections to a single neuron. For SONETs, we can calculate from (2.7) how β depends on divergence

$$\beta = \frac{1}{p(N-1)} \sum_{j \notin \{i,k\}} \Pr(W_{ij} = 1, W_{kj} = 1) \approx p(1 + \alpha_{\text{div}}),$$

where the last approximation is in the limit of large N. As an extreme test to see if the common input determined by α_{div} would lead to increased synchrony, we generated SONETs with connection probability $p = 0.1$ where α_{div} varied over its full range $\alpha_{\text{div}} \in [0, 1/p - 1] = [0, 9]$, leading to shared input in the range $\beta \in [0.1, 1]$. Such large values of α_{div} are not realistic, as shown by the out-degree distributions of figure 4.11A. When α_{div} is near its maximum, the out-degrees are nearly all zero or $N - 1$: most neurons do not have any projections to others neurons while a small fraction (around pN) project to nearly every other neuron. Nonetheless, we wanted to test to see if such extreme cases, where neurons share almost all of their pre-synaptic connections, would lead to higher synchrony. However, as shown in figure 4.11B, even large values of α_{div} had little influence on synchrony. An ensemble average over many networks does unmask a small

increase of synchrony with α_{div} (not shown), but this effect is dwarfed by increases in synchrony due to α_{chain}.

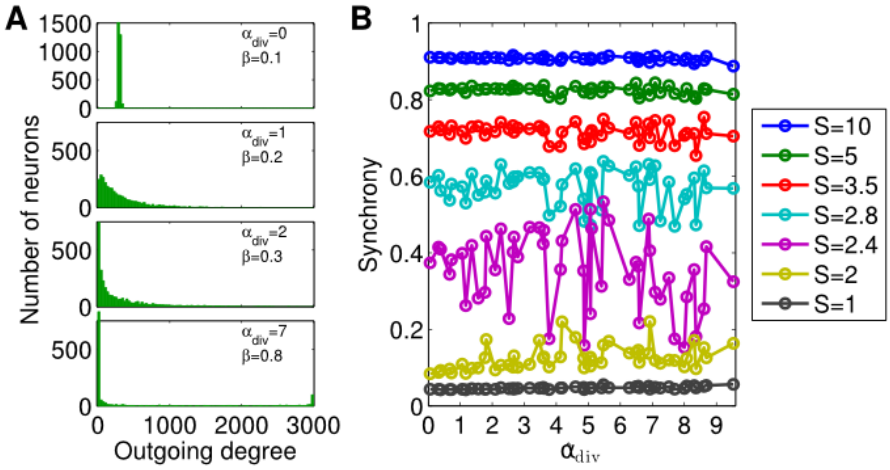


Figure 4.11: Demonstration of the lack of influence of divergent connections on synchrony. Each network was a SONET of $N = 3000$ neurons with $p = 0.1$ and $\alpha_{\text{recip}} = \alpha_{\text{conv}} = \alpha_{\text{chain}} = 0$. **A**. Sample outgoing degree distributions demonstrate how large values of α_{div} are unrealistic. For $\alpha_{\text{div}} = 7$, nearly 2400 neurons have no projections (bar at smallest in-degree is truncated) and nearly 100 neurons project to all other neurons. **B**. For each of 50 networks with different values of α_{div}, the PRC model (4.10) was simulated using 7 different coupling strengths S. Since the same networks are used, fluctuations due to the particular network structure are similar for each connectivity strength. Values of $\widehat{\alpha}_{\text{div}}$ can exceed the theoretical maximum of $1/p - 1 = 9$ as they are estimates from the generated connectivity matrix using (2.22). PRC model parameters are the same as used in figure 4.8B.

Other researchers have also found that high common input may not necessarily lead to large spike correlations.[1] [44] found that changing common input by manipulating the outgoing degree distribution typically had little impact on correlations in networks of excitatory and inhibitory neurons. [38] similarly showed that common input had little influence on spiking correlation, but since they used Erdős-Rényi networks where $\beta \approx p$, they could not isolate changes in common input β from changes in the first order connection probability p. Since our networks do not have inhibition, our results cannot

[1] Since we are looking at oscillating neurons, we obtain equivalent results if we measure spike count correlation rather than synchrony. As long as one counts spikes in bins sufficiently smaller than the oscillation period, the synchrony induces spike correlations.

be explained by the opposing effect of inhibition canceling out correlations in input currents [38]. Instead, we are left to conclude that common input due to divergence only weakly influences correlation, when it isn't combined with convergence and chains.

4.4.4 Higher order network structure

In the SONETs, the synchronizability of the networks was determined by the second order connectivity statistics of convergence and chains. However, higher order network structure can also influence synchronizability. We introduced higher order connectivity statistics of one particular type: the additional structure contained in the power law in-degree and out-degree distributions of scale free networks. Adding this higher order structure did alter the synchrony, even if we kept the second order connectivity statistics fixed. This simple manipulation demonstrated that knowing the second order connectivity statistics may not be sufficient to fully know the effect of the network on synchrony. Nonetheless, within a particular class of networks, we did see the same qualitative dependence of synchrony on the second order statistics, where only convergence and chains had an influence. As long as there weren't too many neurons that received few inputs, synchrony tended to increase with chains and decrease with convergence as it did with the SONETs.

The differences observed from the in-degree distribution of the power law networks could likely be captured by adding one third-order connectivity statistic: the probability of three connections onto a single neuron, which would presumably be linked to the skew of the in-degree distribution in analogy to (3.6). However, there are many more third-order connection motifs (12 in all) so that dipping into third-order statistics would greatly increase the complexity. We expect, though, that only a subset of those third-order statistics would influence synchrony.

Other higher order network statistics that have captured a lot of attention are the clustering coefficient and mean path length often examined in the context of small world networks [49]. A number of studies have looked at the influence of network connectivity on neuronal activity and synchrony in the context of small world models [19, 21, 29]. In the classical small world models, all the second order network statistics α are zero except α_{recip}, so one cannot observe the effect of α_{conv} or α_{chain} in these models. For the

networks parameters we used, the influence of these second order statistics on synchrony was substantially larger than the influence of the small world rewiring parameter.

While the brain may have higher order connectivity statistics, there are several advantages to using second order statistics to describe a network. As in general, lower order statistics requires less data to estimate than do higher order statistics; it may be possible to estimate the five SONET parameters from experimental data but harder to estimate the higher order statistics. Furthermore, higher order statistics should be compared to the null hypothesis that they can be explained by the second order statistics observed. Finally, the low-dimensional framework of SONETs provides a mechanism for systematically exploring a wide range of network structures and quantifying their differences.

Chapter 5

A kinetic theory approach for all-to-all coupled networks

5.1 Kinetic theory approach to a single integrate-and-fire neuron model

Neurons are elementary processing units in nervous systems. The membrane potential of a neuron is the electrical difference between the inside surface of membrane and the surrounding extracellular medium. When a signal arrives at a postsynaptic terminal, it triggers a synaptic event in the postsynaptic neuron, leading to a transient change in its membrane voltage. If the change is positive, the synapse is said to be excitatory. If the membrane potential of a neuron is above a threshold, the neuron generates an action potential (spike). Neuronal networks are made up of biological neurons connected via synapses.

The integrate-and-fire model is a simplified single neuron model. Assume that during the action potential, the membrane potential follows an infinitely fast trajectory. We impose a voltage threshold v_{th}. Once the membrane potential reaches v_{th}, we claim a spike occurs. Then the membrane potential is set to a hyperpolarized voltage v_{reset} with $v_{reset} < v_{th}$. The spiking times of the neuron t_{sp} are recorded as the spike train. $t_{sp}^- = \lim_{\delta t \to 0^+}(t_{sp} - \delta t)$, and $t_{sp}^+ = \lim_{\delta t \to 0^+}(t_{sp} + \delta t)$, which is the time instantly after

the spiking time.

$$\begin{aligned} \tau_m \frac{dV}{dt} &= \varepsilon_r - V + R_m I_e \\ V(t_{sp}^+) &= v_{reset} \text{ when } V(t_{sp}^-) = v_{th} \end{aligned} \tag{5.1}$$

where τ_m is the membrane time constant. The dynamics of the single neuron membrane potential can be illustrated by the following figure.

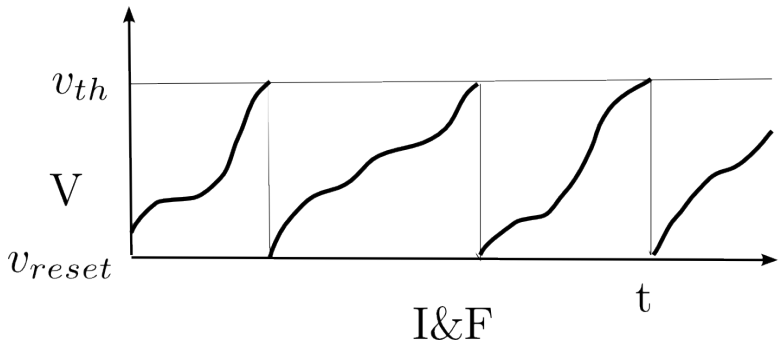


Figure 5.1: Membrane potential of I&F neuron

The population density approach turns the description of single neuron trajectories to a probabilistic description of neuron population dynamics. By constructing an ensemble of trajectories that corresponds to independent realizations of Poisson inputs with mean $\lambda(t)$, $\rho(v,t)dv$ describes the probability that the membrane potential of neurons falls in $[v, v+dv)$.

The evolution equation is based on the conservation of probability, $Pr(V(t) \in (v_0, v)) = \int_{v_0}^{v} \rho(\theta, t)d\theta$. Let $J(v,t)$ describes the probability flux at v at time t, which means the probability per unit time of crossing v. Thus

$$J(v_0,t) - J(v,t) = \frac{\partial}{\partial t}\Pr(V(t) \in (v_0, v)) = \frac{\partial}{\partial t}\int_{v_0}^{v} \rho(\theta, t)d\theta \tag{5.2}$$

Equation(5.2) states that the rate of change of probability in interval $[v_0, v]$ is given by the rate the probability that enters the interval minus the rate that leaves. Take the derivative of (5.2) with respect to v, we have

$$\frac{\partial \rho}{\partial t}(v,t) = -\frac{\partial J}{\partial v}(v,t), \text{ for } \varepsilon_i < v < v_{th} \tag{5.3}$$

Note that whenever the voltage of a neuron reaches threshold, it fires a spike and is reset to v_{reset}. This brings a source of probability at v_{reset}. The evolution of the probability density becomes

$$\frac{\partial \rho}{\partial t}(v,t) = -\frac{\partial J}{\partial v}(v,t) + \delta(v - v_{reset})J(v_{th},t) \tag{5.4}$$

The firing rate at time t is the probability a neuron fires in unit time, and it equals $J(v_{th},t)$.

$J(v,t) = J_l(v,t) + J_e(v,t) + J_i(v,t)$, where J_l, J_e, J_i represent the probability flux due to leakage, excitatory and inhibitory synaptic inputs respectively. $J_l(v,t)$ is the leakage flux toward ε_r, and equals the product of the probability density and the voltage change rate due to leakage.

$$J_l(v,t) = -\frac{1}{\tau}(v - \varepsilon_r)\rho(v,t) \tag{5.5}$$

For voltage v, if a neuron with membrane potential $\theta(\theta < v)$ receives an excitatory synaptic input of size bigger than $(v - \theta)$, it will jump pass the threshold. Let random variable A_e denote the jump size when receiving a synaptic input, and $F_{A_e}(v)$ be the complimentary cumulative distribution of A_e. Since the excitatory synaptic input time is independent of the jump size, the probability of the neuron at θ crossing v in unit time is $\lambda_e(t)F_{A_e}(v - \theta)$. $J_e(v,t)$ is the sum over all contributions of all infinitesimal intervals from ε_r to v:

$$J_e(v,t) = \lambda_e(t)\int_{\varepsilon_i}^{v} F_{A_e}(v - \theta)\rho(\theta,t)d\theta \tag{5.6}$$

Similarly, the probability flux due to inhibitory synaptic input can be obtained:

$$J_i(v,t) = -\lambda_i(t)\int_{v}^{v_{th}} F_{A_i}(\theta - v)\rho(\theta,t)d\theta \tag{5.7}$$

The firing rate $r(t)$ is the probability per unit time that a neuron crosses threshold, thus

$$r(t) = \lambda_e(t)\int_{\varepsilon_i}^{v_{th}} F_{A_e}(v - \theta)\rho(\theta,t)d\theta \tag{5.8}$$

The probability density is governed by

$$\begin{aligned}\frac{\partial \rho}{\partial t} = &-\frac{\partial}{\partial v}[-\frac{1}{\tau}(v - \varepsilon_r)\rho(v,t) + \lambda_e(t)\int_{\varepsilon_i}^{v} F_{A_e}(v - \theta)\rho(\theta,t)d\theta \\ &- \lambda_i(t)\int_{v}^{v_{th}} F_{A_i}(\theta - v,t)\rho(\theta,t)d\theta] \\ &+ \delta(v - v_{reset})r(t)\end{aligned} \tag{5.9}$$

5.2 A kinetic theory approach for all-to-all coupled networks

Traditional population density methods modeling neural networks assume that neurons within a given population are independent, ignoring correlations caused by their interactions, and ignore the interactions among themselves. However, interactions within the same neural population play an important role for the neural dynamics. Thus it is important to study the higher order statistics of the activity of coupled neurons.

We develop a population density method that captures second-order statistics of interacting neuronal populations. Our work is focused on all-to-all networks, where every neuron has a connection onto every other neuron. We base the model on current-driven integrate-and-fire neurons with instantaneous synapses. To make the problem simpler, we assume that there are only excitatory synapses in the model. When a neuron fires, all the other neurons have an equal probability of being affected. When a neuron receives a synaptic input, its membrane potential get a upward kick. Since this may cause the membrane potential of postsynaptic neurons to cross threshold, there may be a "chain reaction": firing neurons promote the membrane potential of postsynaptic neurons, causing them to pass across threshold and fire, and so on. There is a constraint that each neuron can only fire once in a burst, which corresponds to physiological phenomenon of refractoriness of neuron. In our model, the refractory time is 0.

The "chain reaction" of the large population leads to a high dimensional activity. Thus dimensional reduction for the dynamic system is essential. One possibility is to ignore the higher order interactions between neurons. There is some evidence that keeping up to second order statistics (Schneidman et al 2006; Shlens et al 2006, Tang et al 2008, Yu et al 2008)maybe sufficient to predict higher order statistics in the neuronal networks of the brain. Hence, we aim to develop a model that captures such second-order statistics.

In section (5.3), I give the description of the model, and derive third order truncation and fourth order truncation of kinetic theory equation for neural network with population size N. Section (5.4) shows the kinetic theory results for some parameters.

5.3 Kinetic theory equation for all-to-all coupled neural networks

5.3.1 Model description

In our model, we assume there are N neurons with the same biophysical properties and statistics. The neurons receive independent external Poisson inputs, as well as synaptic inputs from presynaptic neurons. The membrane potential V_i of neuron i is governed by

$$\frac{dV_i}{dt} = -\frac{1}{\tau}(V - \varepsilon_r) + \sum_{j \neq i} \sum_{k} A\delta(t - t^k) + \sum_{jk} W\delta(t - t_j^k) \tag{5.10}$$

when $v_0 < V < v_{th}$. where τ denotes the time constant, and ε_r denotes the reversal potential. t^k is the arrival time of the k-th external input to neuron i. t_j^k is the arrival time of the k-th action potential generated by neuron j. Once a neuron receives an external input, its membrane potential will be promoted by a random variable A with probability density distribution f_A. When a neuron receives a synaptic input generated by presynaptic neurons, its membrane potential will jump with size W. The probability density function for random variable W is f_w. The neuron population follows the rule that once the membrane potential of neuron crosses v_{th}, it will be reset to v_0.

