# Topological Data Analysis in Neural Science

Aljosa Marjanovic

17. May 2020

## 1 Introduction and Motivation

This essay has the aim of presenting the article [BSH19] and some background material, that is, to better connect the article [BSH19] to the previous work it was based upon. The article itself has the purpose of demonstrating a possible way of applying topological data analysis (TDA) on the data which arises in the context of modeling and simulating the biological neural networks, or alternatively, their measurements. It is a part of a line of articles following a similar theme and background (differing of course in the case study, the data and the exact methods). A method developed and described by the authors of [BSH19], based on persistent homology, was applied to analyze a known and well-researched classical model (Brunel model) with some simplifications which do not diminish the main purpose of the article: to show the usefulness of this and similar methods and motivate their use in more complex and realistic cases.
Firstly, we introduce the notion of a model of a neural network and its dynamics as it is understood in the field of neural science, as to define the goals for which the TDA is of help. Neural science describes and explains the functioning of large networks of neurons in nature, in order to explain how they process information and perform calculations, i.e. react to input. The manifold of inner states of a network corresponds to its functioning. Even if it is possible to know the state of every single neuron at any given moment, the function and output of the network are the result of its global state. The same global state may have many realizations for different states of singular neurons, which are not (as) important for describing neurological function. This global state is referred to as the dynamics of a neural network. On the other hand, these dynamics arise from their functional base, a single neuron with its signal propagation and its dynamics, usually modeled by systems of differential equations depending on the parameters of the model (concentrations of neurotransmitters, states on the synapses etc). One of the main goals of neuroscience is to examine the global dynamics and its effects. Some simple practical examples include the differences in brain activity between the different stages of sleep or between the normal state of the brain and an epileptic episode.
A proper model describes the network architecture, identifying different layers, regions or sub-populations of neurons (i.e. neurons with distinct behavior), moreover it has to describe the network connectivity, which connections are allowed and how they are created etc. Lastly, it has to explain the behavior of the neurons and their synapses, giving equations that govern the spike generation. A model also has to define the input (stimuli) to the network as well as the data recorded (output) so that it can be compared to the reality (which it ought to model) and be validated. Attempts have been made to standardize representation of models, making them easier to reproduce, verify or falsify [NGP09].

The usual methods employed in order to analyze these models are naturally graph theory, theory of dynamical systems, statistics and statistical mechanics. They, however, sometimes fall short of classifying different states of a network. For example, the averaged quantities of statistical mechanics sometimes do not differentiate between various stable states of the network.
For the sake of demonstrating the TDA method, a Brunel network has been simulated and spike train data produced. The choice of the model is natural, since it is very simple, established as a basic example, precisely defined and reproduced and lastly, despite its simplicity, possesses a wide range of dynamics. The data consists of spiking sequences, i.e. precise moments of spiking for every single neuron in the network. The idea is to embed neurons into a space in which they are closer together the more "similar" their spike trains are. To measure this similarity and distances between neurons three distinct measures will be discussed. Simplicial complexes are constructed for which the Betti-numbers are calculated. This is the core of the method. At this juncture, a statement could be postulated: distinct Betti-curves, derived from spike data, correspond to distinct regimes of the neural network and can be used to classify and recognize the states. The topological properties of the network in an abstract space change with the functioning of the network.
In the article, however, as a further step, distill information out of the Betti-curves and reduce them to a couple of characteristic numbers. Using methods of machine learning, they demonstrate that one can predict the state of a network with high accuracy by only observing these numbers.
The outline of this text is as follows. In the next section, we introduce the Brunel model. The third section describes the simulations and the data to be analyzed from [BSH19]. The Method of data analysis is then described in the fourth and a conclusion follows as well as the relevant references.

