# LEARNING NETWORK STRUCTURES FROM FIRING PATTERNS

Amin Karbasi<sup>†</sup> Amir Hesam Salavati<sup>\*</sup> Martin Vetterli<sup>\*</sup>

<sup>†</sup> School of Engineering and Applied Science, Yale, USA \*School of Computer and Communication Sciences, EPFL, Switzerland

# ABSTRACT

How can we decipher the hidden structure of a network based on limited observations? This question arises in many scenarios ranging from social to wireless and to neural networks. In such settings, we typically observe the nodes' behaviors (e.g., the time a node learns about a piece of information, or the time a node gets infected by a disease), and we are interested in inferring the true network over which the diffusion takes place. In this paper, we consider this problem over a neural network where our aim is to reconstruct the connectivity between neurons merely by observing their firing activity. We develop an iterative NEUral INFerence algorithm (NEUINF) to identify the type of effective neural connections (i.e. excitatory/inhibitory) based on the Perceptron learning rule. We provide theoretical bounds on the average performance of NEUINF as well as numerical analysis to compare the performance of the proposed approach to previous art.

*Index Terms*— Neural network tomography, connectome mapping, inverse signal problems, perceptron rule

### 1. INTRODUCTION

Reconstruction of neuronal network connectivity has been a major challenge for the past decades. Currently, invasive procedures are the only "reliable" approach to map the connectome. However, these approaches are prohibitively complex and time-consuming: it took more than 10 expert/year to map the whole connectome of C. Elegans, comprising only 302 neurons and a few thousands synaptic connections [1]. To map the whole brain of *fruit flies*, with around 10,000 neurons, we would have to spend around 4700 expert/year [2, 3]. Following the same approach and using current technology, it is estimated that it will take around 14 billion man/year to completely map the human brain's connectome [2]. Although there is an increasing effort to make some parts of the invasive procedures automated, such approaches remain impractical for mid-sized/large networks. Furthermore, the current invasive techniques cannot be applied to live animals/humans, as it involves dissecting the brain.

As a result, *inverse* methods with the focus on mapping the (whole or partial) connectome from the activity of the neurons are compelling (or perhaps the only viable) alternatives. They can be applied to live specimen and potentially require much less time and labor. Furthermore, rapid advancements in recording technologies has made it possible to simultaneously monitor the activities of tens [4] to hundreds of neurons [5], [6] (for a good review, see [7]). Upcoming technologies will increase this number significantly in near future [8]. These advancements will soon provide us with abundant neural data that can be used by inference algorithms to inversely map neural connectomes.

In this paper, we propose such an inverse approach for identifying the *type* of (effective) connections between neurons that reflects their causal relationship. More precisely, we consider a neural network where neurons follow the standard Leaky Integrate-and-Fire (LIF) model [9] and connections are associated with random delays. We propose an iterative NEUral INFerence algorithm NEUINF to identify both the effective connections and their types (i.e., excitatory/inhibitory). We provide some theoretical estimates on the performance of the algorithm. Furthermore, through numerical simulations we validate the results and see how they compare to the previous state of the art.

# 2. RELATED WORK

Identifying neural connections from a set of recorded neural activities is an instance of network tomography [10] and has been extensively studied in the past *Cross Correlogram* (CC) is perhaps the most widely-used method to identify (functional) connection between a pair of neurons [7]. However, approaches based on CC usually fall short of identifying *causal* relation or *effective* connectivity of neurons.

Another very popular approach to identify effective connectivities is Generalized Linear Models (GLMs) [11]. Methods based on GLM essentially take the effects of stimulus, self-history of the neurons, and contribution of other neurons into account and calculate the *filters* through which all these factors affect the firing response of each neuron. GLM was recently used in reconstructing a real physiological circuit from recorded neural data [12]. The approaches based on GLMs are provably accurate (i.e. they identify the correct set of connections in the underlying graph) *if* the neural model used to generate spike data matches the one used in GLM. Otherwise, the final estimation will have some bias and variance from the correct results [13]. Furthermore, traditional GLM approaches require high computational cost. To this end, recently, approximation methods have been suggested to resolve this issue [13], [14]. Nevertheless, the convergence still only applies when the model for neurons and that of GLM closely match each other.

Bayesian approaches are also widely used. In [15], a Maximum a Posteriori (MAP) approach is applied to solve the problem of connection identification, with accurate results in the regime of limited data at the expense of prohibitive computational costs.