In our all-to-all coupled network, chain reaction is an important property. Figure(5.2) shows one possibility of the chain reaction for all-to-all coupled network with population size $N = 4$. In figure (5.2), neuron 1 fires, which promotes the membrane potentials of all the other neurons. The membrane potential of neuron 2 crosses threshold while neuron 3 and 4 reach subthreshold voltages. The fire of neuron 2 causes neuron 3 and 4 jump simultaneously, thus neuron 3 and 4 jump past threshold. The four neurons fire at the same time and then their membrane potentials are all set to v_{reset}.

5.3.2 Derivation of kinetic theory equation

The neural network with population size $N = 3$ is a simple setting of the N neural network. By deriving the kinetic theory equation for $N = 3$, the identical terms appear in the general case. In addition, it describes how the simple case pertains to general case. The general case of population N will be presented in section 5.3.

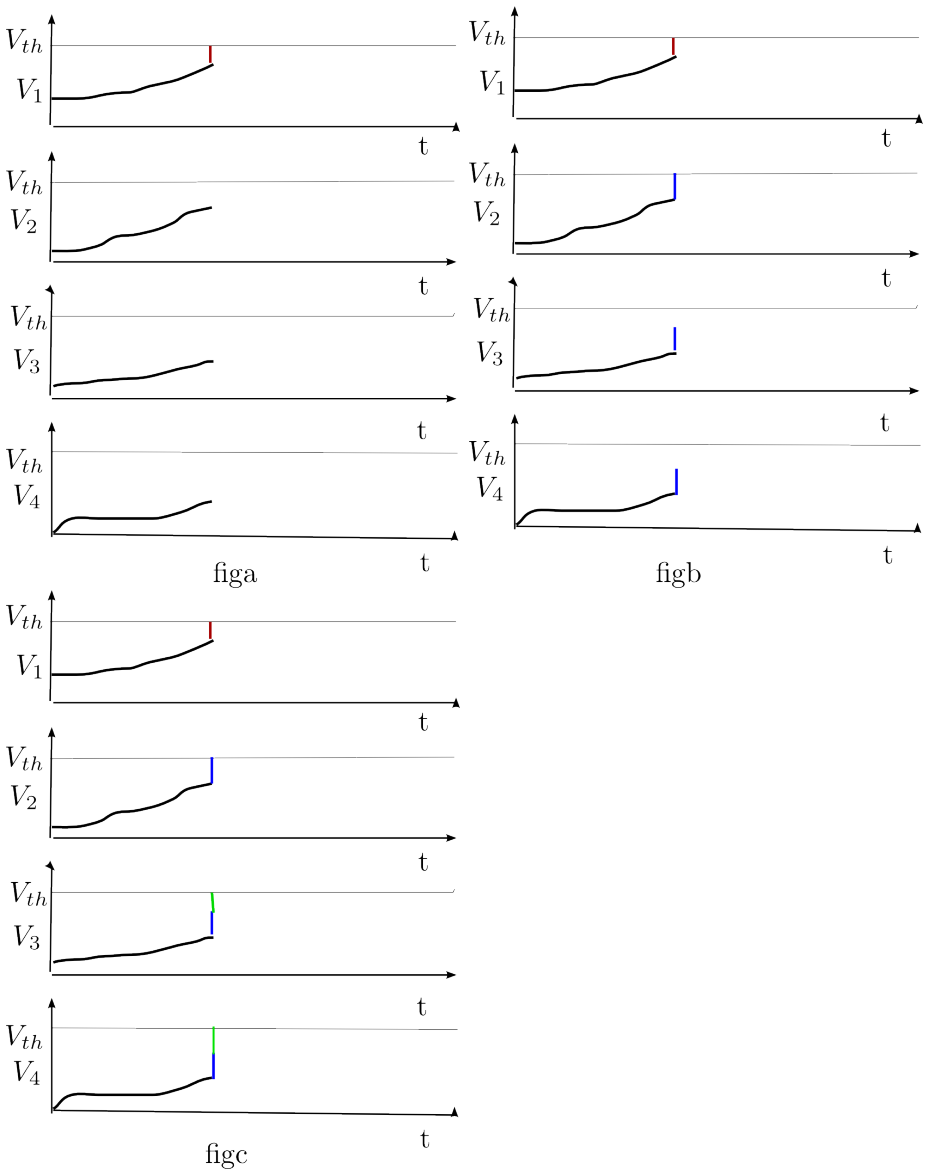


Figure 5.2: One possibility of chain reaction

Below is the network for population size $N = 3$.

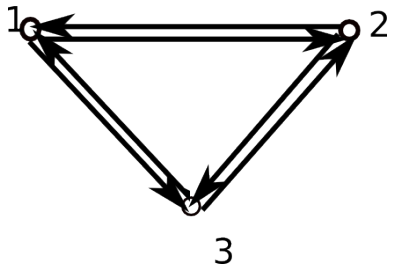


Figure 5.3: All-to-all coupled network with $N = 3$

Kinetic theory equation of $\rho(v_1, v_2, v_3, t)$ for $N = 3$

We represent $V_k(t)$ as the voltage state of neuron k at time t, and it could be described by the population density function $\rho(v_1, v_2, v_3, t)$.

$$\int_\Omega \rho(v_1, v_2, v_3, t) dv_1 dv_2 dv_3 = Pr((V_1(t), V_2(t), V_3(t)) \in \Omega)$$

where Ω is any region in the $v_1 - v_2 - v_3$ space.In a large population,the above integral can be interpreted as the expectation of neuron triples in region Ω.

The history of synaptic input does not influence the probability evolution of the voltage triple $(V_1(t), V_2(t), V_3(t))$. Therefore, $(V_1(t), V_2(t), V_3(t))$ is a Markov process. It consists of deterministic voltage evolution due to leakage, jump processes due to input,

and the voltage reset after firing a spike. The following is the differential Champam-Kolmogorov equation describing this Markov process:

$$\begin{aligned}
&\frac{\partial \rho(v_1,v_2,v_3,t)}{\partial t} = \frac{1}{\tau}\frac{\partial}{\partial v_1}[(v_1-\varepsilon_r)\rho(v_1,v_2,v_3,t)] \\
&+\frac{1}{\tau}\frac{\partial}{\partial v_2}[(v_2-\varepsilon_r)\rho(v_1,v_2,v_3,t)] \\
&+\frac{1}{\tau}\frac{\partial}{\partial v_3}[(v_3-\varepsilon_r)\rho(v_1,v_2,v_3,t)] \\
&+\lambda(t)[\int_{v_0}^{v_1} f_A(v_1-\theta_1)\rho(\theta_1,v_2,v_3,t)d\theta_1 - \rho(v_1,v_2,v_3,t)] \\
&+\lambda(t)[\int_{v_0}^{v_2} f_A(v_2-\theta_2)\rho(v_1,\theta_2,v_3,t)d\theta_2 - \rho(v_1,v_2,v_3,t)] \\
&+\lambda(t)[\int_{v_0}^{v_3} f_A(v_3-\theta_1)\rho(v_1,v_2,\theta_3,t)d\theta_3 - \rho(v_1,v_2,v_3,t)] \\
&+\sum_{i=1}^{3}\delta(v_i-v_0)J_{i,reset} \\
&+\sum_{1\le i<j\le 3}\delta(v_i-v_0)\delta(v_j-v_0)J_{ij,reset} \\
&+\delta(v_1-v_0)\delta(v_2-v_0)\delta(v_3-v_0)J_{123,reset}
\end{aligned} \tag{5.11}$$

The first three lines of equation (5.11) are the advection terms due to leakage which makes the voltage decays exponentially to ε_r. The fourth line is the independent voltage jump of neuron 1 caused by the external Poisson inputs. It contains the contribution corresponding to neuron 1 with voltage θ_1 receiving an external input of size $v_1-\theta_1$ and landing at v_1. The second term of the fourth line is the loss of probability when neuron 1 at voltage v_1 jumps to higher voltage due to external Poisson input. The fifth and sixth line are identical to the fourth line with neurons changed. The seventh line $\delta(v-v_0)J_{i,reset}$ contains the terms due to the firing of neuron i. $J_{1,reset}$ is defined by

$$\begin{aligned}
&J_{1,reset}(v_2,v_3,t) = \lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_2}\int_{v_0}^{v_3} F_A(v_{th}-\theta_1)f_w(v_2-\theta_2)f_w(v_3-\theta_3) \\
&\times \rho(\theta_1,\theta_2,\theta_3)d\theta_1 d\theta_2 d\theta_3
\end{aligned} \tag{5.12}$$

$J_{1,reset}(v_2,v_3,t)$ describes the event that neuron 1 receives an external Poisson input that promotes its voltage past threshold, which causes neuron 2 and neuron 3 to jump

simultaneously. Neuron 1 with voltage $(V_1(t) = \theta_1)$ receives an input of size bigger than $(v_{th} - \theta_1)$, thus its voltage crosses the threshold and is reset to voltage v_0. Due to the network connection, $(V_2(t), V_3(t)) = (\theta_2, \theta_3)$ receives the synaptic input of size $(v_2 - \theta_2, v_3 - \theta_3)$ and arrives at the subthreshold voltages (v_2, v_3). $J_{2,reset}$ and $J_{3,reset}$ are the symmetric terms of independent external input causing neuron 2 or neuron 3 fires.

$J_{12,reset}$ defines the terms due to firing of two neurons 1 and 2.

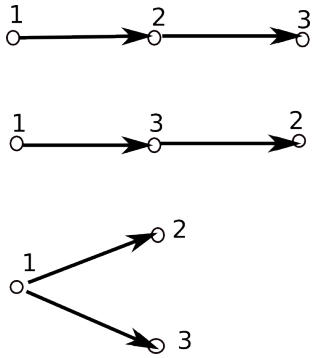
$$
\begin{aligned}
&J_{12,reset}(v_3, t) = \\
&\lambda(t)[\int_{v_0}^{v_{th}} \int_{v_0}^{v_{th}} \int_{v_0}^{v_3} F_A(v_{th} - \theta_1) F_w(v_{th} - \theta_2) f_{2w}(v_3 - \theta_3) \rho(\theta_1, \theta_2, \theta_3, t) d\theta_1 d\theta_2 d\theta_3 \\
&\quad + \int_{v_0}^{v_{th}} \int_{v_0}^{v_{th}} \int_{v_0}^{v_3} F_A(v_{th} - \theta_2) F_w(v_{th} - \theta_1) f_{2w}(v_3 - \theta_3) \rho(\theta_1, \theta_2, \theta_3, t) d\theta_1 d\theta_2 d\theta_3]
\end{aligned}
\tag{5.13}
$$

$J_{12,reset}$ describes the probability that the neuron triple $(V_1(t), V_2(t), V_3(t)) = (\theta_1, \theta_2, \theta_3)$ gets to (v_{th}, v_{th}, v_3). The first term of $J_{12,reset}$ reflects the event that neuron 1 receives an external input bigger than $(v_{th} - \theta_1)$, then its voltage is reset to v_0. It then gives a synaptic input to neuron 2 and neuron 3 simultaneously. The jump size of neuron 2 is bigger than $(v_{th} - \theta_2)$, so that it fires and is reset, and gives another synaptic input to neuron 3. Therefore, neuron 3 has a double jump with a jump size $(v_3 - \theta_3)$. The second term of $J_{12,reset}$ is symmetric to the first term with the roles of neurons 1 and 2 reversed. Let f_{2w} denote the probability density function of the double jump size, that is $f_{2w}(x) = \int_{v_0}^{x} f_w(t) f_w(x - t) dt$. F_{2w} describes its complimentary cumulative distribution function, that is $F_{2w}(x) = \int_{x}^{\infty} f_{2w}(t) dt$.

$J_{13,reset}, J_{23,rest}$ are identical to $J_{12,reset}$ with roles of neurons switched. $J_{123,reset}$ defines the terms due to firing of all three neurons, and it is described by:

$$
\begin{aligned}
&J_{123,reset}(t) = \\
&\Big\{\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_2)[F_{2w}(v_{th}-\theta_3)-F_w(v_{th}-\theta_3)] \\
&+ F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)[F_{2w}(v_{th}-\theta_2)-F_w(v_{th}-\theta_2)] \\
&+ F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_2)F_w(v_{th}-\theta_3) \\
&+ \text{symmetric terms}\Big\}\rho(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3
\end{aligned}
\tag{5.14}
$$

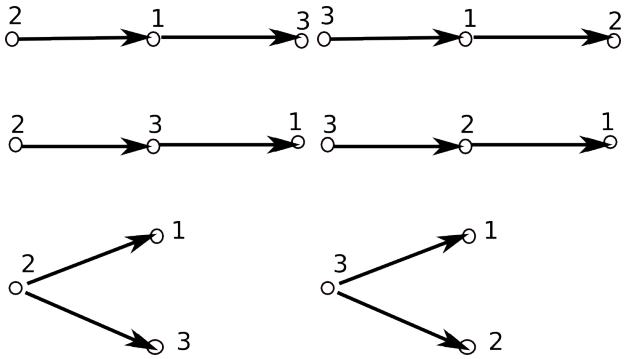
The first term of $J_{123,rest}(t)$ reflects the event that neuron 1 receives an independent external Poisson input which promotes its membrane potential past threshold and resets it to v_0. The resulting synaptic inputs cause neuron 2 and neuron 3 jump simultaneously. Neuron 2 jumps past threshold but neuron 3 jumps to a subthreshold voltage. Then neuron 2 causes neuron 3 to jump past threshold. The second term of $J_{123,reset}(t)$ describes the symmetric chain reaction with the roles of neurons 2 and 3 reversed. The third term is neuron 1 jumping past threshold, giving neuron2 and neuron3 one jump, which causes both of them to fire. Figure (5.3.2) illustrates the first three terms of $J_{123,reset}$:



$J_{123,reset}$ also includes six more terms that neuron 2 or neuron 3 receiving external Poisson inputs and causing the others to fire. They are described by figure(5.3.2):

Kinetic theory equation of $\rho(v_1, v_2, t)$ for $N = 3$

When we study the general case for N neurons, the equation for the full density will become very complicated. However, we are interested in describing just second-order



statistics. We can do this by deriving an equation for the density describing the state of any pair of neurons in the population. We illustrate the approach for the special case of the network containing three neurons.