## 2 Brunel Model and its Properties

The Brunel model is the so-called leaky integrate and fire (LIF) model of neural networks. The model regards a neuron as either incitory or dampening for a signal (arriving at its synapses). A neuron has a membrane potential $V_i(t)$ which is modeled by the differential equation:

$$\tau V_i'(t) = -V_i(t) + RI_i(t) \tag{1}$$

where $I_i(t)$ are synaptic currents given by

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

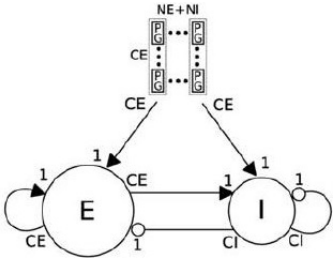

Figure 1: CE and CI are the number of connections

the negative term being the “leak“. Here we sum all of the currents coming to the cell $j$, where $J_{ij}$ are the amplitudes (or maximal efficiency) of the

signal. The second is a sum of different spikes arriving at synapse $j$ given by a $\delta$-function and with delayed moments of spiking $t_j^k + D$, where $D$ is the delay. The $\delta$-function gains a meaning here only after an appropriate integration over it, idea being that the signals are short pulses in infinitely short periods of time. In this way the membrane potential jumps for an appropriate quantity at time $t - t_j^k - D = 0$. $R$ is the cell membrane resistance and $\tau$ is the integration time.
If $V_i$ at some point reaches the threshold (activation) potential $V_\theta$, the neuron fires, i.e. a signal (spike) is sent to all of the neurons synapses, which changes the potential of its post-synaptic neurons and at the same time resets its own potential to $V_r$. The neuron stays inactive for a while.
A possible symbolical representation of a Brunel network is displayed in Figure 1, which is taken from [NGP09]. The network itself consists of $N$ neurons divided, as mentioned before, into the two populations: excitatory (E) and inhibitory (I). In vivo observation of neocortical and hippocampial networks inspired the ratio $N_E = 4N_I$ of the sub-populations (i.e. 80% of neurons are excitatory). Each neuron receives exactly the same number of randomly chosen connections from other neurons, $K_E = PN_E$ being from excitatory and $K_I = PN_I$ from inhibitory neurons. The connectivity parameter $P$ [1] should be rather small ($P << 1$), resulting in a sparsely connected network. There are also $K_{\text{ext}}$ outside connections, the input to the network. These inputs fire randomly, given by a Poisson process. That is, for an outside connection, the spike times $t_j^k$ are governed by a Poisson point process on the real line with parameter $\nu_{\text{ext}}$, which is a model parameter of interest, being compared to $\nu_\theta = \frac{V_\theta}{JK_E\tau}$. $\nu_\theta$ is the minimal external rate necessary to trigger a spike within recurrent connections (i.e. wake the network up, activate it). For the simplicity of the model, it is assumed that $J_{ij}$ are all equal $J > 0$ for excitatory external synapses as well as for excitatory recurrent synapses. For inhibitory synapses they are $-gJ < 0$, where $g > 1$ is another free parameter of interest.
In the article [Bru00], two variants of the model were proposed. In the first one, inhibitory and excitatory neurons have the same characteristics ($\tau$ is the same, efficiencies are either $J$ or $-gJ$ on all synapses etc). The second one takes the differences in the propagation of the signal between different connections into account (inhibitory/inhibitory, inhibitory/excitatory etc). We consider the simpler model, which nevertheless has a large parameter space consisting of $N$ ($N_E$, $N_I$), $P$ ($K_E$, $K_I$), $\tau$, $V_\theta$, $V_r$, $D$, $J$, $\nu_{\text{ext}}$ and $g$.
The original Brunel model has been studied analytically, through numerical integration (of the system of partial differential equations) as well as through numerical simulations of the model. The following regimes have been observed:

- Syncronious regular (SR) states, in which groups of neurons are synchronized and behave as oscillators, if excitation dominates inhibition and synaptic time distributions are sharply peaked
- Asynchronious regular (AR) states, with stationary global activity and regular individual neurons firing, arises when excitation dominates inhibition and synaptic

