Estimating the Topology of Neural Networks from Distributed Observations

Roxana Alexandru, Pranav Malhotra, Stephanie Reynolds and Pier Luigi Dragotti
Electrical and Electronic Engineering Department
Imperial College London
email: {roxana.alexandru12, pranav.malhotra13, stephanie.reynolds09, p.dragotti}@imperial.ac.uk

***Abstract*—We address the problem of estimating the effective connectivity of the brain network, using the input stimulus model proposed by Izhikevich in [1], and adapting this to our framework. We first analyse the temporal dynamics of neural networks, showing that the spike propagation within the brain can be modelled as a diffusion process. This analysis helps prove the suitability of NetRate algorithm proposed by Rodriguez in [2] to infer the structure of biological neural networks. Finally, we present simulation results using synthetic data to verify the performance of the topology estimation algorithm.**

***Index Terms*—Neural networks, network topology inference, stability analysis of spike propagation, Izhikevich neuron model, Brian simulator, NetRate algorithm.**

I. Introduction

Neuronal network connectivity has been studied at multiple spatial scales with the aim of understanding function in both healthy and unhealthy brains [3]. Analysis of connectivity in fMRI data identified network properties that could be used as diagnostic markers for Alzheimer's disease [4] and Schizophrenia [5]. Moreover, the prospects for understanding network connectivity at smaller spatial scales have been significantly expanded recently by the development of tools to simultaneously monitor and stimulate neuronal microcircuit activity [6]. Using multi-photon imaging and genetically encoded sensors, it is possible to simultaneously interrogate the activity of thousands of neurons at single-cell and single-spike resolution [7]. These techniques enable a whole neural circuit to be monitored at sufficient resolution to decode its connectivity.

Early research in the area of network tomography, a term coined by Vardi in [8], focused on inferring link-layer topologies, and the same techniques are currently being applied to estimate the topology of biological neural networks. Many algorithms for effective connectivity inference apply statistical measures such as the Granger Causality test [9]. The Generalized Linear Model (GLM) is also a common technique used to infer neural networks. For example, the successful reconstruction of a physiological circuit was achieved in [10], using coupling functions composed of spline basis functions, and modelling the spike trains as point processes. Research in the area also explored using the perceptron learning rule for neural network inference [11].

In this paper, we present a novel method to solve the biological neural network inference problem, based on adapting the network estimation algorithm NetRate proposed by Rodriguez in [2]. In particular, we focus on determining the effective connectivity of the brain network. We use Izhikevich's input stimulus model to analyse the behaviour of a neuron. The stability of the system of differential equations proposed by Izhikevich helps us underpin the time when a neuron enters the unstable region (i.e. spikes), and hence, the cause for its spike - either pre-synaptic neurons or noise from outside the observed network. Through the analysis of the cause of spikes, we demonstrate that the system can be probabilistically described, even though the individual neuron behaviour is modelled through a set of deterministic differential equations. Simulations help us determine the transmission likelihood between individual neurons within the network. These are then used by the NetRate algorithm, which infers the weighted connections within the neural network.

This paper is organized as follows. In Section II we present an overview of information processing in neural networks, including the notion of connectivity, and the spiking neuron model used. Section III gives an overview of the NetRate algorithm, which is used to infer the connectivity within a graph. Then, in Section IV we prove the feasibility of using NetRate to achieve neural network inference. The analysis revolves around the temporal dynamics of neural networks, such as formalizing the notion of a cascade within the network, using stability analysis to probabilistically describe spike propagation, and finally, formulating the transmission likelihood between the individual neurons. The experimental results in Section V show the performance of the adapted inference algorithm, on synthetic data. Finally, conclusions are presented in Section VI.

II. Problem Formulation

In order to study the problem of neural network topology inference, one has to first understand how neural networks process information. This is influenced by the connectivity of the neurons, as well as by the way in which synaptic wiring may cause a neuron to spike.

A. Definition of Connectivity

With reference to brain networks, there are three main definitions of connectivity. First, the structural connectivity describes the physical connections between different neurons, or parts of the brain. Second, functional connectivity refers to statistical dependencies between different units in the brain.

The statistical associations are described using tools such as correlation [12], coherence [13] and mutual information [14]. Third, effective connectivity describes the causal relationships between neurons, using techniques such as the Granger causality test [15] and Dynamic Causal Models [16].

The method presented in this paper aims to infer the effective connectivity within a static network of neurons, modelled through the adjacency matrix. If i is a pre-synaptic neuron of j, then the (i, j) entry in the adjacency matrix will have a non-zero value proportional to the strength of the directed connection from i to j.