In last section we describe the three dimensional probability density equation for an all-to-all coupled network with N=3. We want to study the statistics of coupled neurons in that network. Thus we integrate it over v_3, and get a two dimensional equation for (v_1, v_2). This reflects evolution of the population density of a neuron pair neuron 1 and neuron 2 at state (V_1, V_2). The random walk of $(V_1(t), V_2(t))$ is illustrated in the following figure:

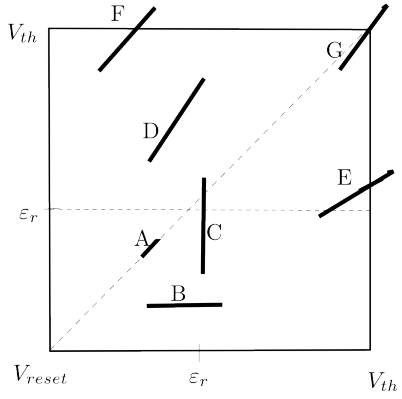


Figure 5.4: Random walk of the voltages of neuron pair 1 and 2

The following is the corresponding kinetic theory equation:

$$\begin{aligned}
\frac{\partial \rho(v_1,v_2,t)}{\partial t} &= \frac{1}{\tau}\frac{\partial}{\partial v_1}[(v_1-\varepsilon_r)\rho(v_1,v_2,t)] + \frac{1}{\tau}\frac{\partial}{\partial v_2}[(v_2-\varepsilon_r)\rho(v_1,v_2,t)] \\
&+ \lambda(t)[\int_{v_0}^{v_1} f_A(v_1-\theta_1)\rho(\theta_1,v_2,t)d\theta_1 - \rho(v_1,v_2,t)] \\
&+ \lambda(t)[\int_{v_0}^{v_2} f_A(v_2-\theta_2)\rho(v_1,\theta_2,t)d\theta_2 - \rho(v_1,v_2,t)] \\
&- \lambda(t)\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)\rho(v_1,v_2,\theta_3,t)d\theta_3 \\
&+ \lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_1}\int_{v_0}^{v_2} F_A(v_{th}-\theta_3)f_w(v_1-\theta_1)f_w(v_2-\theta_2)\rho(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3 \\
&+ \sum_{i=1}^{2}\delta(v_i-v_0)J_{i,reset} + \delta(v_1-v_0)\delta(v_2-v_0)J_{12,reset}
\end{aligned} \tag{5.15}$$

Let $(V_1(t), V_2(t))$ represent the voltage states of neuron 1 and neuron 2. The first three lines are analogous to these in equation (43). The new terms appear in lines four and five describes the influence of neuron 3 to the pair. The fourth line describes the loss of probability of neuron pair at (v_1, v_2) when neuron 3 passing threshold, causing the neuron pair at (v_1, v_2) to jump to higher voltages. The fifth line reflects the event that neuron 3 fires causes neurons at (θ_1, θ_2) to receive a synchronous input and jumps with size $(v_1 - \theta_1, v_2 - \theta_2)$. Note that these terms contain the full density $\rho(v_1, v_2, v_3, t)$ so that the equation is not a closed system for $\rho(v_1, v_2, t)$. We describe how we close the system below.

$J_{1,reset}$ contains terms that neuron 1 fires and neuron 2 jumps to a subthreshold. $J_{2,reset}$ is identical with roles of neurons reversed. $J_{12,reset}$ contains terms when both neurons fire synchronously. They are given in the appendix.

Closure of kinetic theory equation for $N = 3$

Note that the kinetic theory equation for $N = 3$ (5.15) is not closed since it involves the full probability density $\rho(v_1, v_2, v_3, t)$. Therefore, we should find a method to approximate $\rho(v_1, v_2, v_3, t)$ and make the equation closed.

Our closure method is based on the conditional independence. Assume that neuron

2 and neuron 3 are independent given neuron 1. Thus

$$\frac{\rho(v_1,v_2,t)}{\rho(v_1,t)}\frac{\rho(v_1,v_3,t)}{\rho(v_1,t)}=\frac{\rho(v_1,v_2,v_3,t)}{\rho(v_1,t)}$$

We can derive $\rho(v_1,v_2,v_3,t)=\frac{\rho(v_1,v_2,t)\rho(v_1,v_3,t)}{\rho(v_1,t)}$. We can also assume that neuron 1 and neuron 3 are independent given neuron 2, then $\rho(v_1,v_2,v_3,t)=\frac{\rho(v_1,v_2,t)\rho(v_2,v_3,t)}{\rho(v_2,t)}$. To make the probability density for neuron 1 and 2 be symmetric, we approximate the full probability density by

$$\rho(v_1,v_2,v_3,t)=\frac{1}{2}[\frac{\rho(v_1,v_2,t)\rho(v_1,v_3,t)}{\rho(v_1,t)}+\frac{\rho(v_1,v_2,t)\rho(v_2,v_3,t)}{\rho(v_2,t)}]$$

By approximating the full probability density by the terms above, we made the two dimensional kinetic theory equation for $N=3$ closed.

Average firing rate and synchronous firing rate

In the two dimensional population density equation, the reset terms describe the probability density of neurons with voltage reset after firing. The reset terms can be interpreted as

$$\begin{aligned}
&J_{1,reset}(v_2,t)dv_2dt=\Pr\big(\text{neuron 1 fires at } T\in(t,t+dt)\text{and } V_2(T^+)\in(v_2,v_2+dv_2)\big)\\
&J_{2,reset}(v_1,t)dv_1dt=\Pr\big(\text{neuron 2 fires at } T\in(t,t+dt)\text{and } V_1(T^+)\in(v_1,v_1+dv_1)\big)\\
&J_{12,reset}(t)=\Pr\left(\text{both neurons fire simultaneously at } T\in(t,t+dt)\right)
\end{aligned}\tag{5.16}$$

$J_{1,reset}(v_2,t)dv_2$ indicates the probability per unit time that neuron 1 fires while neuron 2 is at v_2, $J_{12,reset}(t)$ describes the probability per unit time that both neurons fire. Thus the average firing rate of neuron 1 is

$$r_{ave}(t)=\int_{v_0}^{v_{th}}J_{1,reset}(v_2,t)dv_2+J_{12,reset}(t)\tag{5.17}$$

The synchronous firing rate $r_{syn}(t)=J_{12,reset}(t)$.

5.3.3 Kinetic theory equation for N all-to-all coupled neurons

In the all-to-all coupled network, once a neuron fires, it will promote the membrane potential of all other neurons that are not firing. If any of these neurons is driven across

threshold, they will instantaneously jump the voltage of the remaining neurons, even higher, possibly causing even more neurons to fire. Therefore, when N gets larger, the "chain reaction" of neurons simultaneously firing will be combinatorically complex.

We initially sketch the form of the equation for N neurons, and then reduce it to the equation for any pair of neurons in the population. To make the resulting equation tractable, we truncate the chain reactions to three or four neurons. In the former case, we will arrive at an equation similar to that form the $N = 3$ case described above.

The population density function for N neurons is N-dimensional. Let V_k represent the membrane potential of the $k-$th neuron at time t.Then we have the population density function $\rho(v_1, \ldots, v_N, t)$.

$$\int_\Omega \rho(v_1, \ldots, v_N, t) = \Pr(((V_1(t), \ldots, V_N(t)) \in \Omega$$

where Ω) is any region of the N dimensional space $v_1 - \cdots - v_N$. Then the evolution of probability density function $\rho(v_1, \ldots, v_N, t)$ is described by an integro-differential equation that is analogous to equation (43) but much more complicated. It has the following form:

$$\begin{aligned}
&\frac{\partial \rho(v_1, \ldots, v_N, t)}{\partial t} = \sum_{i=1}^{N} \frac{1}{\tau} \frac{\partial}{\partial v_i}[(v_i - \varepsilon_r)\rho(v_1, \ldots, v_N, t)] \\
&+ \lambda(t) \sum_{i=1}^{N} \left[\int_{v_0}^{v_i} f_A(v_i - \theta_i)\rho(v_1, \ldots, \theta_i, \ldots, v_N, t) d\theta_i - \rho(v_1, \ldots, v_N, t) \right] \\
&+ \lambda(t) \sum_{i=1}^{N} \delta(v_i - v_0) \int F_A(v_{th} - \theta_i) f_w(v_1 - \theta_1) \ldots f_w(v_N - \theta_N) \\
&\quad \rho(\theta_1, \ldots, \theta_N, t) d\theta_1 \ldots d\theta_N \\
&+ \lambda(t) \sum_{i,j=1}^{N} \delta(v_i - v_0)\delta(v_j - v_0) \int F_A(v_{th} - \theta_i) F_w(v_{th} - \theta_i) \\
&\quad f_{2w}(v_1 - \theta_1) \ldots f_{2w}(v_N - \theta_N)\rho(\theta_1, \ldots, \theta_N) d\theta_1 \ldots d\theta_N \\
&+ \ldots
\end{aligned} \tag{5.18}$$

This equation is high dimensional and the terms grows combinatorically. It is unrealistic to track the full statistics of the spike trains. Thus we study the statistics of spike

trains generated by any pair of the population. We will get a integro-differential equation describing the evolution of probability. By integrating equation(53) over $v_3, \ldots, v_N$, we could obtain a two dimensional probability density function describing the membrane potentials of neuron 1 and neuron 2. However,the resulting equation $\rho(v_1, v_2, t)$ depends on a series of population density functions $\rho(v_1, v_2, t), \ldots, \rho(v_1, \ldots, v_N, t)$. An alternative equation approximating the original one is needed. One possibility is to truncate the equation to a certain level, and make the equation closed. We describe two such truncations below.

The third order truncation of kinetic theory equation: general case

As we know, the exact integro-differential equation for the probability density function $\rho(v_1, v_2, t)$ is very complex. Thus we want to find an approximation to the original equation. If we keep the interactions which involves at most three neurons, we can get a truncation for the kinetic theory equation. The following equation is the third order truncation of probability density function $\rho(v_1, v_2, t)$

$$\begin{aligned}
\frac{\partial \hat{\rho}(v_1, v_2, t)}{\partial t} &= \frac{1}{\tau}\frac{\partial}{\partial v_1}[(v_1 - \varepsilon_r)\hat{\rho}(v_1, v_2, t)] + \frac{1}{\tau}\frac{\partial}{\partial v_2}[(v_2 - \varepsilon_r)\hat{\rho}(v_1, v_2, t)] \\
&+ \lambda(t)[\int_{v_0}^{v_1} f_A(v_1 - \theta_1)\hat{\rho}(\theta_1, v_2, t)d\theta_1 - \hat{\rho}(v_1, v_2, t)] \\
&+ \lambda(t)[\int_{v_0}^{v_2} f_A(v_2 - \theta_2)\hat{\rho}(v_1, \theta_2, t)d\theta_2 - \hat{\rho}(v_1, v_2, t)] \\
&- (N-2)\lambda(t)\int_{v_0}^{v_{th}} F_A(v_{th} - \theta_3)\hat{\rho}(v_1, v_2, \theta_3, t)d\theta_3 \\
&+ (N-2)\lambda(t)\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th} - \theta_3)f_w(v_1 - \theta_1)f_w(v_2 - \theta_2)\hat{\rho}(\theta_1, \theta_2, \theta_3)d\theta_1 d\theta_2 d\theta_3 \\
&+ \delta(v_1 - v_0)J_{1,reset}(v_2, t) + \delta(v_2 - v_0)J_{2,reset}(v_1, t) + \delta(v_1 - v_0)\delta(v_2 - v_0)J_{12,reset}(t)
\end{aligned} \tag{5.19}$$

$J_{1,reset}(v_2, t)$,$J_{12,reset}$ and $r_{ave}(t)$ are given in Appendix B.

The fourth order truncation of kinetic theory equation: general case

The following is the approximation equation truncating to the fourth order interaction.

$$\begin{aligned}
\frac{\partial\hat{\rho}(v_1,v_2,t)}{\partial t} &= \frac{1}{\tau}\frac{\partial}{\partial v_1}[(v_1-\varepsilon_r)\hat{\rho}(v_1,v_2,t)] + \frac{1}{\tau}\frac{\partial}{\partial v_2}[(v_2-\varepsilon_r)\hat{\rho}(v_1,v_2,t)] \\
&+ \lambda(t)[\int_{v_0}^{v_1} f_A(v_1-\theta_1)\hat{\rho}(\theta_1,v_2,t)d\theta_1 - \hat{\rho}(v_1,v_2,t)] \\
&+ \lambda(t)[\int_{v_0}^{v_2} f_A(v_2-\theta_2)\hat{\rho}(v_1,\theta_2,t)d\theta_2 - \hat{\rho}(v_1,v_2,t)] \\
&- (N-2)\lambda(t)\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)\hat{\rho}(v_1,v_2,\theta_3,t)d\theta_3 \\
&+ (N-2)\lambda(t)\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)f_w(v_1-\theta_1)f_w(v_2-\theta_2)\hat{\rho}(\theta_1,\theta_2,\theta_3)d\theta_1 d\theta_2 d\theta_3 \\
&+ (N-2)(N-3)\lambda(t)\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_4) \\
&[f_{2w}(v_1-\theta_1)f_{2w}(v_2-\theta_2) - f_w(v_1-\theta_1)f_w(v_2-\theta_2)]\hat{\rho}(\theta_1,\theta_2,\theta_3,\theta_4,t)d\theta_1 d\theta_2 d\theta_3 d\theta_4 \\
&+ \delta(v_1-v_0)J_{1,reset} + \delta(v_2-v_0)J_{2,reset} \\
&+ \delta(v_1-v_0)\delta(v_2-v_0)J_{12,rest}
\end{aligned} \tag{5.20}$$

We have explained the first three lines in last section. The fourth and fifth line describes the fact that neurons other than 1 and 2 receives an external input and promote all other neurons jump once, then neuron couple (neuron 1 and neuron 2) at (θ_1,θ_2) jumps to (v_1,v_2) while the neuron couple at (v_1,v_2) jumps to higher subthreshold voltage. Since neuron 3,...,neuron N are identical, the coefficient is $N-2$. The sixth and seventh lines reflect the probability that any two neurons other than 1 and 2 fire, causing neuron 1 and 2 at (θ_1,θ_2) jump twice to (v_1,v_2), subtract the probability that first jumps to (v_1,v_2) and then to higher voltages. There are $(N-2)(N-3)$ pairs of neurons other than 1 and 2, so the coefficient of line 6 is $(N-2)(N-3)$.

$J_{1,reset}(v_2,t), J_{2,reset}(v_1,t)$, and $J_{12,reset(t)}$ and r_{ave} are given in the appendix B.

Note that $\hat{\rho}(v_1,v_2,t)$ depends on higher order population density function such as $\hat{\rho}(v_1,v_2,v_3,t)$ or $\hat{\rho}(v_1,v_2,v_3,v_4,t)$. Therefore, we should develop a BBGKY(Bogoliubov-Born-Green-Kirkwood-Yvon) moment hierarchy to approximate the truncated equation and make it closed.

5.4 Results and conclusions

We have derived the third and fourth order truncation for kinetic theory equation of neural networks by stopping at the third or fourth order connectivity. We also approximate the full probability density to make the third order truncation equation closed. The kinetic theory equation should be exact if it is without approximation. Here, we would like to test how well the truncated kinetic theory equations capture the true dynamics of the network.

5.4.1 Testing method

We form the Monte Carlo network of all-to-all coupled neurons. We drive each neuron with Poisson process of rate $\lambda(t)$ and allow for synaptic failure. Let p be the probability of that a spike in one neuron successfully elicits an input event in a given postsynaptic neuron.

In the kinetic theory, we model the jump due to presynaptic input as following the probability distribution function $f_w(x)$. Let $f_w(x) = \delta(x)(1-p) + p\hat{f}_Y(x)$ where $\hat{f}_Y(x)$ is the probability distribution function of jump size given success.

We normalize the voltage range $[v_0, v_{th}] = [0, 1]$, and set the reversal potential $\varepsilon_r = 0.5$. The membrane time constant $\tau = 0.01$s, and the jump size due to external input A follows $f_A = \Gamma(8, 0.1/8)$, where the probability density function of $\Gamma(k, \theta)$ is $f(x; k, \theta) = \frac{1}{\theta^k}\frac{1}{\Gamma(k)}x^{k-1}\exp(-x/\theta)$ for $x \geq 0$ and $k, \theta > 0$. The jump size due to presynaptic input given success also follows $\hat{f}_Y(x) = \Gamma(8, 0.1/8)$. We set the independent Poisson input rate $\lambda(t) = 500$ spikes per second and allow the population density reaches steady state. We keep the input rate until $t = 0.05$ second and we raise the input rate to $\lambda(t) = 1000$ spikes per second. We keep the input rate until $t = 0.2$ second.

For the Monte Carlo simulation, we divide the time interval $[0, 0.2]$s into 200 subintervals with $\bigtriangleup t = 0.001$s. We run 10000 realizations of the all-to-all coupled network and record the firing times of neuron pair 1 and 2. Let vector N_{ave} denote the number of spikes generated by neuron 1 in the 200 time subintervals. Let vector N_{syn} denote the number of spikes generated by both of the two neurons in the time subintervals. Then we form the histogram of the average firing rate $r_{ave} = N_{ave}/(10000 \bigtriangleup t)$, and that of

the synchronous firing rate $r_{syn} = N_{syn}/(10000 \bigtriangleup t)$.