[1] The inverse $\frac{1}{P}$ is also called the relative number of connections

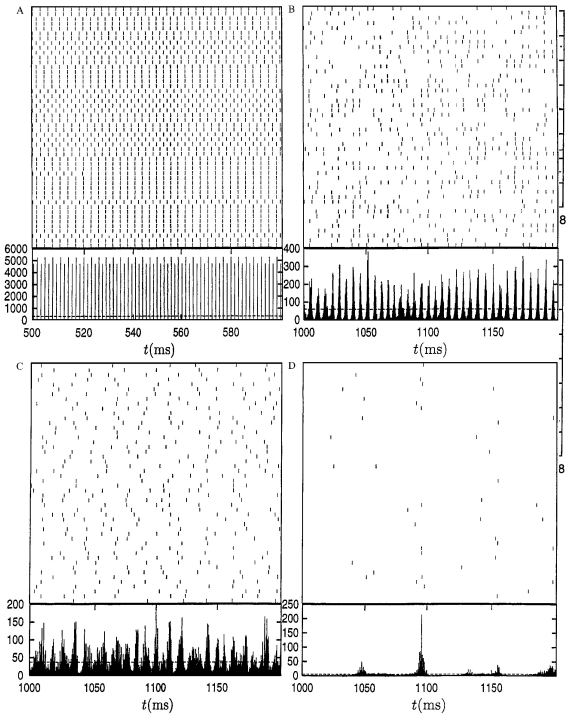


Figure 2: Examples of the simulation data for some regimes of the network, taken from [Bru00]. In the upper part of a single graph, spike trains of 50 randomly chosen neurons are shown, in the lower the global activity, given as the sum of all firings in a time bin. A is SR, B is SI (fast), C is AI, D is SI (slow)

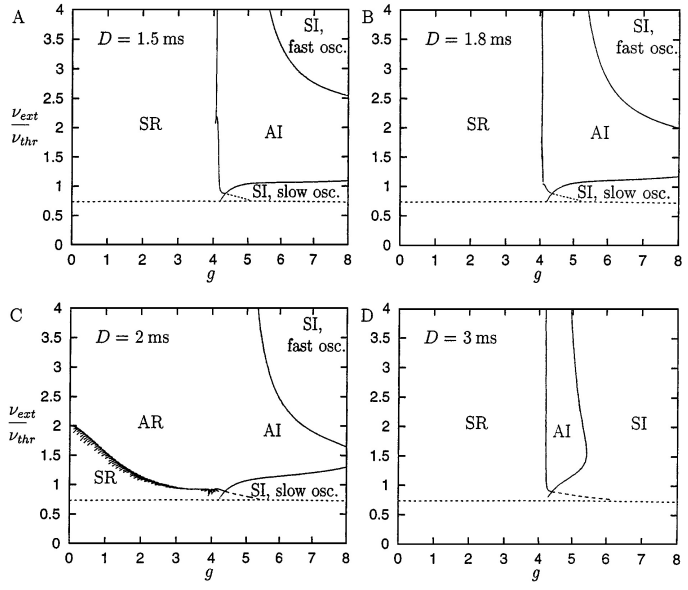


Figure 3: Bifurcation diagrams for some specific values of time delay, taken from [Bru00]

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

Examples of states obtained by a simulation are given in Figure 2. The regions of these regimes in the parameter space have been determined and given illustratively for some cases in the bifurcation graphs in Figure 3.

## 3 The Simulations

The authors of [BSH19] decided to downscale the simpler version of the Brunel model from the original $N = 12500$ neurons to $N = 2500$ out of technical considerations regarding the calculations of homologies for large graphs. They have, however, noticed that lowering the number of neurons while keeping the connectivity parameter $P = \frac{K}{N}$ constant will result in higher correlation between the neurons (meaning they are likelier to spike simultaneously with little time difference) and subsequently in the loss of diversity of the network dynamics. Thus, in order to keep the spectrum of dynamics of the network closer to the original presented in the previous section, the authors decided to tweak the rest of the parameters. They increased $J$ while keeping $JK$ constant and they increased the rate of the external population $\nu_{\text{ext}}$. Further, an additional inhibitory external population has been introduced as to preserve the mean and variance of the external population, all of which has the goal of reducing the correlation of neurons in the network. In order to cover as many regimes as possible from the original model, they came up with three distinct versions of the network:

- Version 1: Sparsely connected ($P = 10\%$) 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 ($D = 3ms,\ J = 0,2mV$)

