

Topological Data Analysis in Neural Science

Aljosa Marjanovic

May 11, 2020

Topological exploration of artificial neuronal network dynamics

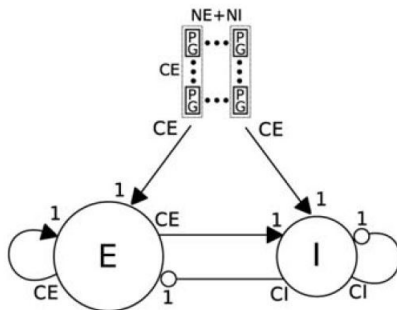
Jean-Baptiste Bardin, Gard Spreemann, and Kathryn Hess

Laboratory for Topology and Neuroscience, Brain Mind Institute, École
Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Table of Contents

- 1 The Brunel Model
- 2 The Simulations and the Data
- 3 Data Analysis

1. The Brunel Model



The potential on the membrane of a single neuron is governed by the equation:

$$\tau V_i'(t) = -V_i(t) + RI_i(t)$$

where

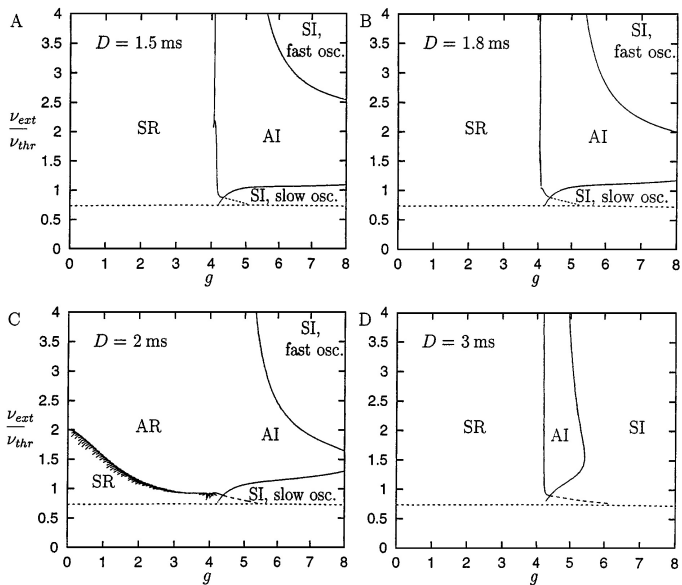
$$RI_i(t) = \tau \sum_j J_{ij} \sum_k \delta(t - t_j^k - D)$$

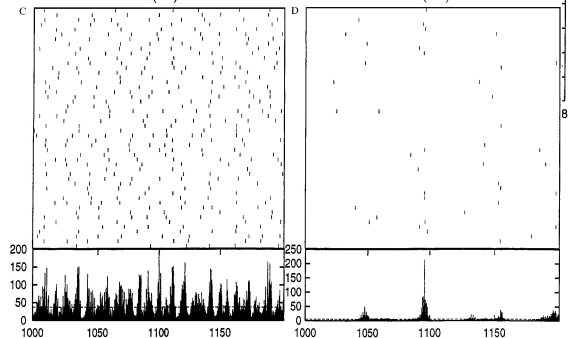
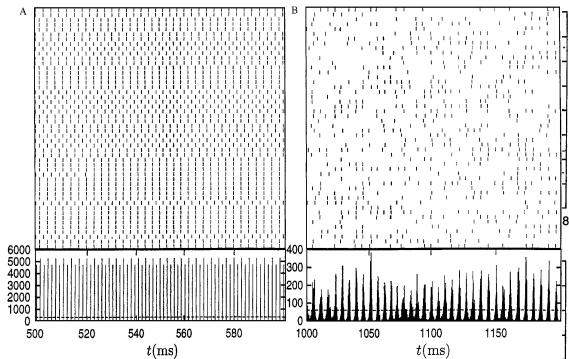
Amplitudes J_{ij} are all either equal J for excitatory or $-gJ$ ($g > 1$) for inhibitory neurons.

| Symbol | Description |
|-------------------------------|----------------------------|
| N_E | # excitatory neurons |
| N_I | # inhibitory neurons |
| C_E | # exc. synapses per neuron |
| C_I | # inh. synapses per neuron |
| J | Synaptic strength |
| D | Synaptic delay |
| g | Rel. synaptic efficiency |
| $\nu_{\text{ext}}/\nu_\theta$ | Rel. external rate |

The network has the following range of different states:

- Synchronious regular (SR) states, in which groups of neurons are synchronized and behave as oscillators when excitation dominates inhibition and synaptic time distributions are sharply peaked
- Asynchronious regular (AR) states, with stationary global activity and regular individual neuron firing, arises when excitation dominates inhibition and synaptic time distribution is widely peaked
- Asynchronious irregular (AI) states, with stationary global activity but strongly irregular individual firing of neurons, when inhibition dominates excitation
- Synchronious irregular (SI) states, with oscillatory global activity but strongly irregular individual firing at low firing rates





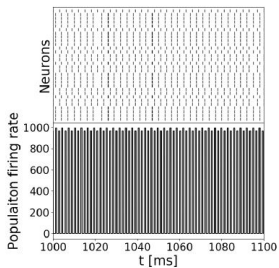
2. The Simulations and the Data

- Authors wanted to analyze a network with lower number of neurons as to reduce the computational complexity, so they picked Brunel model and lowered the number of neurons at $N = 2500$
- This however, proved to be problematic...

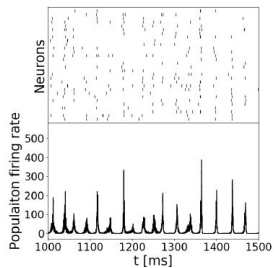
The versions of the network that were simulated are:

- Version 1: Sparsely connected ($P := \frac{C}{N} = 10\%$) and fast synaptic transmission ($D = 1,5ms$, $J = 0,1mV$)
- Version 2: Denser connectivity ($P = 40\%$) and fast and strong synaptic transmission ($D = 1,5ms$, $J = 0,2mV$)
- Version 3: Denser connectivity ($P = 40\%$) and slow and strong synaptic transmission ($D = 3ms$, $J = 0,2mV$)

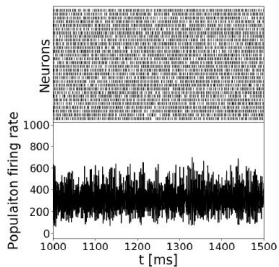
A



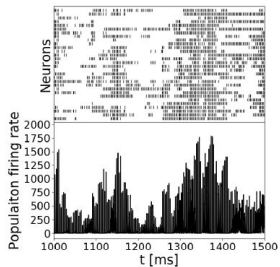
B



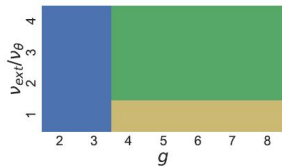
C



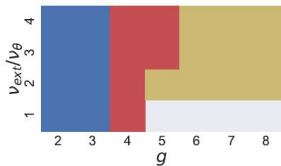
D



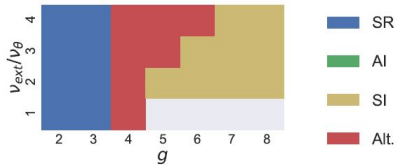
A



B



C



3. Data Analysis

The spike train data for all neurons of a network is given. We want to discriminate between the regimes using TDA. One wants to build appropriate simplicial complexes.

To the end of building weighted graphs, three distance measures were used:

- Pearson correlation
- Spike-synchronicity
- Spike-distance

For two neurons let their spike trains be given by $\{t_i^{(1)}\}$, $i = 1, \dots, M_1$ and $\{t_j^{(2)}\}$, $j = 1, \dots, M_2$, where $t_i^{(1)}$ and $t_j^{(2)}$ are spike moments for the two neurons and M_n the number of spikes for neurons $n = 1, 2$. From the two spike trains we derive spike count vectors a and b having n components, each k -th being the number of spikes in k -th bin. The Pearson correlation is then given by:

$$r(a, b) = \frac{\sum_{k=1}^n (a_k - \bar{a})(b_k - \bar{b})}{\sqrt{\sum_{k=1}^n (a_k - \bar{a})^2} \sqrt{\sum_{k=1}^n (b_k - \bar{b})^2}}$$

Define the distance between a and b as $d := 1 - r(a, b)$

Spike-Synchronicity

Define a coincidence window

$\tau_{ij}^{(1,2)} := \min\{t_{i+1}^{(1)} - t_i^{(1)}, t_i^{(1)} - t_{i-1}^{(1)}, t_{j+1}^{(2)} - t_j^{(2)}, t_j^{(2)} - t_{j-1}^{(2)}\}/2$ and coincidence counter:

$$C_i^{(1)} := \begin{cases} 1, & \min_j |t_i^{(1)} - t_j^{(2)}| < \tau_{ij}^{(2)} \text{ (for a } j) \\ 0, & \text{otherwise} \end{cases}$$

Then count and average all the incidences of synchronous spiking:

$$S_c := \frac{1}{M} \sum_{k=1}^M C(k)$$

where $M = M_1 + M_2$ and $C(k)$ is the enumeration of $C_i^{(1)}, C_j^{(2)}$. Take $d := 1 - S_c$ as distance.