The numerical method solving the third order truncation of kinetic theory equation is described in the Appendix.

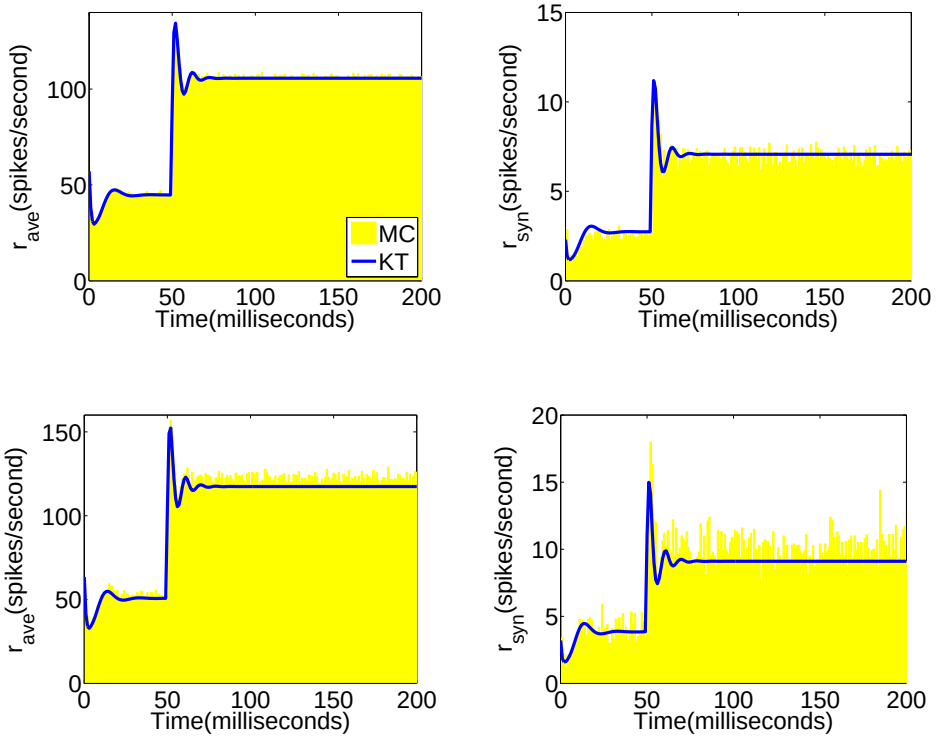


Figure 5.5: Comparison of Monte Carlo simulation and kinetic theory results for $p = 0.3, N = 6$, $p = 0.3, N = 10$. Top: for $p = 0.3, N = 6$, the average firing rate(left), synchronous firing rate(right) are plotted in response to a input jump at $t = 50$ms. Bottom: for $p = 0.3, N = 10$, the average firing rate(left), synchronous firing rate(right) are plotted in response to a input jump at $t = 50$ms.

Figure (5.4.1)(top) shows that in the case of $p = 0.3, N = 6$, the histograms from Monte Carlo coincide with the $r_{ave}(t)$ and $r_{syn}(t)$ obtained from solving the truncated kinetic theory equation. It shows that for a certain range of parameters, the conditional independence approximation did not hurt, and that ignored chain reactions have little effect. Figure (5.4.1)(bottom) shows that in the case of $p = 0.3, N = 10$, the histograms

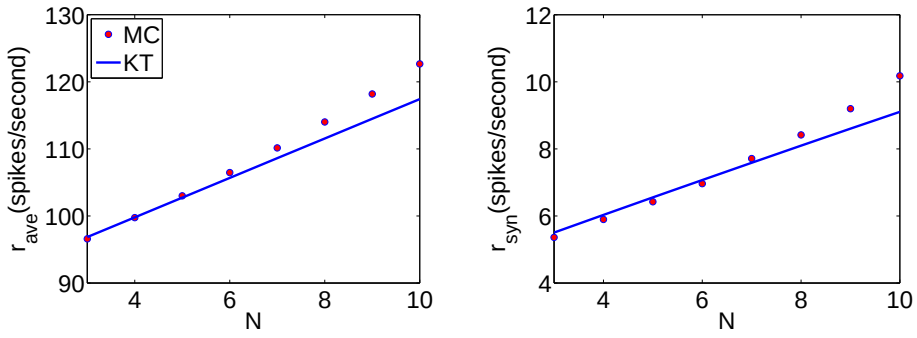


Figure 5.6: The average firing rate and synchronous firing rate in the steady state of the network configuration: $\lambda = 1000$ spikes per second, $p = 0.3, N = 3, 4, 5, 6, 7, 8, 9, 10$. Left: the dots are Monte Carlo simulations for r_{ave} and the plot is the truncated kinetic theory results for r_{ave}. Right: the dots are Monte Carlo simulations for r_{syn} and the plot is the truncated kinetic theory results for r_{syn}.

of the Monte Carlo simulations are higher than the values obtained by the truncated kinetic theory equation. Thus the truncated kinetic theory equation begins to break down. This is because when N gets bigger, it is probably that there are more cascades of many neurons firing together. Thus the third order truncation is not a good approximation to the neural network dynamics. To check this, we can run Monte Carlo simulations and find the histograms of numbers that neurons firing simultaneously.

Figure (5.4.1) shows the average firing rate and synchronous firing rate in the steady state of the networks: $\lambda = 1000$ spikes per second, $p = 0.3$, and $N = 3$ to 10. The dots are Monte Carlo simulation results and the plots are the numerical results of the truncated kinetic theory equation for r_{ave} and r_{syn}. Figure (5.4.1) illustrates that when $p = 0.3$, the third order truncation of kinetic theory equation captures the average firing rate and synchronous firing rate of the network. When N increases, r_{ave} and r_{syn} also increase. The plots start to break down when $N \geq 8$. The third order truncation of kinetic theory equation correctly captures how synchrony emerges as we increase N for small N. But the kinetic theory equation does not capture r_{ave} and r_{syn} as N increases, and it will get even worse if we increase N further.

Next, we discuss how the problem is not necessarily the population size. It can

work well if we decrease coupling to reduce the effect of chain reactions. We do this by decreasing the synaptic success probability p.

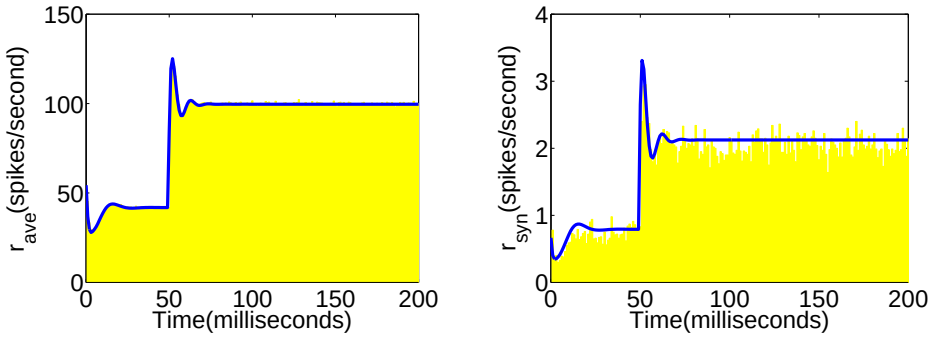


Figure 5.7: Comparison of Monte Carlo simulation and kinetic theory results for $N = 10, p = 0.1$, the average firing rate(left), synchronous firing rate(right) are plotted in response to a input jump at $t = 50$ms.

Figure (5.7) shows that when $N = 10, p = 0.1$, the histograms of the Monte Carlo results coincide with the values of $r_{ave}(t)$ and $r_{syn}(t)$ obtained by solving the truncated kinetic theory equation. This is reasonable since when p is small, the coupling is small, thus the effect of the "chain reaction" is weak and we can approximate the kinetic theory equation by ignoring higher order interactions. Figure (5.7)(below) is the case of $N = 10, p = 0.3$, which breaks down. When p increases, it is more likely that there are many cascades so that we cannot truncate the equation up to the lower order.

Figure (5.8) shows the average firing rate and synchronous firing rate in the steady state of the networks: $\lambda = 1000$ spikes per second, $N = 10$, and $p = 0 : 0.0025 : 0.35$. The dots are Monte Carlo simulations and the plots are numerical results of truncated kinetic theory equation for r_{ave} and r_{syn}. It illustrates that when $p \leq 0.15$, the truncated kinetic theory equation captures the average firing rate and synchronous firing rate of the network. When p increases, r_{ave} and r_{syn} increases too. Figure (5.4.1) also shows that when $p \geq 0.175$, the kinetic theory solution starts to break down.

In an all-to-all coupled network with synaptic failure, once a neuron fires, all the other neurons at subthreshold voltage have the probability of being affected. In a sparsely coupled network, neurons are only affected by their neighbors (Refer to Future work

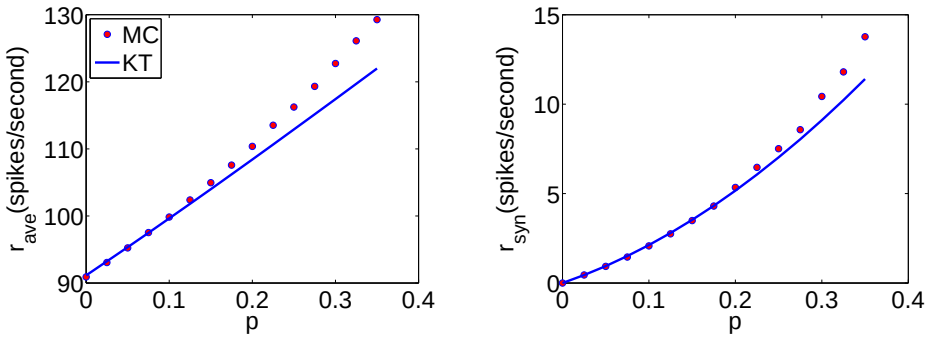


Figure 5.8: The average firing rate and synchronous firing rate in the steady state of the network configuration: $\lambda = 1000$ spikes per second, $N = 10, p = 0 : 0.025 : 0.35$. Left: the dots are Monte Carlo simulations for r_{ave} and the plot is the truncated kinetic theory results for r_{ave}. Right: the dots are Monte Carlo simulations for r_{syn} and the plot is the truncated kinetic theory results for r_{syn}.

for more discussion). Networks in brain are sparsely coupled. It may be that this approach will work to approximate behavior of networks in the brain(i.e., raise questions of importance of chain reactions).

We have derived the third order truncation for the kinetic theory equation of an all-to-all neural network. We numerically solved the equation and compared it with the Monte Carlo simulation. The result shows that the third order truncation for kinetic theory equation captures the average firing rate and synchronous firing rate in a certain parameter regime.

5.4.2 Autocorrelation and cross-correlation with time delay

In our kinetic theory approach, the synchronous firing rate is one kind of second order statistics, but it does not capture all the correlations of the couple of neurons. The outputs are treated as renewal processes, and they are correlated in some time delay.

Let $\rho_{cross}(v_2, \tau; t_0)$ be the density such that $\rho_{cross}(v_2, \tau; t_0)/r_{ave}(t_0)$ is the probability density of $V_2(t+t_0)$ conditioned on neuron 1 firing at t_0. Then $r_{cross}(\tau; t_0)$ can be defined such that $r_{cross}(\tau; t_0)/r_{ave}(t_0)$ is the probability per unit time that neuron 2 fires at $t_0+\tau$ conditioned on neuron 1 firing at t_0. Then the cross-correlation between neuron 1 at t_0

and neuron 2 at $t_0 + \tau$ can be obtained:

$$C(\tau; t_0) = \delta(\tau) r_{syn}(t_0) + r_{cross}(\tau; t_0) - r_{ave}(t_0) r_{ave}(t_0 + \tau)$$

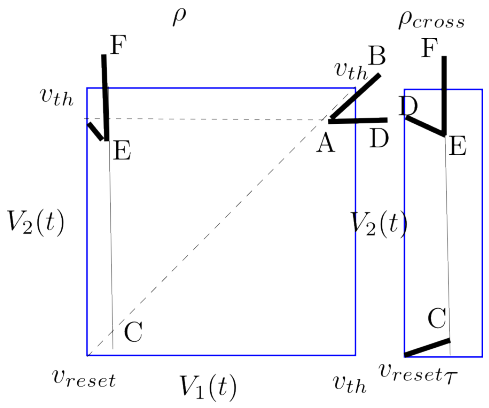


Figure 5.9: Illustration of method to track second order output statistics: for the path ($A \to D \to E \to F$), when neuron 1 fires and neuron 2 is reset to v_2, the firing is added to record $\rho_{cross}(v_2, 0; t_0)$(point D in the right diagram). When neuron 2 evolves from D to E, the time since neuron 1 fires is tracked by τ. Neuron 2 instantaneously cross threshold once it gets to E, which adds a record to $\rho_{cross}(v_{th}, \tau; t_0)$. For the path $A \to B \to C$, the firing is added to record of $\rho_{cross}(v_{th}, 0; t_0)$.

Similarly, we can obtain the autocorrelation of outputs of neuron 1 at delay time τ.

5.4.3 Higher order truncation of kinetic theory equation

In the third order truncation of kinetic theory equation, we made two approximations: truncation up to the third order approximation for the dynamics of network, and conditional independence approximation for the full probability density. We stopped at the third order and did not capture longer chain reactions. We would like to explore the feasibility of higher order truncation of the kinetic theory equation. We derived the fourth order equation (A.66), and use the two approximations described above.

References

[1] Abbott,L.F.,and van Vreeswijk,C.(1993). Asynchronous states in networks of pulse coupled oscillators. *Phys.Rev.E Stat.Phys.Plasmas Fluids Relat.Interdiscip.Topics.*48,1483–1490.

[2] Arenas,A., Díaz-Guilera A., Kurths,J.,Moreno Y.,and Zhou C.(2008).Synchronization in complex networks, *Physics Reports.* 469,93–153.

[3] Axmacher,N.,Mormann, F.,Fernández,G.,Elger,C.E.,and Fell,J.(2006).Memory formation by neuronal synchronization. *Brain Res Rev.* 52,170–182.

[4] Bergman,H.,Feingold,A.,Nini,A.,Raz,A.,Slovin,H.,Abeles,M., and Vaadia,E.(1998). Physiological aspects of information processing in the basal ganglia of normal and parkinsonian primates. *Trends in neurosciences.* 21,32–38.

[5] Bianconi,G.,Gulbahce,N.,Motter, and Adilson E.(2008). Local Structure of Directed Networks. *Phys. Rev. Lett.*100,118701.

[6] Boccaletti,S.,Latora V.,Moreno,Y.,Chavez,M., and Hwang,D.-U.(2006). Complex networks: Structure and dynamics, *Phy. Rep.* 424, 175–308.

[7] Brunel,N.(2000).Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J.Comp.Neurosci.* 8,183–208

[8] Cavazos,J.E.,Golarai,G.,and Sutula,T.P.(1991). Mossy fiber synaptic reorganization induced by kindling:time course of development,progression, and permanence. *J. Neurosci.* 11,2795–2803

[9] Chow,C.C.(1998).Phase-locking in weakly heterogeneous neuronal networks. *Physica A* 118,343–370.

[10] Chung,F., and Lu,L.(2002). Connected components in random graphs with given expected degree sequences. *Ann. Comb.* 6, 125–145.

[11] Denker,M.,Timme,M.,Diesmann,M.,Wolf,E.,and Geisel,T.(2004). Breaking synchrony by heterogeneity in complex networks. *Phys.Rev.Lett.* 92,074103.

[12] Diesmann,M.,Gewaltig,M.O., and Aertsen,A.(1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature*402, 529–533.