They simulate each network for a period of 20 seconds for a given set of free parameters $g$ and $\nu_{\text{ext}}/\nu_\theta$. Parameter $g$ and $\nu_{\text{ext}}/\nu_\theta$ take discrete values 2-8 and 1-4, respectively, so that each version of the network has to be tested for 28 pairs of parameters. Since network connections are randomly initialized every time, they repeated the simulation for each version and each pair of parameters 10 times, totaling 280 simulations for every

version. Spike trains are recorded for every single neuron, so that the data set of a simulation has a form similar to an EEG, as if every neuron were to have an interface.
The Brunel network is allegedly robust, in the sense that the randomness of connections

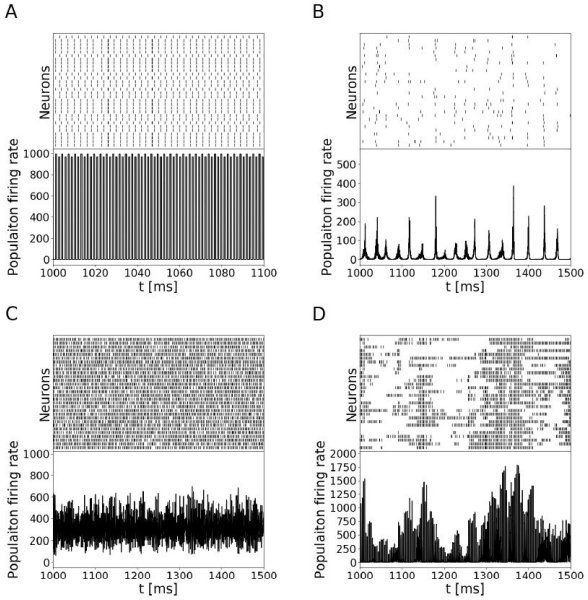


Figure 4: Four regimes of the downscaled network, taken from [BSH19]. The upper part of a graph is a raster of 30 randomly chosen spike trains, the lower part is the global activity. A is SR, B is SI, C is AI and D is Alt

and of the input shouldn't influence the regime to which the network is to switch for the given parameters. The authors, however, report that 40 simulations resulted in no neurons firing, which was not expected. Those simulations were not considered.
They observed four different regimes in the downscaled tweaked network. These are SR, SI and AI of the original model and Alt, which the original does not possess. In the Alt-regime neurons alternate between inactive state for a longer period of time and short rapid firing. For a qualitative impression and a comparison to the original, the simulation data examples for the regimes of the downscaled network are given in Figure 4. Regions of different regimes in the parameter space are given in Figure 5 and they remind us of the bifurcation diagrams of the original.

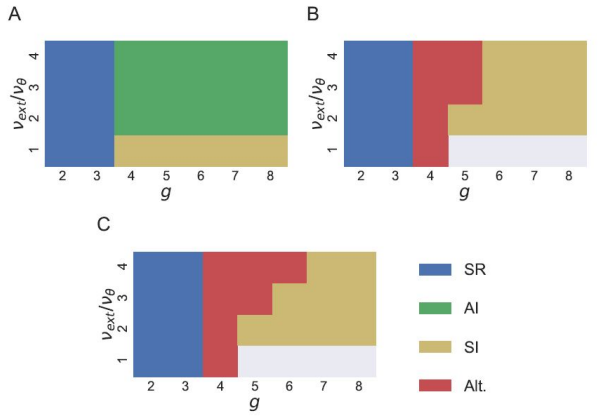


Figure 5: Diagrams of the parameter space for versions 1, 2 and 3, respectively, taken from [BSH19]. For the white areas the data is either not available or the network did not activate.

## 4 The Method

The first step of the method for analyzing spike train data presented in [BSH19] is to create simplicial complexes out of the simulated data. The relative distances of neurons between each two neurons have to be calculated using an adequately chosen "metric". They propose three: the Pearson correlation, spike synchronicity and spike distance, presented for example in [KMB15], as follows.

For two neurons let their spike trains be given by $\{t_i^{(1)}\}$, $i = 1, ..., M_1$ and $\{t_j^{(2)}\}$, $j = 1, ..., M_2$, where $t_i^{(1)}$ and $t_j^{(2)}$ are spike moments for the two neurons and $M_n$ the number of spikes for neurons $n = 1, 2$ (Figure 6).