B. Spiking Neuron Model

One of the most commonly used spiking models in computational neuroscience is the Izhikevich model [1], which is able to replicate thoroughly the spiking behavior of biological neurons. The dynamics of the spiking events are given by the following two equations:

$$\frac{dv(t)}{dt} = 0.04v^2(t) + 5v(t) + 140 - u(t) + I, \quad (1)$$

$$\frac{du(t)}{dt} = a(bv(t) - u(t)), \quad (2)$$

where:

$v(t)$ = membrane potential,
$u(t)$ = membrane recovery,
I = input noise due to unobserved neurons.

The four static parameters a, b, c and d are used to model different types of neurons. The two main classes are inhibitory and excitatory neurons, which we can further classify into regular spiking neurons and chattering neurons. The method presented in this paper focuses on inference of a network consisting of only excitatory regular spiking neurons. To model this, the tuning parameters are set as in [1], to the values: $a = 0.02, b = 0.2, c = -65$, and $d = 8$.

III. NetRate Algorithm for Network Inference

The neuron connections are inferred using the NetRate algorithm proposed by Rodriguez in [2]. The algorithm is designed in the framework of diffusion processes over static directed networks, with unknown connections. The spreading model is the susceptible-infected one, where a node is initially susceptible and once it becomes infected, it cannot recover from the disease.

NetRate relies on modelling the pairwise interactions between the nodes using a probabilistic approach. In this sense, each directed edge from j to i is assigned the conditional likelihood $f(t_i|t_j, \alpha_{j,i})$, of node i to be infected at time t_i given node j was infected at time t_j. The parameters $\alpha_{j,i}$ represent the transmission rates associated with edges, and we note that a rate $\alpha_{j,i} = 0$ represents the absence of an edge between the two nodes.

In terms of data required, the algorithm assumes access to multiple independent cascades of information. Each cascade C is generated by randomly selecting a source node, which is infected at time $t = 0$, and monitoring the entire network within the observation window $[0, T_c]$ after the infection starts. The information is allowed to spread through the network according to the likelihoods $f(t_i|t_j, \alpha_{j,i})$ of each pair of nodes in the network. The generated cascade C contains the first infection absolute times of all the N nodes in the network, within a window of observation, $\boldsymbol{t}^c := (t_1^c, ..., t_N^c)$. We note that NetRate is based on the assumption of complete knowledge of the infections occurring during the observation window.

The aim of the NetRate algorithm is to infer the transmission edges $\alpha_{j,i}$ using multiple independent cascades of information (i.e. observing when nodes in the network get the infection), and knowledge of the shape of the likelihood function $f(t_i|t_j, \alpha_{j,i})$. The method used is the maximum likelihood (ML) estimation, which finds the optimal network connections and their corresponding weights, such that the likelihood of the observed cascades is maximized.

The likelihood of a cascade is calculated based on the survival function, which is the probability that a node is uninfected until the end of observation, T^c, given the infections at the other nodes. The probability that node i is not infected by node j by time t_i is given by the survival function:

$$S(t_i|t_j, \alpha_{j,i}) = 1 - F(t_i|t_j, \alpha_{j,i}), \quad (3)$$

where $F(t_i|t_j, \alpha_{j,i})$ is the cumulative density function, calculated from the transmission likelihood $f(t_i|t_j, \alpha_{j,i})$.
Moreover, the hazard function is defined as the instantenous infection rate, and given by:

$$H(t_i|t_j, \alpha_{j,i}) = \frac{f(t_i|t_j, \alpha_{j,i})}{S(t_i|t_j, \alpha_{j,i})}. \quad (4)$$

Assuming the infections happen independently across different network edges, the likelihood of a cascade of infection times $\boldsymbol{t}^c$ is:

$$f(\boldsymbol{t}^c; A) = \prod_{t_i < T^c} \prod_{t_m > T^c} S(T|t_i, \alpha_{i,m}) \times \prod_{k:t_k<t_i} S(t_i|t_k, \alpha_{k,i}) \sum_{j:t_j<t_i} H(t_i|t_j, \alpha_{j,i}) \quad (5)$$

The cascades are assumed independent, and hence the likelihood of all cascades is the product of individual likelihoods. Hence, the NetRate algorithm aims to solve the network inference problem given by:

$$\begin{aligned} &\underset{A}{\text{minimize}} && -\sum_{c \in C} log f(\boldsymbol{t}^c; A) \\ &\text{subject to} && \alpha_{j,i} \geq 0, j = 1, \ldots, N, i \neq j, \\ &\text{where} && A := \{\alpha_{j,i} | i, j = 1, \ldots, N, i \neq j\} \end{aligned} \quad (6)$$

The optimization problem defined by Eq.(6) is convex when $f(t_i|t_j, \alpha_{j,i})$ follows the exponential, power-law or Rayleigh models and a detailed proof of this result is found in [2].