[13] Erdős,P., and Rényi,A.(1959). On Random Graphs. I. *Publ.Math.*6, 290–297

[14] Ermentrout,G. B., Kopell,N.(1991). Multiple pulse interactions and averaging in systems of coupled neural oscillators, *J. Math. Biol.*29, 195-217.

[15] Jaynes,E.T.(1978) Where do we stand on maximum entropy inference. In R.D. Levine and M.Tribus, editors, *The Maximum Entropy Formalism.* MIT Press, Cambridge, MA.

[16] Hansel,D.,and Sompolinsky,H.(1992). Synchronization and computation in a chaotic neural network. *Phys.Rev.Lett.*68,718–721.

[17] Hoffman,A.J.,and Wielandt,H.W.(1953). The variation of the spectrum of a normal matrix, *Duke Math. J.*20, 37–39.

[18] Jaynes,E.T.(1957). Information theory and statistical mechanics. *Phys.Rev.*106,62–79

[19] Kitano,K., and Fukai,T.(2007).Variability v.s. synchronicity of neuronal activity in local cortical network models with different wiring topologies. *J.Comp.Neurosci.*23,237–250.

[20] Kramer,M.A.,Chang,F.L., Cohen,M.E., Hudson,D.,and Szeri,A.J.(2007). Synchronization measures of the scalp electroencephalogram can discriminate healthy from Alzheimer's subjects. *Int J Neural Syst*17, 61–69.

[21] Kriener,B.,Helias,M.,Aertsen,A.,and Rotter,S.(2009).Correlations in spiking neuronal networks with distance dependent connections. *J.Comp.Neurosci.*27,177-200.

[22] Kuramoto,Y.(1984). Chemical oscillations, waves, and turbulence. Berlin:Springer-Verlag.

[23] LaMar,M.D., and Smith,G.D.(2010). Effect of node-degree correlation on synchronization of identical pulse-coupled oscillators. *Phys.Rev.E Stat.Nonlin.Soft Matter Phys.*81,046206.

[24] Lefort,S.,Tomm,C.,Sarria,J.-C.F.,and Petersen,C.C.(2009).The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex.*Neuron*61,301-316.

[25] Liu,C., and Nykamp,D.Q.(2009). A kinetic theory approach to capturing interneuronal correlation: The feed-forward case. *J.Comp.Neurosci.*26,339–368.

[26] Macke,J.H.,Opper,M., and Bethge, M. The effect of pairwise neural correlations on global population statistics,*Max Planck Institute for Biological Cybernetics Technical Reports*, 183

[27] Morris,C., and Lecar,H. Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J* 35, 193-213.

[28] Motter,A.E.,Zhou,C.,and Kurths,J.(2005). Network synchronization, diffusion, and the paradox of heterrogeneity. *Phys.Rev.E Stat.Nonlin.Soft Matter Phys.* 71,016116.

[29] Netoff,T.I., Clewley,R.,Arno,S.,Keck,T., and White,J.A.(2004). Epilepsy in small-world networks. *J.Neurosci.*24,8097-8083.

[30] Netoff,T. I. and Schiff,S. J.(2002).Decreased neuronal synchronization during experimental seizures, *J. Neurosci.*22,7297-7307.

[31] Newman,M.E.J.(2003). The Structure and Function of Complex Networks. *SIAM Rev.Soc.Ind.Appl.Math.* 45, 167–256.

[32] Nishikawa,T.,and Motter,A.E.(2010). Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions. *P.N.A.S.* 107, 10342–10347.

[33] Ogata,K.(2010). *Modern Control Engineering.* Boston:Prentice-Hall.

[34] Oswald,A.-M.M.,Doiron,B.,rinzel,J.,and Reyes,A.D.(2009).Spatial profile and differential recruitment of $GABA_{\beta}$ modulate oscillatory activity in auditory cortex. *J.Neurosci.* 29,1031-10334.

[35] Parent,J.M., Yu,T.W., Leibowitz,R.T., Geschwind,D.H., Sloviter,R.S., and Lowenstein,D.H.(1997) Dentate granule cell neurogenesis is increased by seizures and contributes to aberant network reorganization in the adult rat hippocampus. *J.Neurosci.* 17,3727–3738.

[36] Pecora,L.M., Carroll,T.(1998). Master Stability Functions for Synchronized Coupled Systems. *Phys. Rev. Let.* 80, 2109–2112.

[37] Perin,R.,Berger,T.K.,and Markram,H.(2011). A synaptic organizing principle for cortical neuronal groups. *Proc.Natl.Acad.Sci.U.S.A.* 108,5419-5424.

[38] Renart,A., and de la Rocha,J.,Bartho, P.,Hollender, L.,Parga, N.,Reyes, A., and Harris,K. D.(2010). The Asynchronous State in Cortical Circuits. *Science.*327, 587–590.

[39] Restrepo,J.G.,Ott,E.,and Hunt,B.R.(2006). Emergence of Coherence in Complex Networks of Heterogeneous Dynamical Systems. *Phys. Rev. Let.* 96,254103.

[40] Restrepo,J.G., Ott,E., and Hunt,B.R.(2006). Emergence of synchronization in complex networks of interacting dynamical systems. *Physica D* 224, 114-122.

[41] Restrepo,J.G., Ott,E., and Hunt,B.R.(2007). Approximating the largest eigenvalue of network adjacency matrices. *Phys.Rev.E Stat.Nonlin.Soft Matter Phys.* 76, 056119.

[42] Reyes,A.D.(2003). Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. *Nat. Neurosci.* 6, 593-599.

[43] Rosenbaum,R.J., Trousdale,J.,and Josić,K.(2010). Pooling and correlated neural activity. *Front. Comput. Neurosci.* 4:9. doi:10.3389/fncom.2010.00009

[44] Roxin,A.(2011). The role of degree distribution in shaping the dynamics in networks of sparsely connected spiking neurons. *Front. Comput. Neurosci.*5:8. doi:10.3389/fncom.2011.00008

[45] Shadlen,N., and Newsom,W.T.(1998). The variable discharge of cortical neurons: implications for for connectivity, computation and information coding. *J.Neurosci.* 18, 3870-3896.

[46] Song,S. and Sjőstrom,P.J. and Reigl,M. and Nelson,S. and Chklovskii,D.B.(2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3,e68.doi:10.1371/journal.pbio.0030068

[47] Strogatz,S. H. and Mirollo,R. E.(1991). Stability of incoherence in a population of coupled oscillators. *J.Stat.Phys.*63,613-635.

[48] S. Strogatz.(2000). From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D* 143,1-20.

[49] Watts, D. J. and Strogatz, S. H.(1998). Collective dynamics of 'small-world' networks. *Nature* 394, 440–442.

[50] Tetzlaff, T. and Buschermohle, M. and Geisel, T. and Diesmann, M.(2003). The spread of rate and correlation in stationary cortical networks. *Neurocomputing.*52, 949–954

[51] Uhlhaas,P.J., and Singer, W(2010). Abnormal neural oscillations and synchrony in schizophrenia. *Nature Rev Neurosci.*11,100–113.

[52] W. Van Drongelen and H. Koch and C. Marcuccilli and F. Pena and J.-M. Ramirez.(2003). Synchrony levels during evoked seizure-like bursts in mouse neocortical slices. *J Neurophysiol*, 90, 1571–1580

[53] Rossum,v.,M.C.W., Gina G, Turrigiano, Nelson,S.B.(2002). Fast Propagation of Firing Rates through Layered Networks of Noisy Neurons. *J. Neurosci.* 22, 1956–1966

[54] Zeng,L.H., Xu,L., Rensing,N.R., Sinatra,P.M., Rothman,S.M., and Wong, M.(2007). Kainate seizures cause acute dendritic injury and actin depolymerization in vivo. *J. Neurosci.* 27, 11604–11613.

[55] Zhan,C., Chen,G., and Yeung,L.F.(2010). On the distributions of Laplacian eigenvalues versus node degrees in complex networks. *Physica A* 389, 1779-1788.

Appendix A

Glossary and Acronyms

A.1 Dichomotized method generating the second order networks

In this section, we discuss the method for generating second order networks via dichotomizing a multivariate normal distribution. For a SONET with given first order statistic p and second order statistics $\alpha's$, we obtain the corresponding multivariate normal distribution $\mathcal{N}(\Gamma, \Sigma)$. Γ is a $N(N-1)$ dimensional vector whose elements equal γ and Σ is a $N(N-1) \times N(N-1)$ matrix where

$$\begin{aligned}
&\Sigma_{ij,ij} = 1 \\
&\Sigma_{ij,ji} = \rho_{\text{recip}} \\
&\Sigma_{ij,ik} = \rho_{\text{conv}} \\
&\Sigma_{ji,ki} = \rho_{\text{div}} \\
&\Sigma_{ij,ki} = \rho_{\text{chain}} \\
&\Sigma_{ij,\tilde{i}\tilde{j}} = 0 \text{ for the rest pairs of connections}
\end{aligned} \tag{A.1}$$

where γ and ρ_{recip}, ρ_{conv}, ρ_{div} and ρ_{chain} are precomputed by equation(3.17) to match the desired p and corresponding $\alpha's$.

A natural approach for simulating random variables of $N(N-1)$ dimensional multivariate normal distribution $(\tilde{\Gamma}, \tilde{\Sigma})$ is to use the decomposition of $\tilde{\Sigma}$ such that $\tilde{\Sigma} = \tilde{B}\tilde{B}^T$.

Let u be a vector of size $N(N-1)$ that are drawn from a standard normal distribution, $\tilde{\gamma}+\tilde{B}u$ is the desired sample. One special case is the Cholesky decomposition where $\tilde{B}$ is a lower triangle matrix. However, the Cholesky decomposition method is time-cosuming. Our objective of the dichotomized Guassian method is to find an efficient method to obtain the decomposition of the covariance matrix Σ of the Second order networks. Since Σ only consists 0,1, ρ_{recip}, ρ_{conv}, ρ_{div} and ρ_{chain}, and has the structure described by (A.1), we could decompose Σ by

$$\Sigma = B^2 \tag{A.2}$$

where B is a $N(N-1)\times N(N-1)$ matrix, and $B_{ij,ij} = b_{\text{diag}}$, $B_{ij,ji} = b_{\text{recip}}$, $B_{ij,ik} = b_{\text{conv}}$, $B_{ji,ki} = b_{\text{div}}$, $B_{ij,ki} = b_{\text{chain}}$ for any $i \neq j \neq k$ and $B_{ij,\tilde{i}\tilde{j}} = 0$ for the rest possible pairs of conections. It is easy to show that B is symmetric, then $B = B^T$. By equation (A.1) (A.2), we have

$$\begin{aligned}
\Sigma_{ij,ij} &= \sum_{kl} B^2_{ij,kl} \\
&= B^2_{ij,ij} + B^2_{ij,ji} + \sum_{l\neq i,j} B^2_{ij,il} + \sum_{k\neq i,j} B^2_{ij,kj} + \sum_{k\neq i,j} B^2_{ij,ki} + \sum_{l\neq i,j} B^2_{ij,jl} \\
&= b^2_{\text{diag}} + b^2_{\text{recip}} + (N-2)b^2_{\text{conv}} + (N-2)b^2_{\text{div}} + 2(N-2)b^2_{\text{chain}}
\end{aligned} \tag{A.3}$$

$$\begin{aligned}
\Sigma_{ij,ji} &= \sum_{kl} B_{ij,kl}B_{ji,kl} \\
&= B_{ij,ij}B_{ji,ij} + B_{ij,ji}B_{ji,ji} + \sum_{k\neq i,j} B_{ij,ki}B_{ji,ki} + \sum_{k\neq i,j} B_{ij,kj}B_{ji,kj} \\
&+ \sum_{l\neq i,j} B_{ij,il}B_{ji,il} + \sum_{l\neq i,j} B_{ij,jl}B_{ji,jl} \\
&= 2b_{\text{diag}}b_{\text{recip}} + 2(N-2)b_{\text{chain}}(b_{\text{conv}} + b_{\text{div}})
\end{aligned} \tag{A.4}$$

$$\begin{aligned}\Sigma_{ij,ik} &= \sum_{mn} B_{ij,mn}B_{ik,mn} \\ &= B_{ij,ij}B_{ik,ij} + B_{ij,ji}B_{ik,ji} + B_{ij,ik}B_{ik,ik} + B_{ij,ki}B_{ik,ki} \\ &+ B_{ij,jk}B_{ik,jk} + B_{ij,kj}B_{ik,kj} + \sum_{m\neq i,j,k} B_{ij,mi}B_{ik,mi} + \sum_{n\neq i,j,k} B_{ij,in}B_{ik,in} \\ &= 2b_{\text{diag}}b_{\text{conv}} + 2b_{\text{recip}}b_{\text{chain}} + 2b_{\text{div}}b_{\text{chain}} + (N-3)b_{\text{div}}^2 + (N-3)b_{\text{chain}}^2 \end{aligned} \tag{A.5}$$

$$\begin{aligned}\Sigma_{ji,ki} &= \sum_{mn} B_{ji,mn}B_{ki,mn} \\ &= B_{ji,ij}B_{ki,ij} + B_{ji,ji}B_{ki,ji} + B_{ji,ik}B_{ki,ik} + B_{ji,ki}B_{ki,ki} + B_{ji,jk}B_{ki,jk} + B_{ji,kj}B_{ki,kj} \\ &+ \sum_{m\neq i,j,k} B_{ji,mi}B_{ki,mi} + \sum_{n\neq i,j,k} B_{ji,in}B_{ki,in} \\ &= 2b_{\text{diag}}b_{\text{div}} + 2b_{\text{recip}}b_{\text{chain}} + 2b_{\text{conv}}b_{\text{chain}} + (N-3)b_{\text{conv}}^2 + (N-3)b_{\text{chain}}^2 \end{aligned} \tag{A.6}$$

$$\begin{aligned}\Sigma_{ij,jk} &= \sum_{mn} B_{ij,mn}B_{jk,mn} \\ &= B_{ij,ij}B_{jk,ij} + B_{ij,ji}B_{jk,ji} + B_{ij,ik}B_{jk,ik} + B_{ij,ki}B_{ij,ki} + B_{ij,jk}B_{jk,jk} + B_{ij,kj}B_{jk,kj} \\ &+ \sum_{m\neq i,j,k} B_{ij,mj}B_{jk,mj} + \sum_{n\neq i,j,k} B_{ij,jn}B_{jk,jn} \\ &= 2b_{\text{diag}}b_{\text{chain}} + b_{\text{recip}}b_{\text{conv}} + b_{\text{recip}}b_{\text{div}} + b_{\text{conv}}b_{\text{div}} + b_{\text{chain}}^2 \\ &+ (N-3)b_{\text{conv}}b_{\text{chain}} + (N-3)b_{\text{div}}b_{\text{chain}} \end{aligned} \tag{A.7}$$

We should find solutions to the system (A.3), (A.4), (A.5), (A.6), (A.7). There are many solutions to the system. Let b_{conv}, b_{div} and b_{chain} have the order $O(\frac{1}{\sqrt{N}})$, we could approximate equations (A.5), (A.6) and (A.7) by

$$\frac{\rho_{\text{conv}}}{N} = b_{\text{conv}}^2 + b_{\text{chain}}^2 \tag{A.8}$$

$$\frac{\rho_{\text{div}}}{N} = b_{\text{div}}^2 + b_{\text{chain}}^2 \tag{A.9}$$

$$\frac{\rho_{\text{chain}}}{N} = b_{\text{conv}}b_{\text{chain}} + b_{\text{div}}b_{\text{chain}} \tag{A.10}$$

for large N.