We first define the Pearson correlation distance measure. Let the whole measurement interval of spike trains be divided into $n$ bins, the interval length of a bin being $\frac{T}{n}$. From the two spike trains we derive spike count vectors $a$ and $b$ having $n$ components, each $k$-th component being the number of spikes in the $k$-th bin. The Pearson correlation is then given by:

$$r(a,b) = \frac{\sum_{k=1}^{n}(a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\sum_{k=1}^{n}(a_i - \overline{a})^2}\sqrt{\sum_{k=1}^{n}(b_i - \overline{b})^2}}$$

$\overline{a}$ and $\overline{b}$ being averages. The time binning interval they have chosen for calculating the Pearson correlation was 2 ms.

The idea behind spike-synchronicity is to count (and average) the simultaneous appearances of spikes in the two spike trains. For a definition of what locally counts as simultaneous, we introduce a coincidence window $\tau_{ij}^{(1,2)} := \min\{t_{i+1}^{(1)} - t_i^{(1)}, t_i^{(1)} - t_{i-1}^{(1)}, t_{j+1}^{(2)} -$

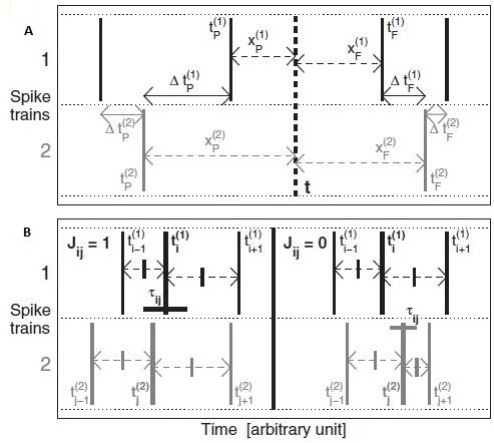


Figure 6: Examples of (small portions of) spike trains and clarification of different intervals, taken from [KMB15]. A for spike distance, B for spike synchronicity

$t_j^{(2)}, t_j^{(2)} - t_{j-1}^{(2)}\}/2$. The coincidence indicator is defined as

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

(analogously for $C_j^{(2)}$), which returns 1 if $i$-th spike is considered coincidental to some $j$ by this definition. We define lastly $S_c := \frac{1}{M} \sum_{k=1}^{M} C(k)$, where $M = M_1 + M_2$ is the total number of spikes in both spike trains and $C(k)$ is the enumeration of $C_i^{(1)}$, $C_j^{(2)}$. $S_C$ is called spike synchronization and is an intuitive measure of similarity between two spike trains. We notice that both the Pearson correlation as well as spike synchronicity are bounded in $[0, 1]$ with them being 1 for identical spike trains. In order to get a viable metric, we define $D_p := 1 - r_{ij}$ and $D_C := 1 - S_C$.
The third measure is called spike distance. For a given moment $t$ and each of the two spike trains we denote $t_P^{(1)}$, $t_P^{(2)}$, $t_F^{(1)}$ and $t_F^{(2)}$ as the first spike previous to and following the moment $t$. Then we define the difference to the nearest spike of the other spike train $\Delta t_P^{(1)}(t) := \min_i |t_P^{(1)}(t) - t_i^{(2)}|$ and analogously $\Delta t_P^{(2)}$, $\Delta t_F^{(1)}$, $\Delta t_F^{(2)}$. We further define the differences between the moment $t$ and first previous and analogously, first next spike, $x_P^{(n)}(t) := t - t_P^{(n)}(t)$ and $x_F^{(n)}(t) := t_F^{(n)}(t) - t$, $n = 1, 2$, as well as the distance between two spikes containing $t$: $x_{ISI}^{(n)} := t_F^{(n)}(t) - t_P^{(n)}(t)$ (so called interspike intervals). The graphical example is given in Figure 6. Then we define the "dissimilarity measure"

$$S_1(t) := \frac{\Delta t_P^{(1)}(t) x_F^{(1)}(t) + \Delta t_F^{(1)}(t) x_P^{(1)}(t)}{x_{ISI}^{(1)}(t)}$$

using $x_F$, $x_P$ as weights, normalizing with $x_{ISI}$ and analogously define $S_2$ for the other spike train. Lastly, we define the "dissimilarity profile" as the weighted average of both