Recently, another line of work has focused on the connection mapping problem that are mostly tailored to LIF neurons. In particular, Bussel et al. [16] addresses the problem by converting the non-linear firing behavior of LIF neurons into a set of linear equations, which can be solved given a sufficient number of recorded samples. While being efficient, this algorithm is highly sensitive to the accuracy of spike time and heavily relies on the knowledge of model parameters, e.g. synaptic delays, which are very difficult to obtain. Additionally, Memmesheimer et al. [17] proposed an inference algorithm based on the Perceptron learning rule, similar to Baldassi et al. [18], for which they proved that under accurate spike times it identifies a simple n-to-1 feed forward network. They also proposed a heuristic extension that works with *finite* precision in recorded spike times. Nevertheless, their model does not take into account (random) synaptic delays.

Finally, we should mention that the *consistency problem* even for a *n*-to-1 feed forward network is NP-hard. In words, determining whether or not there exists a set of delays and weights such that we can fully match the set of input firing patterns to the output is very difficult [19]. Although this result does not necessarily mean finding such a configuration is impossible, it shows that finding provable "positive learning results" for the case of spiking neuron is quite difficult.

Our proposed approach NEUINF in this paper is similar to GLMs, namely, it does not rely on the knowledge of propagation delays. However, it differs in terms of the iterative algorithm and the provided theoretical analysis as it only relies on very general assumptions about the nature of connections. Furthermore, while the algorithm update rule is inspired by the Perceptron learning rule, it differs from previous similar approaches as the effect of propagation delays are considered in NEUINF.

#### 3. MODEL AND PROBLEM STATEMENT

As mentioned earlier, in this paper we consider a Leaky Integrate-and-Fire (LIF) model for the neurons [9]. In this model, each neuron accumulates the incoming (weighted) spikes from all of its neighbors and fires if the net sum exceeds a threshold  $\theta$ . Otherwise, the membrane voltage decays exponentially fast.

As for the graph, we assume that there are two types of connections: *excitatory* and *inhibitory*. In accordance with biological neuronal networks, we assume that the excitatory connections are more numerous than the inhibitory ones. Furthermore, we assume that the weight of connections is



**Fig. 1**: Network model: a recurrent neural network where we try to identify the incoming connections of node b by observing the spike trains  $x_1(t), \ldots, x_5(t)$  and y(t).

fixed and to keep the network balanced, we require for the inhibitory connections to have a larger weight (in magnitude) than excitatory ones. We typically set the weight of an excitatory connection to +1mV and that of an inhibitory connection to  $-\delta$ mV, where  $\delta = n_{exc}/n_{inh}$ , and  $n_{exc}$  and  $n_{inh}$  are the number of excitatory and inhibitory neurons, respectively. Following Dale's principle [20], we assume that all outgoing connections of a neuron have the same type, i.e., they are all either excitatory or inhibitory.

We also assume that neural connections have intrinsic delays which represent the time it takes for the information to propagate through the axons and synapses. The delay for each link is assumed to be a random number in the interval  $[0, d_{\max}]$ , where  $d_{\max} > 0$  is the maximum delay. The delays do not change and once assigned, remains fixed. Figure 1 illustrates the model used to generate data in the simplest case considered in this paper.

The goal of this paper is to infer the underlying connectivity by only observing the spike trains. In this paper, we are only interested in identifying the synaptic connections among the neurons for which we have recorded spike trains. Identifying the connections to *hidden* neurons is outside the scope of this paper and still remains one of the major challenges in connection mapping from spikes.<sup>1</sup> We propose an algorithm called NEUINF that iteratively identifies the connections by producing an *analogue* association matrix which reflects the accumulated belief of each connection. This matrix can then be transformed into a ternary adjacency matrix to infer the type of connections: void (no connection), excitatory, or inhibitory.