Let $C = \frac{1}{N}\begin{pmatrix} \rho_{\text{conv}} & \rho_{\text{chain}} \\ \rho_{\text{chain}} & \rho_{\text{div}} \end{pmatrix}$ and $D = \begin{pmatrix} b_{\text{conv}} & b_{\text{chain}} \\ b_{\text{chain}} & b_{\text{div}} \end{pmatrix}$. By equations (A.8), (A.9) and (A.10), we have

$$C = D^2 \tag{A.11}$$

Conduct singular value decomposition of the symmetric matrix C, we obtain $C = U\Lambda U^T$. $\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ where λ_1 and λ_2 are eigenvalues of C and the columns of U are the corresponding eigenvectors. Then one of the solutions satisfying equation(A.11) is

$$U\Lambda^{\frac{1}{2}}U^T \tag{A.12}$$

Since C is symmetric, the eigenvectors should be real. In this case, the eigenvalues of C should be nonnegative, which requires

$$\rho_{\text{chain}}^2 < \rho_{\text{conv}}\rho_{\text{div}} \tag{A.13}$$

Pluging (A.12) into (A.3) and (A.4), we can solve for b_{diag} and b_{recip} as long as

$$1 - (N-2)(b_{\text{conv}}^2 + b_{\text{div}}^2 + 2b_{\text{chain}}^2) \geq \rho_{\text{recip}} - (N-2)(2b_{\text{conv}}b_{\text{chain}} + 2b_{\text{div}}b_{\text{chain}}) \tag{A.14}$$

For large N, equations(A.3)(A.4) reduce to

$$1 = b_{\text{diag}}^2 + b_{\text{recip}}^2 + \rho_{\text{conv}}^2 + \rho_{\text{div}}^2 \tag{A.15}$$

$$\rho_{\text{recip}} = 2b_{\text{diag}}b_{\text{recip}} + 2\rho_{\text{chain}} \tag{A.16}$$

The solution b_{diag} and b_{recip} exists if

$$\rho_{\text{recip}} + \rho_{\text{conv}} + \rho_{\text{div}} - 2\rho_{\text{chain}} \leq 1 \tag{A.17}$$

Thus the condition of covariance matrix for the underlying multivariate normal distribution is

$$\begin{cases} \rho_{\text{conv}} \geq 0 \\ \rho_{\text{div}} \geq 0 \\ \rho_{\text{conv}}\rho_{\text{div}} - \rho_{\text{chain}}^2 \geq 0 \\ \rho_{\text{recip}} + \rho_{\text{conv}} + \rho_{\text{div}} - 2\rho_{\text{chain}} \leq 1 \end{cases} \tag{A.18}$$

For the second order networks, we obtain a sample X of $N(N-1)$ dimensional correlated Gaussian by $X = BU + \Gamma$, where U is a $N(N-1)$ vector drawn from standard normal distribution. It is obvious that X is draw from a multivariate normal distribution with mean γ and covariance matrix B^2. The next step is to threshold X to a vector by mapping all the nonnegative entries of X to be 1 and the rest to be 0. Then rearrange the resulting vector so that it forms a adjacency matrix of a N node network.

To sum up, the procedure for simulating a second order network of size N with given first order statistic p and second order statistics α_{recip}, α_{conv}, α_{div} and α_{chain} by dichotomized Gausian is

- Calculate the mean γ and covariance matrix Σ of the underlying multivariate normal distribution
- Compute a symmetric matrix B such that $\Sigma = B^2$
- Obtain X by $X = BU + \Gamma$ is the desired sample, where U is a $N(N-1)$ vector drawn from standard normal distribution
- Map entries of X to 1 if it is nonnegative and 0 otherwise
- Rearrange the vector so that it forms the adjacency matrix of a N-node network

A.2 Derivation of joint degree distribution

In this section, we show the formula for the joint degree distribution of second order networks, which is derived by dichotomizing the underlying multivariate normal distribution such that the first order and second order statistics match those of the correlated Bernoulli distribution.

Given two sets of Bernoulli random variables $(A_1, \ldots, A_N)$ and $(B_1, \ldots, B_N)$ as defined in section(3.1.3), $E(A_i) = p$, $E(B_i) = p$ for $i = 1 \ldots N$, $\text{corr}(A_i, A_j) = \alpha_1$, $\text{corr}(B_i, B_j) = \alpha_2$, and $\text{corr}(A_i, B_j) = \alpha_3$ where $i, j \in \{1 \ldots N\}$. By solving the system of equations (3.17), we obtain the corresponding mean of the latent normal distribution γ and correlation ρ_1, ρ_2 and ρ_3. The vector of correlated Bernoulli distribution of $(A_1, \ldots, A_N, B_1, \ldots, B_N)$ is modeled by the latent multivariate normal distribution

$(X_1, \ldots, X_N, Y_1, \ldots, Y_N)$. Then the multivariate normal distribution has mean γ and covariance matrix Σ where $\Sigma_{ii} = 1$ for $1 \leq i \leq N$. $\mathrm{cov}X_i, X_j = \rho_1, \mathrm{cov}Yi, Y_j = \rho_2$ for $i \neq j$. $\mathrm{cov}X_i, Y_j = \rho_3$. Then (X, Y) can be described by the equations (A.19).

$$\begin{aligned} X_i &= \gamma + a_1 U_1 + b_1 U_2 + \sqrt{1-\rho_1} V_i \\ Y_i &= \gamma + a_2 U_1 + b_2 U_2 + \sqrt{1-\rho_2} W_i \\ U_1 &\sim \aleph(0,1) \\ U_2 &\sim \aleph(0,1) \\ V_i &\sim \aleph(0,1) \\ W_i &\sim \aleph(0,1) \end{aligned} \tag{A.19}$$

where

$$\begin{aligned} a_1^2 + b_1^2 &= \rho_1 \\ a_2^2 + b_2^2 &= \rho_2 \\ a_1 b_1 + a_2 b_2 &= \rho_3 \end{aligned} \tag{A.20}$$

which can be written as

$$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_3 \\ \rho_3 & \rho_2 \end{pmatrix} \tag{A.21}$$

We set $A_i = 1$ for $X_i \geq 0$ and $A_i = 0$ otherwise. Similarly, $B_i = 1$ for $Y_i \geq 0$ and $B_i = 0$ if $Y_i < 0$. Let $(K_1, K_2) = (\sum_{i=1}^N A_i, \sum_{j=1}^N B_j)$ and $(R_1, R_2) = (\frac{K_1}{N}, \frac{K_2}{N})$ be the fraction of active Bernoulli random variables in their corresponding population. $p(r_1, r_2)$

is defined as the probability density function for (R_1, R_2). For fixed r_1 and r_2, we have

$$
\begin{aligned}
p(r_1, r_2) &= \Pr(\sum_{i=1}^{N} A_i = Nr_1, \sum_{i=1}^{N} B_i = Nr_2) \\
&= \binom{N}{Nr_1}\binom{N}{Nr_2}\Pr(A_1 = \cdots = A_{Nr_1} = 1, A_{Nr_1+1} = \cdots = A_N = 0, \\
&B_1 = \cdots = B_{Nr_2} = 1, B_{Nr_2+1} = \cdots = B_N = 0) \\
&= \binom{N}{Nr_1}\binom{N}{Nr_2}\Pr(X_1 \geq 0, \ldots, X_{Nr_1} \geq 0, X_{Nr_1+1} < 0, \ldots, X_N < 0, \\
&Y_1 \geq 0, \ldots, Y_{Nr_2} \geq 0, Y_{Nr_2+1} < 0, \ldots, Y_N < 0) \\
&= \binom{N}{Nr_1}\binom{N}{Nr_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(s)\phi(t)\Pr(X_1 \geq 0, \ldots, X_{Nr_1} \geq 0, X_{Nr_1+1} < 0, \ldots, X_N < 0, \\
&Y_1 \geq 0, \ldots, Y_{Nr_2} \geq 0, Y_{Nr_2+1} < 0, \ldots, Y_N < 0|s,t)dsdt \\
&= \binom{N}{Nr_1}\binom{N}{Nr_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(s)\phi(t)\Pr(X_1 \geq 0|s,t)\ldots\Pr(X_{Nr_1} \geq 0|s,t)\Pr(X_{Nr_1+1} < 0|s,t)\ldots \\
&\Pr(X_n < 0|s,t)\Pr(Y_1 \geq 0|s,t)\ldots\Pr(Y_{Nr_2} \geq 0|s,t)\Pr(Y_{Nr_2+1} < 0|s,t)\ldots\Pr(Y_N < 0|s,t)dsdt \\
&= \binom{N}{Nr_1}\binom{N}{Nr_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(s)\phi(t)f^{Nr_1}(s,t)(1-f(s,t))^{N(1-r_1)} \\
&g^{Nr_2}(s,t)(1-g(s,t))^{N(1-r_2)}dsdt
\end{aligned} \tag{A.22}
$$

where

$$
\begin{aligned}
f(s,t) &= \Pr\Big(X_i \geq 0|s,t) = \Pr(\gamma + a_1 s + b_1 t + \sqrt{1-\rho_1}V_i \geq 0|s,t\Big) \\
&= \Pr\left(V_i \geq -\frac{\gamma + a_1 s + b_1 t}{\sqrt{1-\rho_1}}\right) \\
&= \Phi\left(\frac{\gamma + a_1 s + b_1 t}{\sqrt{1-\rho_1}}\right)
\end{aligned} \tag{A.23}
$$

and

$$
\begin{aligned}
g(s,t) &= \Pr\Big(Y_i \geq 0|s,t) = \Pr(\gamma + a_2 s + b_2 t + \sqrt{1-\rho_2}W_i \geq 0|s,t\Big) \\
&= \Phi\left(\frac{\gamma + a_2 s + b_2 t}{\sqrt{1-\rho_2}}\right)
\end{aligned} \tag{A.24}
$$

Then equation(A.22) becomes

$$
\Pr(r_1, r_2) = \binom{N}{Nr_1}\binom{N}{Nr_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(s)\phi(t)\exp\big(Nh(s,t)\big)dsdt \tag{A.25}
$$

where

$$h(s,t) = r_1 \log(f(s,t)) + (1-r_1)\log(1-f(s,t)) \\ + r_2 \log(g(s,t)) + (1-r_2)\log(1-g(s,t)) \tag{A.26}$$

By the saddle point approximation, the equation (A.25) becomes

$$\Pr(r_1, r_2) \approx \binom{N}{Nr_1}\binom{N}{Nr_2}\phi(s_0)\phi(t_0)\exp(Nh(s_0,t_0))\frac{2\pi}{N\sqrt{|\det H(s_0,t_0)|}} \tag{A.27}$$

where (s_0, t_0) is the global maximum of h and $H(s_0, t_0)$ is the Hessian of h at the maximum.

$$\frac{\partial h}{\partial s} = \left(\frac{r_1}{f} - \frac{1-r_1}{1-f}\right)\frac{a_1}{\sqrt{1-\rho_1}}\phi\left(\frac{\gamma + a_1 s + b_1 t}{\sqrt{1-\rho_1}}\right) + \left(\frac{r_2}{g} - \frac{1-r_2}{1-g}\right)\frac{a_2}{\sqrt{1-\rho_2}}\phi\left(\frac{\gamma + a_2 s + b_2 t}{\sqrt{1-\rho_2}}\right)$$
$$\frac{\partial h}{\partial t} = \left(\frac{r_1}{f} - \frac{1-r_1}{1-f}\right)\frac{b_1}{\sqrt{1-\rho_1}}\phi\left(\frac{\gamma + a_1 s + b_1 t}{\sqrt{1-\rho_1}}\right) + \left(\frac{r_2}{g} - \frac{1-r_2}{1-g}\right)\frac{b_2}{\sqrt{1-\rho_2}}\phi\left(\frac{\gamma + a_2 s + b_2 t}{\sqrt{1-\rho_2}}\right) \tag{A.28}$$

Thus we have

$$f(s_0, t_0) = r_1 \\ g(s_0, t_0) = r_2 \tag{A.29}$$

Let $\eta_1 = \Phi^{-1}(r_1), \eta_2 = \Phi^{-1}(r_2)$, by equation(A.23),(A.24), we have

$$\eta_1 = \frac{\gamma + a_1 s_0 + b_1 t_0}{\sqrt{1-\rho_1}} \\ \eta_2 = \frac{\gamma + a_2 s_0 + b_2 t_0}{\sqrt{1-\rho_2}} \tag{A.30}$$

Let $\eta = (\eta_1, \eta_2)^T$,$D = \begin{pmatrix} \sqrt{1-\rho_1} & 0 \\ 0 & \sqrt{1-\rho_2} \end{pmatrix}$ and $C = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$, we have

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = D^{-1}\begin{pmatrix} \gamma \\ \gamma \end{pmatrix} + D^{-1}C\begin{pmatrix} s_0 \\ t_0 \end{pmatrix} \tag{A.31}$$

and we can obtain

$$\begin{pmatrix} s_0 \\ t_0 \end{pmatrix} = C^{-1}\left(D\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} - \begin{pmatrix} \gamma \\ \gamma \end{pmatrix}\right) \tag{A.32}$$

By pluging (s_0, t_0) into equation(A.26), we obtain $h(s_0, t_0) = 0$. We also would like to find the Hessian of h at (s_0, t_0).

$$\begin{aligned}
\frac{\partial^2 h}{\partial s^2}(s_0, t_0) &= -\frac{1}{r_1(1-r_1)}\frac{a_1^2}{1-\rho_1}\phi^2(\eta_1) - \frac{1}{r_2(1-r_2)}\frac{a_2^2}{1-\rho_2}\phi^2(\eta_2) \\
\frac{\partial^2 h}{\partial s \partial t}(s_0, t_0) &= -\frac{1}{r_1(1-r_1)}\frac{a_1 b_1}{1-\rho_1}\phi^2(\eta_1) - \frac{1}{r_2(1-r_2)}\frac{a_2 b_2}{1-\rho_2}\phi^2(\eta_2) \\
\frac{\partial^2 h}{\partial t^2}(s_0, t_0) &= -\frac{1}{r_1(1-r_1)}\frac{b_1^2}{1-\rho_1}\phi^2(\eta_1) - \frac{1}{r_2(1-r_2)}\frac{b_2^2}{1-\rho_2}\phi^2(\eta_2)
\end{aligned} \tag{A.33}$$

and

$$H(s_0, t_0) = \frac{\big((a_1 b_2 - a_2 b_1)\phi(\eta_1)\phi(\eta_2)\big)^2}{r_1 r_2 (1-r_1)(1-r_2)(1-\rho_1)(1-\rho_2)} \tag{A.34}$$

Plugging the equation(A.21)(A.34) into (A.27), we obtain

$$\Pr(r_1, r_2) \approx \sqrt{\frac{(1-\rho_1)(1-\rho_2)}{\rho_1\rho_2 - \rho_3^2}}\frac{\phi(s_0)\phi(t_0)}{\phi(\eta_1)\phi(\eta_2)} \tag{A.35}$$

Then the joint probability distribution of (r_1, r_2) is

$$\Pr(r_1, r_2) = \sqrt{\frac{(1-\rho_1)(1-\rho_2)}{\rho_1\rho_2 - \rho_3^2}}\exp\big(-\frac{1}{2}\big((D\eta-\gamma)^T(CC^T)^{-1}(D\eta-\gamma) - \eta^T\eta\big)\big) \tag{A.36}$$

where $CC^T = \begin{pmatrix} \rho_1 & \rho_3 \\ \rho_3 & \rho_2 \end{pmatrix}$.

