Interaction graph estimation in a neural network

Brochini, Galves, Hodara, Ost, Pouzat

An interaction graph estimation procedure proposed by Duarte, Galves, Löcherbach and Ost (2016).

Application on real data and simulations with additional theoretical results. This requires a pre-treatment of the raw data called spike sorting.
$I$ is the countable set of neurons.
$I$ is the countable set of neurons. Time is indexed by $\mathbb{Z}$. 
The model

$I$ is the countable set of neurons.

Time is indexed by $\mathbb{Z}$.

$(X_t)_{t \in \mathbb{Z}}$ is a stochastic chain taking values in $\{0, 1\}^I$.

For each $i \in I$, and each $t \in \mathbb{Z}$, $X_t(i) = 1$ if neuron $i$ spiked at time $t$ and $X_t(i) = 0$ otherwise.
The model

$I$ is the countable set of neurons. Time is indexed by $\mathbb{Z}$.

$(X_t)_{t \in \mathbb{Z}}$ is a stochastic chain taking values in $\{0, 1\}^I$.

For each $i \in I$, and each $t \in \mathbb{Z}$, $X_t(i) = 1$ if neuron $i$ spiked at time $t$ and $X_t(i) = 0$ otherwise.

For all finite subset $J \subset I$,

$$P \left( X_t(j) = x_j, j \in J/X_{-\infty}^{t-1}(I) \right) = \prod_{j \in J} P \left( X_t(j) = x_j/X_{-\infty}^{t-1}(I) \right).$$
Let $L^i_t$ be the last spike time of neuron $i$ before time $t$ defined by:

$$L^i_t = \sup\{s < t : X_s(i) = 1\}.$$ 

$$P \left( X_t(i) = x_i / X_{t-1}(I) \right) = P \left( X_t(i) = x_i / X_{L^i_t t-1}(I) \right).$$
Let $L^i_t$ be the last spike time of neuron $i$ before time $t$ defined by:

$$L^i_t = \sup\{s < t : X_s(i) = 1\}.$$ 

$$P \left( X_t(i) = x_i/X_{t-1}(I) \right) = P \left( X_t(i) = x_i/X_{L^i_t}(I) \right).$$

Each neuron $i \in I$ possesses an interaction neighborhood $V_i \subset I$ satisfying for all $A \subset I$ with $V_i \subset A$,

$$P \left( X_t(i) = x_i/X_{L^i_t}(A) \right) = P \left( X_t(i) = x_i/X_{L^i_t}(V_i) \right).$$
**Figure:** Sample of $X$ process with size $n$
**Figure**: What is the probability of neuron $i$ to spike at time $t$?
**Figure:** Last time neuron $i$ spiked before $t$:

$L_t^i = \sup\{ s < t : X_s(i) = 1 \}$
Decision rule

**Figure:** The pattern $w$ is the portion of the past containing the information for the decision of the state of neuron $i$ at time $t$. 

<table>
<thead>
<tr>
<th></th>
<th>$l$</th>
<th>$k$</th>
<th>$j$</th>
<th>$i$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 1</td>
<td>0 1 0 0 1 0 0 0 0 1 0 0 1 0 0 1</td>
<td>0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0</td>
<td>0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 1</td>
</tr>
</tbody>
</table>

$w$ is the portion highlighted in blue. 

Neurons $i$, $j$, $k$, and $l$ are shown in the table with their corresponding states at different times.
Does $k$ influences $i$?

**Figure:** If $k$ influences $i$, then a modification of the pattern concerning only $k$ should lead to a different realisation for $i$ in the following time step.
Empirical probability

For a given neuron $i$ and a given pattern $w$, we denote by $N_i(w)$ the number of occurrences of $w$, and by $N_i(w, 1)$ the number of occurrences of $w$ followed by a spike of neuron $i$. For $w$ such that $N_i(w) > n^{\xi + \frac{1}{2}}$ for some parameter $\xi \in ]0, \frac{1}{2}[$, we compute

$$\hat{p}_{(i, n)}(1/w) = \frac{N_i(w, 1)}{N_i(w)}.$$
Does $k$ influences $i$?