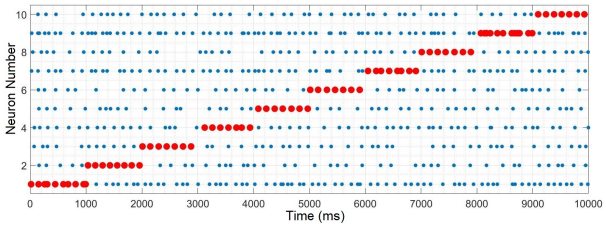


Fig. 1: Cascades generated by sequentially stimulating each node in the network. Each time a neuron is given a constant input, it periodically spikes for 1000ms (illustrated in red), generating a different cascade in the network.

IV. Temporal Dynamics of Neural Networks

In this section we show that interactions within a neural network can be modelled as a diffusion process equivalent to the one assumed by NetRate.
First, the propagation of action potentials within a neural network is governed by the transmission rates between individual nodes, and creates a cascade of initial infection times.
Second, the neuron has a binary state: it is either spiking or not. In order to comply to the framework imposed by the NetRate algorithm, we need to ensure a neuron only spikes once, which is achieved by limiting the observation horizon.
Third, although the spike propagation model proposed by Izhikevich illustrates a deterministic behaviour of individual neurons, we show that it is feasible to describe a neural network probabilistically.

A. Cascade Generation

The cascades represent the input to the NetRate algorithm, which uses these initial infection times to determine the pairwise transmission rates within the network. Supplying an excitatory regular spiking neuron with a constant input in Eq.(1) will make neurons spike periodically, generating an independent cascade within each period. This is illustrated in Fig. 1, where node 1 spikes 10 times within the interval $[0, 1000]$ms, generating 10 cascades. For all the other neurons, the input I will be Gaussian random noise. The proposed input stimulus model ensures that the generated cascades are independent. A constant input is achievable in practice, using recent advances in optogenetic actuators [6].

B. Stability Region of Regular Spiking Excitatory Neuron

To identify the causes of neuron spikes, we analyze the stability of the dynamical system of Eq.(1) and (2), around the equilibirium points. The random noise I is an input to the system and it can therefore be excluded from the stability analysis. The solutions to the set of equations is given by the following two equilibria: $(v_1, u_1) = (-70, -14)$ and $(v_2, u_2) = (-50, -10)$. The Jacobian helps us find the stability of the linearized system, around the equilibrium points. The eigenvalues of the Jacobian for the first equilibrium, $(v_1, u_1) = (-70, -14)$, are negative real numbers, and hence, this point is stable. On the other hand, one of the eigenvalues corresponding to the second equilibrium $(v_2, u_2) = (-50, -10)$, is positive, making this point unstable. This means that a small perturbation around this point would make the neuron's potential diverge to infinity, which is equivalent to a spike. When this happens, we reset the membrane potential v to the value of $c = -65$, once this value goes over the threshold value of $v = 30$, shown in Fig. 2. Furthermore, the second eigenvalue corresponding to $(v_2, u_2) = (-50, -10)$ is negative and real, which allows us to determine the stability boundary of the system, as illustrated in Fig. 2.

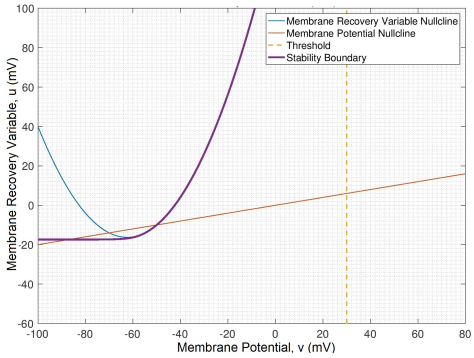


Fig. 2: Stability boundary of excitatory regular spiking neurons.