A.3 Clustering coefficients and transitivity

In this section, we would like to derive an approximation of the clustering coefficients of SONETs with given statistics p and α's. There are four types of clustering coefficients for directed graph. Let C_l^i denote the i-th local clustering coefficients:

$$C_1^i = \frac{\text{number of Type I Triangle with one vertex at } i}{\text{number of chain motif with center at } i} \tag{A.37}$$

$$C_2^i = \frac{\text{number of Type II Triangle with } i \text{ as the center of convergent motif}}{\text{number of convergent motif with center at } i} \tag{A.38}$$

$$C_3^i = \frac{\text{number of Type II Triangle with } i \text{ as the center of divergent motif}}{\text{number of divergent motif with center at } i} \tag{A.39}$$

$$C_4^i = \frac{\text{number of Type II Triangle with } i \text{ as the center of chain motif}}{\text{number of chain motif with center at } i} \tag{A.40}$$

The approximation of local clustering coefficients could be obtained by dichotomizing Gaussian method. The clustering coefficients of the network is defined as the average of local clustering coefficients of all nodes in the network:

$$C_l = \frac{1}{N}\sum_{i=1}^{N} C_l^i \text{ for} l = 1 \dots 4 \tag{A.41}$$

Let's take an example of C_1^i, which is the proportion of type I triangles divided by the number of chain motifs centered at node i. From p, we can get γ. From $\alpha' s$, we could get the corresponding ρ_{conv}, ρ_{div} and ρ_{chain}. In figure(3.7) **A**, the tree edges W_{ji},W_{ik} and W_{kj} are based on a underlying multivariate normal distribution with mean $\Gamma_1 = \begin{pmatrix} \gamma \\ \gamma \\ \gamma \end{pmatrix}$, and covariance matrix $\Sigma_1 = \begin{pmatrix} 1 & \rho_{\text{chain}} & \rho_{\text{chain}} \\ \rho_{\text{chain}} & 1 & \rho_{\text{chain}} \\ \rho_{\text{chain}} & \rho_{\text{chain}} & 1 \end{pmatrix}$. Let's denote the random variables with the multivariate normal distribution as X_{ji},X_{ik} and X_{kj}. Then we have

$$\begin{aligned} &\Pr(W_{ji} = W_{ik} = W_{kj} = 1) = \Pr(X_{ji} \geq 0, X_{ik} \geq 0, X_{jk} \geq 0) \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma_1|^{\frac{1}{2}}} \\ &\exp\left(-\frac{1}{2}(x_{ji}-\gamma, x_{ik}-\gamma, x_{jk}-\gamma)\Sigma_1^{-1}(x_{ji}-\gamma, x_{ik}-\gamma, x_{jk}-\gamma)^T\right) dx_{ji}dx_{ik}dx_{jk} \end{aligned} \tag{A.42}$$

$\Pr(W_{kj} = 1 | W_{ji} = W_{ik}) = \frac{\Pr(W_{ji}=W_{ik}=W_{kj}=1)}{\Pr(W_{ji}=1, W_{ik}=1)}$. By equation(A.42)(2.12), we could

approximate the mean of C_1^i by

$$\frac{1}{p^2(1+\alpha_{\text{chain}})}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma_1|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x_{ji}-\gamma, x_{ik}-\gamma, x_{kj}-\gamma)\Sigma_1^{-1}(x_{ji}-\gamma, x_{ik}-\gamma, x_{kj}-\gamma)^T\right) dx_{ji}dx_{ik}dx_{kj} \tag{A.43}$$

We can obtain the approximation of other three clustering coefficients:

$$\bar{C}_2 \approx \frac{1}{p^2(1+\alpha_{\text{conv}})}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma_2|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x_{ij}-\gamma, x_{ik}-\gamma, x_{kj}-\gamma)\Sigma_2^{-1}(x_{ij}-\gamma, x_{ik}-\gamma, x_{kj}-\gamma)^T\right) dx_{ij}dx_{ik}dx_{kj} \tag{A.44}$$

$$\bar{C}_3 \approx \frac{1}{p^2(1+\alpha_{\text{div}})}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma_2|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x_{ji}-\gamma, x_{ki}-\gamma, x_{kj}-\gamma)\Sigma_2^{-1}(x_{ji}-\gamma, x_{ki}-\gamma, x_{kj}-\gamma)^T\right) dx_{ji}dx_{ki}dx_{kj} \tag{A.45}$$

$$\bar{C}_4 \approx \frac{1}{p^2(1+\alpha_{\text{chain}})}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma_2|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x_{ji}-\gamma, x_{ik}-\gamma, x_{jk}-\gamma)\Sigma_2^{-1}(x_{ji}-\gamma, x_{ik}-\gamma, x_{jk}-\gamma)^T\right) dx_{ji}dx_{ik}dx_{jk} \tag{A.46}$$

where $\Sigma_2 = \begin{pmatrix} 1 & \rho_{\text{chain}} & \rho_{\text{conv}} \\ \rho_{\text{chain}} & 1 & \rho_{\text{div}} \\ \rho_{\text{conv}} & \rho_{\text{div}} & 1 \end{pmatrix}$

Estimate clustering coefficients From a given network connectivity matrix W, we could estimate the four clustering coefficients directly.

$$\hat{C}_1^i = \frac{\sum_{j,k\neq i} W_{ji}W_{ik}W_{kj}}{d^i_{\text{in}}d^i_{\text{out}}} = \frac{W^3_{ii}}{d^i_{\text{in}}d^i_{\text{out}}} \tag{A.47}$$

$$\hat{C}_2^i = \frac{\sum_{j,k\neq i} W_{ij}W_{ik}W_{kj}}{d_{\text{in}}(d_{\text{in}}-1)} = \frac{(W^2W^T)_{ii}}{d^i_{\text{in}}d^i_{\text{out}}} \tag{A.48}$$

$$\hat{C}_3^i = \frac{\sum_{j,k\neq i} W_{ji}W_{ki}W_{kj}}{d_{\text{out}}(d_{\text{out}}-1)} = \frac{(W^TW^2)_{ii}}{d_{\text{out}}(d_{\text{out}}-1)} \tag{A.49}$$

$$\hat{C}_4^i = \frac{\sum_{j,k\neq i} W_{ji}W_{ik}W_{jk}}{d_{\text{in}}d_{\text{out}}} = \frac{(WW^TW)_{ii}}{d_{\text{in}}d_{\text{out}}} \tag{A.50}$$

By taking the average of local clustering coefficients, we obtain the measured clustering coefficients.

$$\hat{C}_l = \frac{\sum_{i=1}^N \hat{C}_l^i}{N} \text{ for } l = 1\dots 4 \tag{A.51}$$

A.4 Generate samples from given pattern of probability distribution and statistics

In this section, we would like to introduce how we simulate networks with desired marginal degree distribution pattern and statistics. Assume the network has N vertices, the first order statistic is p and second order statistics α_{conv}, α_{div} and α_{chain} are given. Let p_{in} denote the probability density function for the in-coming degree. Thus p_{in} should satisfy the equations

$$\int_0^N p_{\text{in}}(x)dx = 1 \tag{A.52}$$

$$\int_0^N xp_{\text{in}}(x)dx = (N-1)p \tag{A.53}$$

$$\int_0^N x(x-1)p_{\text{in}}(x)dx = (N-1)(N-2)p^2(1+\alpha_{\text{conv}}) \tag{A.54}$$

Now let's look at the truncated scale-free network for given N, first order statistic p, second order statistics α_{conv} and α_{div}. The probability density function for the in-coming degree distribution has the form

$$p_{\text{in}}(x) = \begin{cases} c_{\text{in}}x^{-\beta_{\text{in}}} & x \in [a_{\text{in}}, b_{\text{in}}] \\ 0 & \text{otherwise} \end{cases} \tag{A.55}$$

where β_{in} is the exponent of the power law distribution, a_{in} and b_{in} are the leftcut and rightcut, respectively. For a network of N nodes, $0 \leq a_{\text{in}} < b_{\text{in}} \leq N$. c_{in} is a constant that normalize the distribution. For given N, p, α_{conv} and exponent β_{in}, a_{in}, b_{in} and c_{in} could be obtained by plugging the probability density function of truncated power law (A.55) into equations(A.52),(A.53) and (A.54).

$$c\frac{b_{\text{in}}^{1-\beta_{\text{in}}} - a_{\text{in}}^{1-\beta_{\text{in}}}}{1-\beta} = 1 \tag{A.56}$$

$$c\frac{b_{\text{in}}^{2-\beta_{\text{in}}} - a_{\text{in}}^{2-\beta_{\text{in}}}}{2-\beta} = (N-1)p \tag{A.57}$$

$$c\frac{b_{\text{in}}^{3-\beta_{\text{in}}} - a_{\text{in}}^{3-\beta_{\text{in}}}}{3-\beta} = (N-1)(N-2)p^2(1+\alpha_{\text{conv}}) + (N-1)p \tag{A.58}$$

Networks with exponential in-coming and out-going degree distribution could also be simulated. The truncated exponential distribution can be defined by

$$p_{\text{in}}(x) = \begin{cases} c_{\text{in}} \exp(-\beta_{\text{in}}(x - a_{\text{in}})) & x \in [a_{\text{in}}, b_{\text{in}}] \\ 0 & \text{otherwise} \end{cases} \tag{A.59}$$

where b_{in} is the exponent, a_{in}, b_{in} are the left cutoff and right cutoff of the degree distribution. c_{in} is a constant that normalize the probability density function. By plugging the probability density function (A.59) into equations(A.52)(A.53)(A.54), we can solve a_{in}, b_{in} and c_{in}.

Figure(A.1) illustrates a probability density function for truncated power law distribution and exponential distribution.

Given α_{div} and exponent β_{out}, we could also find the out-going degree distribution of truncated power law or exponential distribution. Once we have the marginal in-coming and out-going degree distribution, we can draw N samples from the two probability density functions, respectively. Let $(d_{\text{in}}^1, d_{\text{in}}^2, \ldots, d_{\text{in}}^N)$ denote samples drawn from $p_{\text{in}}(x)$, which gives an in-coming degree sequence. Similarly, the samples $(d_{\text{out}}^1, d_{\text{out}}^2, \ldots, d_{\text{out}}^N)$ are the samples draw from $p_{\text{out}}(x)$ and denote the out-going degree sequence. The in-coming degree sequence and out-going degree sequence are independent.

To create dependence of the in-coming degree and out-going degree, a Gaussian copula is applied. For $i = 1 \ldots N$, draw samples (s_i, t_i) from multivariate gaussian distribution with mean 0 and covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ where ρ is any chosen correlation

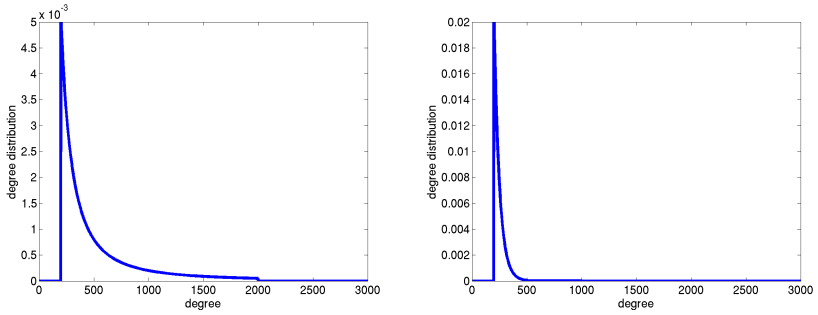


Figure A.1: Probability density function. Left: scale free with $N = 3000, \beta = 2$, $a = 200$ and $b = 2000$ Right: network with exponential distribution that has $N = 3000$, $\beta = 0.02$, $a = 200$ and $b = 1000$

coefficient of the multivariate normal distribution. By solving the equations (A.60) and (A.61):

$$\int_0^{\tilde{d}^i_{\text{in}}} p_{\text{in}}(x)dx = \Phi(s_i) \tag{A.60}$$

$$\int_0^{\tilde{d}^i_{\text{out}}} p_{\text{out}}(x)dx = \Phi(t_i) \tag{A.61}$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution, we could find the correlated samples $(\tilde{d}^i_{\text{in}}, \tilde{d}^i_{\text{out}})$ for $i = 1 \dots N$.

Then the weight of connectivity from j to i can be described as $w_{ij} = c\frac{\tilde{d}^i_{\text{in}}\tilde{d}^j_{\text{out}}}{\sum_{ij}\tilde{d}^i_{\text{in}}\tilde{d}^j_{\text{out}}}$. The adjacency matrix has $W_{ij} = 1$ with probability w_{ij}. c is chosen so that the expectation of in-coming and out-going degree is p. Note that it should satisfy the condition: $c\frac{\tilde{d}^i_{\text{in}}\tilde{d}^j_{\text{out}}}{\sum_{ij}\tilde{d}^i_{\text{in}}\tilde{d}^j_{\text{out}}} \leq 1$ for$i = 1 \dots N$.

Thus we generated scale-free networks for given $p, \alpha_{\text{conv}}, \alpha_{\text{chain}}$ and exponent of the power law distribution. We could create dependence by the Gaussian copula. Once the network is generated, its $\widehat{\alpha}_{\text{chain}}$ can be obtained from the adjacency matrix.

A.5 The kinetic theory equation for N=3

$$
\begin{aligned}
J_{1,reset}(v_2,t) = \lambda(t)[&\int_{v_0}^{v_{th}}\int_{v_0}^{v_2} F_A(v_{th}-\theta_1) f_w(v_2-\theta_2)\rho(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+ \int_{v_0}^{v_{th}}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)(f_{2w}(v_2-\theta_2)-f_w(v_2-\theta_2)) \\
&\rho(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3 \\
&+ \int_{v_0}^{v_{th}}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_1) f_{2w}(v_2-\theta_2)\rho(\theta_1,\theta_2,\theta_3,t)]d\theta_1 d\theta_2 d\theta_3
\end{aligned}
\tag{A.62}
$$

$J_{1,reset}(v_2,t)$ consists of three parts. The first line reflects the event that neuron 1 receives an external input and fires, which causes neuron 2 jumps to a subthreshold voltage v_2. The second line describes the event that neuron 1 receives an independent Poisson input and fires, which causes neuron3 fires and neuron2 jumps to a subthreshold voltage. Neuron 3 generates a synaptic input and promotes the membrane potential of neuron 2 to a subthreshold voltage v_2, which is described by the first term of line 2. The second term of line 2 is the loss of probability that neuron 2 at v_2 jumps to a higher voltage due to the fire of neuron 3. The third line reflects that neuron 3 receives an external input and fires, which cause neuron 1 fires and neuron 2 jumps to a subthreshold voltage. Firing of neuron1 cause a second jump of neuron2 to v_2. $J_{2,reset}(v_1,t)$ is identical to $J_{1,reset}(v_2,t)$ with roles reversed.

$$
\begin{aligned}
&J_{12,reset}(t) = \\
&\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}[F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_2)+F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_1)]\rho(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)[F_{2w}(v_2-\theta_2)-F_w(v_2-\theta_2)] \\
&+F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_3)[F_{2w}(v_1-\theta_1)-F_w(v_1-\theta_1)] \\
&+F_A(v_{th}-\theta_3)[F_w(v_{th}-\theta_1)F_{2w}(v_{th}-\theta_2) \\
&+F_w(v_{th}-\theta_2)F_{2w}(v_{th}-\theta_1)-F_w(v_{th}-\theta_1)F_w(v_{th}-\theta_2)]\rho(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3
\end{aligned}
\tag{A.63}
$$

$J_{12,reset}(t)$ contains the situations which are illustrated by figure(**??**).