Figure: We will compare the empirical probabilities for couples of patterns that are identical outside $k$.
The sensitivity parameter $\epsilon$.

Figure: If $|\hat{p}_{(i,n)}(1|w) - \hat{p}_{(i,n)}(1|v)| > \epsilon$ for any couple $(w, v)$ satisfying $w \setminus \{k\} = v \setminus \{k\}$ we accept $k$ in the estimated interaction neighborhood.
Results and discussion

- simulation to explore sensitivity and cutoff parameters
- procedure to deal with small sample sizes
- issue of partially observed networks
- experimental results
- limitations of the method
Exploring parameters $\epsilon$ and $\xi$ with simulations

Figure: Color code: white—correct absent, grey—correct present, red—false positive, blue—false negative. light — blue and light — red inconclusive (Network 1 with $\mu = 1$ and $n = 1e5$). Original Duarte et al estimator
Limitation: Small sample size and/or greater number of neurons: Too many inconclusives
Pruning procedure

• Limitation : Small sample size and/or greater number of neurons : Too many inconclusives
• Pruning procedure : re-estimate graph disconsidering neurons that the original estimator says is NOT presynaptic
• pruning procedure is due to the reduction in the number of presynaptic candidate neurons while maintaining the same sample size, leading to the improvement of the estimation quality.
• Analytical results : iterative pruning procedure conserves the consistency of the estimation
Pruning procedure

Before pruning

Postsynaptic

Presynaptic

After pruning

Postsynaptic

Presynaptic

Brochini, Galves, Hodara, Ost, Pouzat

Interaction graph estimation in a neural network
A key assumption in Duarte et al when determining the neighbourhood of neuron $i$ is that we have access to the activity of all neurons presynaptic to $i$. It is not realistic experimentally! What happens when this condition is not met?

Numerical experimentation: estimated the connection graph of subgraphs compared to true connections yielded:

- All connections identified as false were indeed false.
- All connections identified as true were either indeed true or a false positive due to a projected connection.

Analytical Results: guarantee that if there is NO PATH between neuron $j$ and $i$ involving an unobserved neuron, then the estimator is not expected to produce a connection.
A key assumption in Duarte et al when determining the neighbourhood of neuron \( i \) is that we have access to the activity of all neurons presynaptic to \( i \)

Not realistic experimentally!

What happens when this condition is not met?

Numerical experimentation: estimated the connection graph of subgraphs compared to true connections yielded:
- All connections identified as false were indeed false
- All connections identified as true were either indeed true or a false positive due to a projected connection

Analytical Results: guarantee that if there is NO PATH between neuron \( j \) and \( i \) involving an unobserved neuron, then the estimator is not expected to produce a connection.
**Figure:** Network 2. Complete graph recovered by procedure that identifies false positives due to projections. This procedure can be used to deal with large N.
Binning: Spike train $\rightarrow$ symbolic sequence of 0’s and 1’s: divide sample in small and attribute 1 when a spike occurs inside that window.

- Cannot be too small: not enough repetitions of patterns
- Cannot be too large: two spikes of the same neuron in the same window
- Choose the smallest possible window allowing at most 1% superpositions
Results on real data

Figure: Estimated connection graph for spontaneous activity of projection neurons of the antennal lobe of *Schistocerca americana*. 
Method Limitations

- Results concerning partially observed networks require stationary.
- Requires huge amount of pattern repetitions: Strong connections and sparse activity is a bad combo!
Estimation of neuronal interaction graph from spike train data
Ludmila Brochini, Antonio Galves, Pierre Hodara, Guilherme Ost, Christophe Pouzat
https://arxiv.org/abs/1612.05226
Thank You!