A neuron in a state above the stability boundary will converge to the equilibrium point $(v_1, u_1) = (-70, -14)$, whereas the points below the boundary will result in spiking of the neuron and the resetting of the v and u states. The trajectory of the neuron's state is illustrated in Fig. 3, where the time to spike is approximately 7s when the neuron's initial unstable state is $(v_{init}, u_{init}) = (-80, -20)$. On the other hand, if the initial state was $(v_{init}, u_{init}) = (-40, -30)$, the time to spike would be 1s. This highlights the fact that the neuron's initial values of the membrane potential and recovery will determine the time this neuron takes to spike.

Moreover, the pre-synaptic neuron j which triggers i, spikes at the moment when i enters the unstable region. If no such neuron j exists, then neuron i fires as a result of unobserved neuron activity.

C. Formulation of Transmission Likelihood

Given the general shape of the underlying distribution $f(t_i|t_j, \alpha_{j,i})$ which describes how infections spread within the network, the NetRate algorithm is able to infer the values of the transmission rates $\alpha_{j,i}$.
For the purpose of this analysis, it is assumed that the topology of the network is known.
Then, a histogram of time delays between the pre-synaptic and post-synaptic spikes can be formed for each value of $\alpha_{j,i}$, using the following method. For each neuron i that spikes, we find the neuron j which spiked at the moment when i became unstable. We treat j as a cause for neuron i's spike, and measure the time delay between the two spikes as $t_{j,i} = t_j - t_i$. This firing event is then placed into the bin corresponding to the time delay $t_{j,i}$ and the value $\alpha_{j,i}$. If no neuron is found to have caused i to enter the unstable region, then the firing

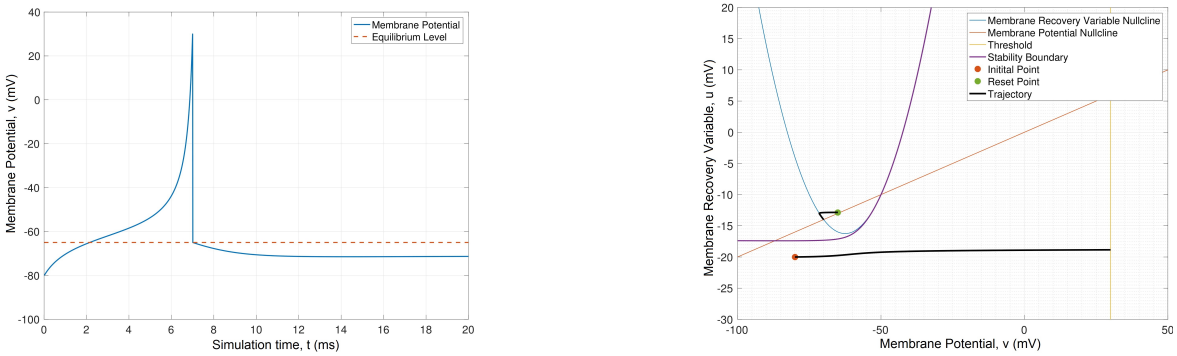


Fig. 3: Trajectories of a neuron's state for a given initial point.

event is not included in any of the histograms.
Fig. 4 explains how the shape of the transmission likelihood is formed. For example, neuron 5 fires at time $t = 257$. It enters the unstable region at time $t = 254$, the exact time when neuron 6 fires. The transmission rate from neuron 6 to neuron 5 is $\alpha_{6,5} = 21$, and the time delay between the two spikes is $t_{6,5} = 3$. Hence, we place this event in the third bin of the histogram corresponding to parameter value $\alpha = 21$.

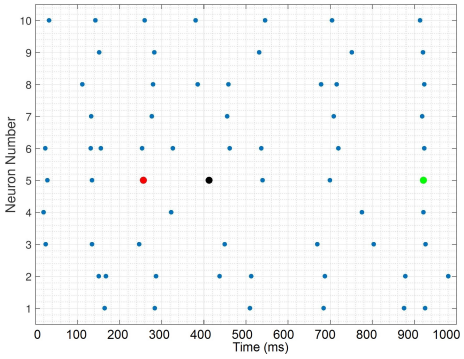


Fig. 4: Spike propagation used to infer the shape of the transmission likelihood.

The histograms in Fig. 5 show the conditional likelihood function $f(t_i|t_j, \alpha_{j,i})$, for different values of the transmission rate parameter. For small values of $\alpha_{j,i}$, the distribution is approximately exponential, whereas for large values of the parameter, the distribution becomes approximately Rayleigh. NetRate algorithm assumes a unique distribution for all the transmission rates, and in this context, we choose Rayleigh distribution to model the propagation of spikes through the neural network. This is because we aim to identify with greater accuracy the edges that have larger α weights, rather than minor edges.