Spike-Distance is given by:

$$D_S := \frac{1}{T} \int_0^T S(t) dt$$

where

$$S(t) := \frac{S_1(t)x_{ISI}^{(2)}(t) + S_2(t)x_{ISI}^{(1)}(t)}{2(\overline{x_{ISI}})^2}$$

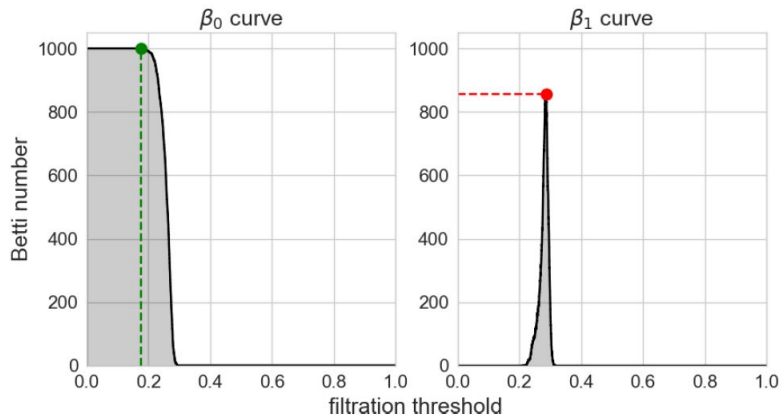
for

$$S_i(t) := \frac{\Delta t_P^{(i)}(t)x_F^{(i)}(t) + \Delta t_F^{(i)}(t)x_P^{(i)}(t)}{x_{ISI}^{(1)}(t)}$$

Simplicial Complexes

- Select a distance measure.
- For a chosen distance measure d the edge between vertices i and j is given the weight $w(i, j) := d(i, j)$
- The 2-simplices are given weight $w(i, j, k) := \max\{w(i, j), w(j, k), w(i, k)\}$.
- All the vertices, edges and 2-simplices build simplicial complex K .
- A filtration of this complex for a threshold ϵ we get by including only the simplices with weight less than ϵ . Thus we get the filtration $V \subset K_1 \subset \dots \subset K$
- The zeroth and first Betti-number of filtrations are calculated and recorded
- Betti-curves are plotted (i.e. dependance of the Betti-numbers on ϵ)

Examples of Betti-Curves and characteristic numbers



The following numbers are derived from the Betti-curves:

- The area under the Betti-0 curve
- The threshold at which Betti-0 curve starts to decrease
- The area under the Betti-1 curve
- The maximum of Betti-1 curve

Thus for every set of spike train data and every one of distance measures these four numbers are derived, totaling in 12 numbers per simulation.

The next step of their procedure was to use a machine learning method called the support vector machine (SVM) classifier to distinguish between the regimes using only these characteristic numbers.

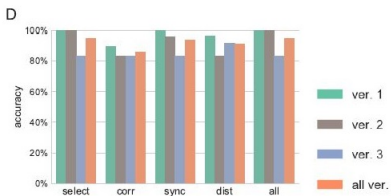
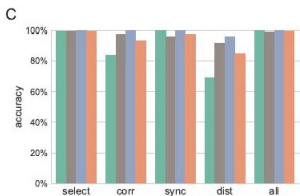
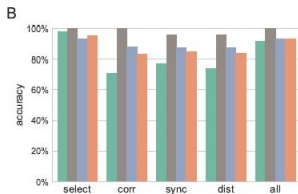
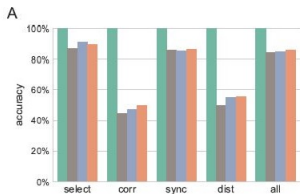
To prevent the overlapping of the clusters, out of these 12 numbers, 6 are optically selected for machine learning method:

- The area under Betti-0 for all measures (3 numbers)
- The area under Betti-1 for correlation and spike synchronicity (2)
- The maximum for Betti-1 for spike distance

The training sets were divided into four categories, composing of samples from the simulations of versions 1-3 and the last one composed of samples of all versions. The classifier for version k was then tested on the samples either from the version k unused for the training, or on all the valid samples from other versions (with regimes occurring in version k)

The Results of the ML Classifier

| Training data | Test data | | | |
|---------------|-----------|-----------|-----------|--------|
| | Version 1 | Version 2 | Version 3 | All |
| Version 1 | 100% | 86,67% | 91,18% | 89,68% |
| Version 2 | 97,69% | 100% | 93,33% | 95,18% |
| Version 3 | 99,23% | 99,17% | 100% | 99,23% |
| All | 100% | 100% | 100% | 100% |



Thanks for listening!