## 4. RECONSTRUCTION ALGORITHM: NEUINF

In order to identify the connections, we can consider each post-synaptic neuron separately and find its incoming connections. Let  $x_i(t)$  and y(t) denote the firing state of the *i*-th pre-synaptic and the post-synaptic neuron at time *t*. For the ease of presentation, let us assume momentarily that the firing state of the post-synaptic neuron at time *t* depends only

<sup>&</sup>lt;sup>1</sup>Note that we still consider the effect of hidden neurons through a *random* "outside traffic" affecting the states of neurons in the network.

on the states of its neighbors at time t. Thus, we effectively ignore the propagation delay and the integration procedure. Our goal is to find a set of weights W such that the average quadratic error E over a recording period T is minimized:

$$\min_{W} E = \min_{W} \frac{1}{T} \sum_{t=1}^{T} \left| f\left( \sum_{i=1}^{n} W_{i} x_{i}(t) - \theta \right) - y(t) \right|^{2}.$$
 (1)

Here,  $W_i$  is our *belief* about the connection weight from the *i*-th pre-synaptic to the post-synaptic neuron and f(.) is the Heaviside step function (or its continuous approximation). By taking the derivative, we obtain

$$\nabla E = \frac{2}{T} \sum_{t=1}^{T} \left( f\left( \langle W, X_t \rangle - \theta \right) - y(t) \right) f'\left( \langle W, X_t \rangle - \theta \right) X_t^{\top}$$
(2)

where  $X_t = (x_1(t), \ldots, x_n(t))$  is the state of pre-synaptic neurons at time t. Now, we can iteratively update our beliefs about W according to the following update rule:

$$W(\tau + 1) = W(\tau) - \alpha_{\tau} \nabla E(\tau), \qquad (3)$$

where  $\alpha_{\tau} > 0$  is a small number representing the learning rate. The above update rule follows the standard gradient descent approach. However, a closer look at (2) reveals some interesting characteristics: the weight  $W_i$ , will be updated (for a given t) if both of the following conditions are satisfied: 1)  $x_i(t) = 1$ , and 2)  $f(\langle W(\tau), X_t \rangle - \theta) \neq y(t)$ . In words, we update the weight  $W_i$  in round  $\tau$  if and only if the pre-synaptic neuron *i* fires and  $W(\tau)$  does not correctly predict the state y(t) of the post-synaptic neuron. Note that these two conditions are in fact reminiscent of the well-known Perceptron learning rule in neural networks [21].

Equipped with these observations, we propose NEUINF, as defined in Algorithm 1, that does not use the derivative of function f(.) and takes into account both the unknown propagation delays and the neural integration procedure. Since we do not know the synaptic delays, we define an *integration window*  $\Delta$  such that only the firings during this window are considered for the updates.

Algorithm 1 is deterministic. In practice, we found that a stochastic version, called STOCHASTIC NEUINF, works much better. The only difference is in the weight update rule. In each iteration  $\tau$ , the stochastic version updates the weights  $\Delta W_i \neq 0$  with some probability  $\beta$  independently at random. In the deterministic version,  $\beta = 1$ .

In order to analyze the performance of NEUINF we make the following mild assumptions:

- (A1) The probability of firing for a (pre-synaptic) neuron does not depend on it being excitatory or inhibitory.
- (A2) Excitatory (inhibitory) connections increase (resp. decrease) the probability of the post-synaptic neuron to fire.

#### Algorithm 1 NEUINF

**Input:** The observations  $\{x_i(t)\}$  and  $\{y(t)\}$  over period T, a maximum number of iterations  $\tau_{\text{max}}$ .

**Output:** Connections Belief Matrix W

Initialize  $W(0) = (0, ..., 0)_{1 \times n}$ for  $\tau = 1 \rightarrow \tau_{\max}$  do for t = 1, ..., T do set  $v_i(t) = \sum_{t_i=t-\Delta}^t x_i(t), \forall i$ set  $\hat{y}(t) = f(\sum_i W_i(t)v_i(t) - \theta)$ if  $\hat{y}(t) \neq y(t)$  then  $\Delta W = (\hat{y}(t) - y(t)) V_t^T$ Break end if end for Update:  $W(\tau + 1) = W(\tau) - \alpha_\tau \Delta W$ end for

With the above assumptions we can prove that NEUINF provides the desired ordering for the *average* values of different connection types, i.e., it returns higher values for excitatory connections than inhibitory or non-existent (void) connections. Note that we obtain this guarantee under the general (and somehow trivial) assumptions A1 and A2. Therefore, our result can be applied to a wide range of scenarios. We can have tighter guarantees (namely, with high probability, etc) if we make more strict assumptions about the type of the neurons and statistical properties for the firing patterns (as in the case of GLM) or the structure of underlying graph (e.g., random or small-world graphs).