V. Experimental Results

The following results demonstrate the suitability of NetRate for biological neural network inference. The performance is analysed using accuracy, recall and precision.
Accuracy is defined as $\frac{\sum_{j,i}|J(\alpha_{j,i})-J(\alpha^*_{j,i})|}{\sum_{j,i}J(\alpha_{j,i})+\sum_{j,i}J(\alpha^*_{j,i})}$, where $J(\alpha) = 1$ if $\alpha > 0$ and $J(\alpha) = 0$ otherwise. This metric captures the situations when an edge is wrongly inferred, e.g. it is inferred as absent, but is actually present.
Recall is calculated as the ratio of true edges in the inferred network, to the absolute number of true edges.
Finally, precision is the ratio of true inferred edges to the absolute number of inferred edges.

In order to preserve the sparsity of the inferred network, we set some of the estimated edges to 0. In the case of an N nodes small-world network of average node degree $d = 2$, we only keep the $d \times N$ major edges. Similarly, if the sparseness of a random graph as defined in [17] is s, then we consider only the first $s \times N^2$ major edges. Fig. 6 shows the results for a random geometric graph, and small-world network of 10 nodes each. The software used to generate the spike data is Brian Simulator [17]. This package also generates the random graph, whilst the small-world network was created using the Watts-Strogatz model [18]. In addition, the optimisation problem imposed by NetRate is solved using the CVX package [19]. The sparseness of the random graph is 0.1 and the small-world network has average node degree $d = 2$ and rewiring probability $\beta = 0.2$. Each node is stimulated for 1000ms, as described in Section IV-A. The experiment is repeated over 10 different networks and the results are then averaged. From the metrics in Fig. 6 it is evident that NetRate is indeed a suitable algorithm for biological neural network inference, achieving more than 65% accuracy and 60% precision for both small-world and random graphs. The recall is 65% for the small-world network, and 74% for a random graph, proving that the majority of the detected edges are indeed true edges. Finally, Fig. 7 shows that the performance of the algorithm is dependent on the absolute number of cascades. The performance generally increases as we lenghten the simulation time for each node in the network, and hence, as we increase the number of cascades.

VI. Conclusions

In this paper we proposed a novel method to infer the topologies of biological neural networks, using the NetRate algorithm. We have shown that it is possible to model propagation of neuron action potentials as a diffusion process. Moreover, we have demonstrated the probabilistic nature of

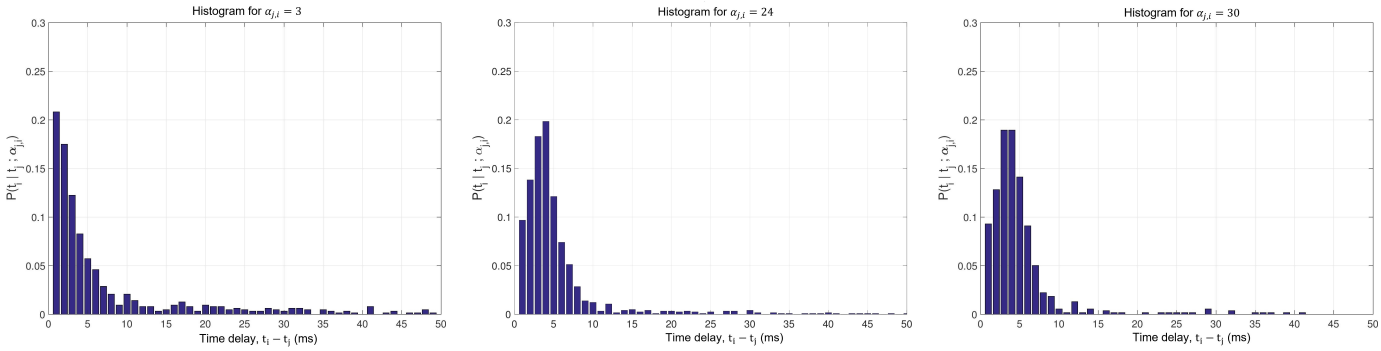


Fig. 5: Empirical node-to-node transmission probability distribution.