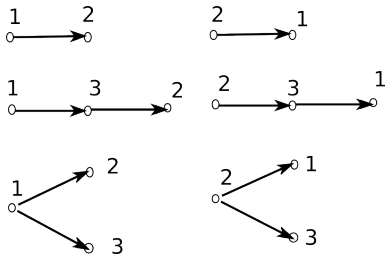


Figure A.2: Illustration of $J_{12,reset}$ for all-to-all coupled network with $N = 3$

A.6 The third order truncation for the kinetic theory equation

$$
\begin{aligned}
J_{1,reset}(v_2,t) &= \lambda(t)\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_1)\hat{\rho}(\theta_1,v_2,t)d\theta_1 \\
&+ \int_{v_0}^{v_{th}}\int_{v_0}^{v_2} F_A(v_{th}-\theta_1)f_w(v_2-\theta_2)\hat{\rho}(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+ (N-2)\int_{v_0}^{v_{th}}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} [F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_1)f_{2w}(v_2-\theta_2) \\
&+ F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)(f_{2w}(v_2-\theta_2)-f_w(v_2-\theta_2))]\hat{\rho}(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3
\end{aligned}
\tag{A.64}
$$

$J_{2,reset}(v_1,t)$ is similar to $J_{1,reset}$ with neuron 1 and neuron 2 reversed.

$$
\begin{aligned}
J_{12,reset}(t) &= \lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_2)\hat{\rho}(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_1)\hat{\rho}(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+ (N-2)\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} \{F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)F_{2w}(v_{th}-\theta_2) \\
&+ F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_3)F_{2w}(v_{th}-\theta_1) \\
&+ F_A(v_{th}-\theta_3)[F_w(v_{th}-\theta_1)F_{2w}(v_{th}-\theta_2)+F_w(v_{th}-\theta_2)F_{2w}(v_{th}-\theta_1) \\
&- F_w(v_{th}-\theta_1)F_w(v_{th}-\theta_2)]\}\hat{\rho}(\theta_,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3
\end{aligned}
\tag{A.65}
$$

We discretized the time interval $[0, 0.2]$s into subintervals of size $\bigtriangleup t = 0.001$s. And we discretized the domain $[v_0, v_{th}]\times[v_0, v_{th}]$ in v_1-v_2 space into squares of size $\bigtriangleup v \times \bigtriangleup v$ with $\bigtriangleup v = 0.02$. Let $v_{i,j} = ((i-\frac{1}{2})\bigtriangleup v, (j-\frac{1}{2})\bigtriangleup v)$, then $v_{i,j}$ is the center of square (i, j). Each

point $\rho_{i,j}(t)$ represents the integral of $\rho(v_1, v_2, t)$ over the square (i, j). When solving the evolution of probability density due to leakage, we use the finite-volume method. And we use midpoint rule to estimate the integrals in the equation.

Note that the third order truncation of kinetic theory equation is nonlinear. We use the simple explicit method to solve the nonlinear equation.

A.7 The fourth order truncation for the kinetic theory equation

$J_{1,reset}(v_2, t)$ is described as:

$$
\begin{aligned}
&J_{1,reset}(v_2,t) = \lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_2} F_A(v_{th}-\theta_1) f_w(v_2-\theta_2)\hat{\rho}(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+ (N-2)\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\int_{v_0}^{v_2} F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)[f_{2w}(v_2-\theta_2) - f_w(v_2-\theta_2)] \\
&\hat{\rho}(\theta_1,\theta_2\theta_3,t)d\theta_1 d\theta_2 d\theta_3 \\
&+ (N-2)\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_1) f_{2w}(v_2-\theta_2)\hat{\rho}(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3 \\
&+ (N-2)(N-3)\int_{v_0}^{v_{th}}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} \\
&\{F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)F_{2w}(v_{th}-\theta_4)[f_{3w}(v_2-\theta_2) - f_{2w}(v_2-\theta_2)] \\
&- \frac{1}{2}F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)F_w(v_{th}-\theta_4)[f_{3w}(v_2-\theta_2) - f_w(v_2-\theta_2)] \\
&+ F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_1)F_{2w}(v_{th}-\theta_4)[f_{3w}(v_{th}-\theta_2) - f_{2w}(v_{th}-\theta_2)] \\
&+ F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_4)[F_{2w}(v_{th}-\theta_1) - F_w(v_{th}-\theta_1)]f_{3w}(v_2-\theta_2)\} \\
&\hat{\rho}(\theta_1,\theta_2,\theta_3,\theta_4,t)d\theta_1 d\theta_2 d\theta_3 d\theta_4
\end{aligned}
\tag{A.66}
$$

It consists terms that are caused by external input directly, caused by one, two and three synaptic neurons respectively.

$J_{2,reset}(v_1,t)$ is symmetric to $J_{1,reset}(v_2,t)$ with roles of neuron 1 and neuron 2 reversed. $J_{12,reset}(t)$ is described by the following:

$$
\begin{aligned}
&J_{12,reset} = \lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_2)\hat{\rho}(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_1)\hat{\rho}(\theta_1,\theta_2,t)d\theta_1 d\theta_2 \\
&+(N-2)\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\{F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)[F_{2w}(v_{th}-\theta_2)-F_w(v_{th}-\theta_2)] \\
&+F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_3)[F_{2w}(v_{th}-\theta_1)-F_w(v_{th}-\theta_1)] \\
&+F_A(v_{th}-\theta_3)[F_w(v_{th}-\theta_1)F_{2w}(v_{th}-\theta_2)+F_w(v_{th}-\theta_2)F_{2w}(v_{th}-\theta_1) \\
&-F_w(v_{th}-\theta_1)F_w(v_{th}-\theta_2)]\}\hat{\rho}(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3 \\
&+(N-2)(N-3)\lambda(t)\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}}\{F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)F_{2w}(v_{th}-\theta_4) \\
&[F_{3w}(v_{th}-\theta_2)-F_{2w}(v_{th}-\theta_2)] \\
&-\frac{1}{2}F_A(v_{th}-\theta_1)F_w(v_{th}-\theta_3)F_w(v_{th}-\theta_4)[F_{3w}(v_{th}-\theta_2)-F_w(v_{th}-\theta_2)] \\
&+F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_1)F_{2w}(v_{th}-\theta_4)[F_{3w}(v_{th}-\theta_2)-F_{2w}(v_{th}-\theta_2)] \\
&+F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_3)F_{2w}(v_{th}-\theta_4)[F_{3w}(v_{th}-\theta_1)-F_{2w}(v_{th}-\theta_1)] \\
&-\frac{1}{2}F_A(v_{th}-\theta_2)F_w(v_{th}-\theta_3)F_w(v_{th}-\theta_4)[F_{3w}(v_{th}-\theta_1)-F_w(v_{th}-\theta_1)] \\
&+F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_2)F_{2w}(v_{th}-\theta_4)[F_{3w}(v_{th}-\theta_1)-F_{2w}(v_{th}-\theta_1)] \\
&+F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_4)[(F_{2w}(v_{th}-\theta_1)-F_w(v_{th}-\theta_1))F_{3w}(v_{th}-\theta_2) \\
&+(F_{2w}(v_{th}-\theta_2)-F_w(v_{th}-\theta_2))F_{3w}(v_{th}-\theta_1)+F_w(v_{th}-\theta_1)F_w(v_{th}-\theta_2) \\
&-F_{2w}(v_{th}-\theta_1)F_{2w}(v_{th}-\theta_2)]\}\hat{\rho}(\theta_1,\theta_2,\theta_3,\theta_4,t)d\theta_1 d\theta_2 d\theta_3 d\theta_4
\end{aligned}
\tag{A.67}
$$

$J_{12,reset}(t)$ contains the following situations, which are illustrated by figure(A.3).

The average firing rate of neuron 1 is

$$r_{ave}(t) = \int_{v_0}^{v_{th}} J_{1,reset}(v_2,t)dv_2 + J_{12,reset}$$

It consists of terms describing all possibilities causing a single neuron to fire. Neuron 1 can fire due to its external poisson input, or by spikes generated by up to three other

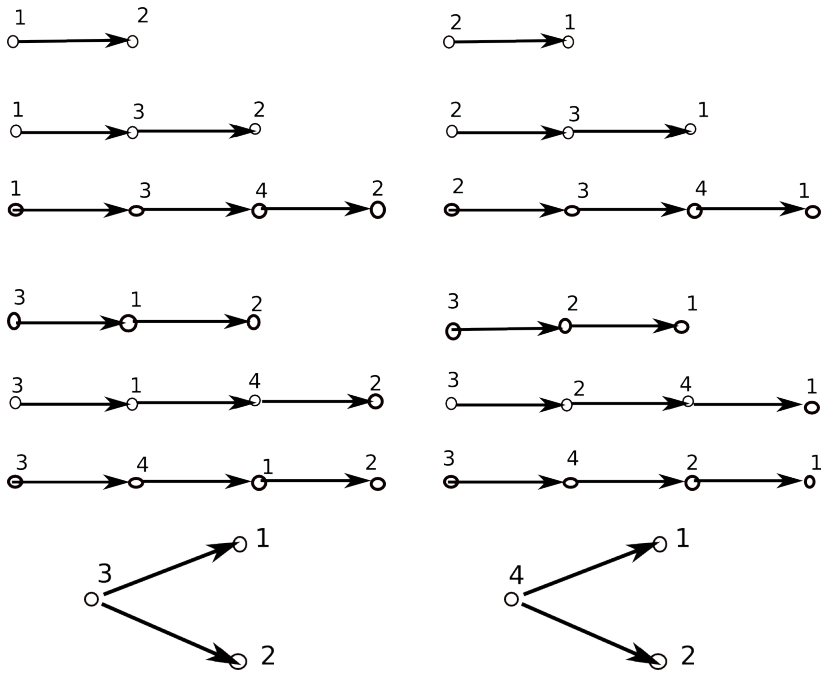


Figure A.3: Illustration of $J_{12,\text{reset}}$ for all-to-all coupled network with $N = 4$

neurons. For example, the last term of r_{ave} reflects the event that neuron 3 when receiving its external input, causing neuron 4 fires while neuron 1 and 2 at a subthreshold. The fire of neuron 4 causes neuron 1 and 2 jump again, which makes neuron 2 fire and neuron 1 still below threshold. Neuron 1 jumps for the third time and finally crosses threshold.

Now show that the population density function is conservative and can be rewritten in the conservative form: $\frac{\partial \rho}{\partial t} = -\nabla \cdot \vec{J}$ where $\vec{J}$ is the probability flux density. Note that the advection term is already in the divergence form. Then look at the term due to external poisson input:

$$\int_{v_0}^{v_1} f_A(v_1-\theta_1)\hat{\rho}(\theta_1, v_2, t)d\theta_1 - \hat{\rho}(v_1, v_2, t) = -\frac{\partial}{\partial v_1}\int_{v_0}^{v_1} F_A(v_1-\theta_1)\hat{\rho}(\theta_1, v_2, t)d\theta_1 \quad \text{(A.68)}$$

$$\int_{v_0}^{v_2} f_A(v_2-\theta_2)\hat{\rho}(v_1, \theta_2, t)d\theta_2 - \hat{\rho}(v_1, v_2, t) = -\frac{\partial}{\partial v_2}\int_{v_0}^{v_2} F_A(v_2-\theta_2)\hat{\rho}(v_1, \theta_2, t)d\theta_2 \quad \text{(A.69)}$$

The above equation hold since f_A is the probability density function for jump size A, and F_A is its complimentary cumulative distribution function.

If only one neuron in the population other than neuron 1 and 2 fires, the probability density is described below.

$$
\begin{aligned}
&\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)f_w(v_1-\theta_1)f_w(v_2-\theta_2)\hat{\rho}(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3 \\
&-\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)\hat{\rho}(v_1,v_2,\theta_3)d\theta_3 \\
&=-\frac{1}{2}\frac{\partial}{\partial v_1}[\int_{v_0}^{v_1}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_1-\theta_1)\hat{\rho}(\theta_1,v_2,\theta_3,t) \\
&+\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_1-\theta_1)f_w(v_2-\theta_2)\hat{\rho}(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3] \\
&-\frac{1}{2}\frac{\partial}{\partial v_2}[\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_2-\theta_2)\hat{\rho}(v_1,\theta_2,\theta_3,t) \\
&+\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_2-\theta_2)f_w(v_1-\theta_1)\hat{\rho}(\theta_1,\theta_2,\theta_3,t)d\theta_1 d\theta_2 d\theta_3]
\end{aligned}
\tag{A.70}
$$

The following is the conservative form of terms that firing of two other neurons causing neuron 1 and 2 to jump.

$$
\begin{aligned}
&\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_4) \\
&\quad[f_{2w}(v_1-\theta_1)f_{2w}(v_2-\theta_2)-f_w(v_1-\theta_1)f_w(v_2-\theta_2)]\hat{\rho}(\theta_1,\theta_2,\theta_3,\theta_4,t)d\theta_1 d\theta_2 d\theta_3 d\theta_4 \\
&=-\frac{1}{2}\frac{\partial}{\partial v_1}\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_4) \\
&[F_{2w}(v_1-\theta_1)f_{2w}(v_2-\theta_2)-F_w(v_1-\theta_1)f_w(v_2-\theta_2)]\hat{\rho}(\theta_1,\theta_2,\theta_3,\theta_4,t)d\theta_1 d\theta_2 d\theta_3 d\theta_4 \\
&-\frac{1}{2}\frac{\partial}{\partial v_2}\int_{v_0}^{v_1}\int_{v_0}^{v_2}\int_{v_0}^{v_{th}}\int_{v_0}^{v_{th}} F_A(v_{th}-\theta_3)F_w(v_{th}-\theta_4) \\
&[F_{2w}(v_2-\theta_2)f_{2w}(v_1-\theta_1)-F_w(v_2-\theta_2)f_w(v_1-\theta_1)]\hat{\rho}(\theta_1,\theta_2,\theta_3,\theta_4,t)d\theta_1 d\theta_2 d\theta_3 d\theta 4
\end{aligned}
\tag{A.71}
$$

$$
\begin{aligned}
&J_1(v_1, v_2, t) = -\frac{1}{\tau}(v - \varepsilon_r)\rho(v_1, v_2, t) \\
&+ \lambda(t) \int_{v_0}^{v_1} F_A(v_1 - \theta_1)\rho(\theta_1, v_2, t) \\
&+ \frac{N-2}{2}\lambda(t)[\int_{v_0}^{v_1} \int_{v_0}^{v_{th}} F_A(v_{th} - \theta_3)F_w(v_1 - \theta_1)\hat{\rho}(\theta_1, v_2, \theta_3, t) \\
&+ \int_{v_0}^{v_1} \int_{v_0}^{v_2} \int_{v_0}^{v_{th}} F_A(v_{th} - \theta_3)F_w(v_1 - \theta_1)f_w(v_2 - \theta_2)\hat{\rho}(\theta_1, \theta_2, \theta_3, t)d\theta_1 d\theta_2 d\theta_3] \\
&+ \frac{(N-2)(N-3)}{2}\lambda(t) \int_{v_0}^{v_1} \int_{v_0}^{v_2} \int_{v_0}^{v_{th}} \int_{v_0}^{v_{th}} F_A(v_{th} - \theta_3)F_w(v_{th} - \theta_4) \\
&[F_{2w}(v_1 - \theta_1)f_{2w}(v_2 - \theta_2) - F_w(v_1 - \theta_1)f_w(v_2 - \theta_2)]\hat{\rho}(\theta_1, \theta_2, \theta_3, \theta_4, t)d\theta_1 d\theta_2 d\theta_3 d\theta_4
\end{aligned}
\tag{A.72}
$$

J_2 is symmetric to J_1 with neuron 1 and neuron 2 reversed. Let $\vec{J} = (J_1, J_2)$.It is obvious that the equation can be written as

$$
\frac{\partial \rho}{\partial t} = -\nabla \cdot \vec{J} + \delta v_1 - v_0 J_{1,reset}(v_2, t) + \delta(v_2 - v_0)J_{2,reset}(v_1, t) + \delta(v_1 - v_0)\delta(v_2 - v_0)J_{12,reset}(t)
\tag{A.73}
$$

The equation is globally conservative by checking

$$
\int_{v_0}^{v_{th}} \int_{v_0}^{v_{th}} J(v_1, v_2, t)dv_1 dv_2 + \int_{v_0}^{v_{th}} J_{1,reset}(v_2, t)dv_2 + \int_{v_0}^{v_{th}} J_{2,reset}(v_1, t)dv_1 + J_{12,reset}(t) = 0
\tag{A.74}
$$