**Theorem 1.** Under the assumptions A1 and A2, and for a sufficiently small learning step  $\alpha_{\tau}$ , NEUINF outputs

$$\bar{w}_{exc} \ge \bar{w}_{void} \ge \bar{w}_{inh},$$

where  $\bar{w}_{exc} \ \bar{w}_{void}$  and  $\bar{w}_{inh}$  denote the expected values returned by NEUINF for excitatory, void, and inhibitory connections, respectively. Here, the expectation is taken over the randomness of pre-synaptic spike patterns.

The proof of the theorem is provided in the longer version of this paper [22]. To give the reader an intuition, we use induction to show that the correct ordering for the *average value* of our beliefs about the type of synaptic connections is preserved in each iteration of the inference algorithm NEUINF.

## 5. NUMERICAL EXPERIMENTS

We generated several Erdos-Renyi *directed random graphs* with a connection probability p. In accordance with real biological networks, we assume that the probability of a connection being excitatory  $p_+$  is five times more than that of being inhibitory  $p_-$ . We set the firing threshold to  $\theta = 5$ mV and the maximum delay to  $d_{\text{max}} = 10$ ms. Further, we assume that neurons have a refractory period of 1ms and the membrane potential is reset after a spike. We also consider the



Fig. 2: Effect of recording duration T on beliefs returned by STOCHASTIC NEUINF for a recurrent network of 60 neurons.

"outside" traffic, which are the incoming spikes from neurons whose activities are not recorded. This traffic is modeled as a random process that sets each recorded neuron to fire with some probability q independent of the activity of other recorded neurons<sup>2</sup>. Once the data is generated, we compare the performance of NEUINF with Cross Correlogram (CC) and GLM. For GLM, we use the code provided in [11]<sup>3</sup> and modified it to serve our setup. We use the K-Means algorithm (with K = 3) to transform the analog association matrix into the digital adjacency matrix.

We evaluate the performance of the considered algorithms according to the following criteria: 1) The average values of the *association matrix* for the excitatory/void/inhibitory connections, 2) The *precision* and *recall* of the algorithm over the *ternary adjacency matrix*.

Numerical Results: We only report the results for the more general case of recurrent networks. Results for the feed-forward topology is available in the longer version [22]. Concretely, we consider a network of n = 60 neurons, 50 excitatory and 10 inhibitory, with a connection probability of 0.2.

Figure 2 illustrates the effect of the observation period T on the average value of beliefs returned by the considered algorithms. Desirably, STOCHASTIC NEUINF outputs an asso-



<sup>2</sup>The data/code is available at http://rr.epfl.ch/KSV2015 <sup>3</sup>The code is available at http://pillowlab.princeton.edu



Fig. 4: Performance comparison after "ternarification"

ciation matrix with the desired ordering, i.e. higher values for excitatory connections and lower values for inhibitory ones. Figure 3 shows the gap between the beliefs for each pair of connection types (i.e., between excitatory and "void" as well as "void" and inhibitory). Clearly, STOCHASTIC NEUINF outperforms the others.

The *precision* and *recall* of different approaches, after "ternarification" using the K-Means algorithm, is shown in Figure 4. In identifying excitatory and "void" connections, STOCHASTIC NEUINF outperforms both CC and GLM. Further, the seemingly good performance of CC in identifying inhibitory connections comes at the expense of low precision.



Fig. 5: The gap between beliefs of STOCHASTIC NEUINF

We also observe that sparsity, both in the firing patterns and network topology, improves the performance. Figure 5 illustrates the performance of STOCHASTIC NEUINF in differentiating connection types in feed-forward/recurrent networks for different values of connection probability p, and probability of being triggered by outside traffic, q.