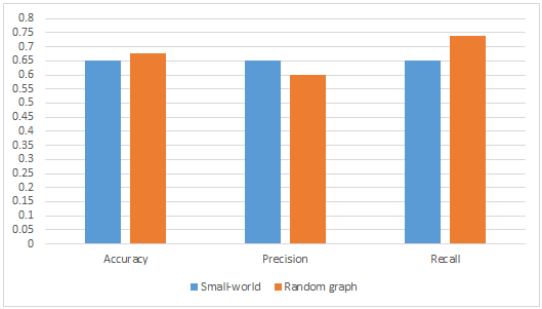


Fig. 6: Averaged performance metrics for inference of 10-node networks.

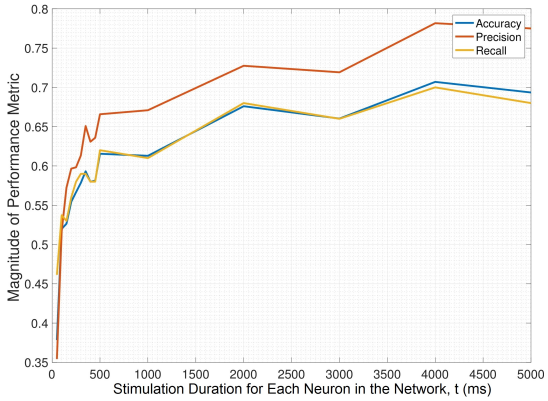


Fig. 7: Averaged performance metrics as a function of the stimulation time of each neuron in the network.

spike propagation, and empirically found the shape of the likelihood function. This is approximately Rayleigh for large values of the transmission rate parameter, which ensures the optimisation problem NetRate aims to solve is convex. Finally, experimental results further indicate the suitability of the proposed method for brain structure estimation.

References

[1] Eugene M Izhikevich. Simple Model of Spiking Neurons. *IEEE Transactions on Neural Networks*, 14(6), 2003.

[2] Manuel Gomez Rodriguez, David Balduzzi, and Bernhard Schölkopf. Uncovering the Temporal Dynamics of Diffusion Networks. may 2011.

[3] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, mar 2009.

[4] Kaustubh Supekar, Vinod Menon, Daniel Rubin, Mark Musen, and Michael D Greicius. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS computational biology*, 4(6):e1000100, jun 2008.

[5] Yong Liu, Meng Liang, Yuan Zhou, Yong He, Yihui Hao, Ming Song, Chunshui Yu, Haihong Liu, Zhening Liu, and Tianzi Jiang. Disrupted small-world networks in schizophrenia. *Brain*, 131(4):945–961, apr 2008.

[6] Adam M Packer, Lloyd E Russell, Henry W P Dalgleish, and Michael Häusser. Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. *Nature Methods*, 12(2):140–146, feb 2015.

[7] Simon P. Peron, Jeremy Freeman, Vijay Iyer, Caiying Guo, and Karel Svoboda. A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. *Neuron*, 86(3):783–799, may 2015.

[8] Y. Vardi. Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data. *Journal of the American Statistical Association*, 91(433):365, mar 1996.

[9] Steven L. Bressler and Anil K. Seth. WienerGranger Causality: A well established methodology. *NeuroImage*, 58(2):323–329, sep 2011.

[10] Felipe Gerhard, Tilman Kispersky, Gabrielle J. Gutierrez, Eve Marder, Mark Kramer, and Uri Eden. Successful Reconstruction of a Physiological Circuit with Known Connectivity from Spiking Activity Alone. *PLoS Computational Biology*, 9(7):e1003138, jul 2013.

[11] Amin Karbasi, Amir Hesam Salavati, and Martin Vetterli. Learning network structures from firing patterns. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 699–703. IEEE, mar 2016.

[12] Emery N Brown, Robert E Kass, and Partha P Mitra. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, 7(5):456–461, may 2004.

[13] Pascal Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10):474–480, oct 2005.

[14] Elad Schneidman, Susanne Still, Michael J Berry Ii, and William Bialek. Network Information and Connected Correlations. 2003.

[15] Andrea Brovelli, Mingzhou Ding, Anders Ledberg, Yonghong Chen, Richard Nakamura, and Steven L Bressler. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26):9849–54, jun 2004.

[16] K J Friston, L Harrison, and W Penny. Dynamic causal modelling. *NeuroImage*, 19(4):1273–302, aug 2003.

[17] Dan F M. Goodman and Romain Brette. The Brian simulator. *Frontiers in Neuroscience*, 3(2):192–197, sep 2009.

[18] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of small-world' networks. *Nature*, 393(6684):440–442, jun 1998.

[19] Michael Grant, Stephen Boyd, and Yinyu Ye. Cvx: Matlab software for disciplined convex programming.