#### 6. REFERENCES

- Duncan J Watts and Steven H Strogatz, "Collective dynamics of small-worldnetworks," *Nature*, vol. 393, no. 6684, pp. 440– 442, 1998.
- [2] Stephen M Plaza, Louis K Scheffer, and Dmitri B Chklovskii, "Toward large-scale connectome reconstructions," *Curr. Opin. Neurobiology*, vol. 25, pp. 201–210, 2014.
- [3] Ann-Shyn Chiang, Chih-Yung Lin, Chao-Chun Chuang, Hsiu-Ming Chang, Chang-Huain Hsieh, Chang-Wei Yeh, Chi-Tin Shih, Jian-Jheng Wu, Guo-Tzau Wang, Yung-Chang Chen, et al., "Three-dimensional reconstruction of brain-wide wiring networks in drosophila at single-cell resolution," *Current Biol*ogy, vol. 21, no. 1, pp. 1–11, 2011.
- [4] Rodrigo Perin, Thomas K Berger, and Henry Markram, "A synaptic organizing principle for cortical neuronal groups," *Proceedings of the National Academy of Sciences*, vol. 108, no. 13, pp. 5419–5424, 2011.
- [5] György Buzsáki, "Large-scale recording of neuronal ensembles," *Nature neuroscience*, vol. 7, no. 5, pp. 446–451, 2004.
- [6] Benjamin F Grewe, Dominik Langer, Hansjörg Kasper, Björn M Kampa, and Fritjof Helmchen, "High-speed in vivo calcium imaging reveals neuronal network activity with nearmillisecond precision," *Nature methods*, vol. 7, no. 5, pp. 399– 405, 2010.
- [7] 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*, vol. 7, no. 5, pp. 456–461, 2004.
- [8] Gaute T Einevoll, Felix Franke, Espen Hagen, Christophe Pouzat, and Kenneth D Harris, "Towards reliable spike-train recordings from thousands of neurons with multielectrodes," *Current opinion in neurobiology*, vol. 22, no. 1, pp. 11–17, 2012.
- [9] Wulfram Gerstner and Werner M Kistler, *Spiking neuron models: Single neurons, populations, plasticity*, Cambridge university press, 2002.
- [10] Rui Castro, Mark Coates, Gang Liang, Robert Nowak, and Bin Yu, "Network tomography: recent developments," *Statistical science*, pp. 499–517, 2004.
- [11] Jonathan W Pillow, Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M Litke, EJ Chichilnisky, and Eero P Simoncelli, "Spatio-temporal correlations and visual signalling in a complete neuronal population," *Nature*, vol. 454, no. 7207, pp. 995–999, 2008.
- [12] 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*, vol. 9, no. 7, pp. e1003138, 2013.
- [13] Alexandro D Ramirez and Liam Paninski, "Fast inference in generalized linear models via expected log-likelihoods," *Journal of computational neuroscience*, vol. 36, no. 2, pp. 215–234, 2014.

- [14] YuryV. Zaytsev, Abigail Morrison, and Moritz Deger, "Reconstruction of recurrent synaptic connectivity of thousands of neurons from simulated spiking activity," *Journal of Computational Neuroscience*, vol. 39, no. 1, pp. 77–103, 2015.
- [15] Ian H Stevenson, James M Rebesco, Nicholas G Hatsopoulos, Zach Haga, Lee E Miller, and Konrad P Körding, "Bayesian inference of functional connectivity and network structure from spikes," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 17, no. 3, pp. 203–213, 2009.
- [16] Frank Van Bussel, Birgit Kriener, and Marc Timme, "Inferring synaptic connectivity from spatio-temporal spike patterns," *Frontiers in computational neuroscience*, vol. 5, 2011.
- [17] Raoul-Martin Memmesheimer, Ran Rubin, Bence P Ölveczky, and Haim Sompolinsky, "Learning precisely timed spikes," *Neuron*, vol. 82, no. 4, pp. 925–938, 2014.
- [18] Carlo Baldassi, Alfredo Braunstein, Nicolas Brunel, and Riccardo Zecchina, "Efficient supervised learning in networks with binary synapses," *BMC Neuroscience*, vol. 8, no. Suppl 2, pp. S13, 2007.
- [19] Wolfgang Maass and Michael Schmitt, "On the complexity of learning from spiking neurons with temporal coding," *Information and Computation*, vol. 153, no. 1, pp. 26–46, 1999.
- [20] Henry Dale, "Pharmacology and nerve-endings," J. Royal Society of Medicine, vol. 28, no. 3, pp. 319–332, 1935.
- [21] Frank Rosenblatt, *The perceptron, a perceiving and recognizing automaton Project Para*, Cornell Aeronautical Laboratory, 1957.
- [22] Amin Krabsi, Amir Hesam Salavati, and Martin Vetterli, "Learning network structures from firing patterns," http: //http://rr.epfl.ch/paper/KSV2015, 2015.