$S_1$ and $S_2$, normalized through the average interspike distance for both spike trains:

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

We average once again over the whole time interval of the spike train measurements $T$, defining the spike train distance:

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

$S$ is bounded by $[0, 1]$, therefore, so is $D_S$, being zero only for identical spike trains.
The motivation for having different measures is not immediately clear. Each one of them omits an aspect of functional similarity between neurons. For example, two neurons firing consistently one after another, but rarely and with longer time differences would have Pearson correlation zero and therefore maximal distance. Furthermore, since the distances between two neurons calculated in different ways, seen as random variables, are apparently uncorrelated and carry independent information, the authors decided not only to compare them but to use them simultaneously, essentially handling the data set in three different spaces.
We only have relative distances of neurons and can not embed them directly into a higher dimensional space. Using any of the measures, for a filtration threshold $\epsilon$ one gets a simplicial complex as follows. Enumerated neurons are the vertices. For a chosen metric the edge between vertices $i$ and $j$ is given the weight $w(i, j) = d(i, j)$, so that every 1-simplex has a value equaling the distance between the neurons. 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 a simplicial complex $K$. We get a filtration of this complex by the threshold $\epsilon$ by including only the simplices with weight less than $\epsilon$. From the filtrations only the number of connected components and 1-holes are recorded, i.e. Betti-numbers $\beta_0$ and $\beta_1$ in dependence of $\epsilon$. Examples of Betti-curves for spike data are given in Figure 7.
The authors then extract the following numbers from the zeroth and the first Betti-curves for a single filtration: the area under the Betti-0 curve and the threshold at which it starts to decrease. From Betti-1 curve we take the area and the maximum. This was repeated for all the distance measures, so that for every simulation there are (as distilled end-result) 12 numbers.
The next steps are somewhat sketchy and experimental. The idea is to show that it is possible 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: area under Betti-0 for all measures, area under Betti-1 for correlation and spike synchronicity and the maximum for Betti-1 for spike distance. They seek to maximize the mutual information score (exclude the numbers, which do not carry additional information). The choice is also justified from experience and the fact that the classifier worked very well.

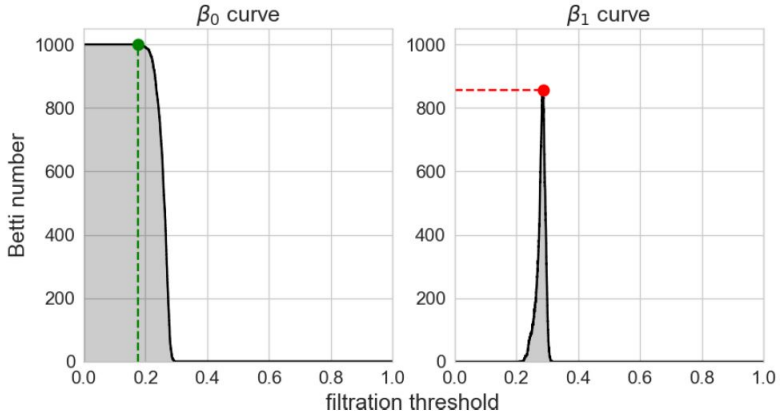


Figure 7: Examples of Betti-curves and characteristic numbers for the data of a simulation

| | Test data | | | |
|---|---|---|---|---|
| Training 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% |

Table 1: Classification accuracy over different training and testing data

Without going into details of the classifier, 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 performance of classifiers is given in the contingency Table 1. The percentage of correctly classified test-cases for the classifiers trained on data-sets from different versions is given. They also compare the use of classifiers with only one measure involved with the use of a combination of measures (selection of features of Betti-curves under different measures) and consistently get better results with a combination.

## 5 Conclusion

In the article [BSH19] the authors have shown that it is possible to recognize the regimes of the Brunel neural network using some topological features. Obvious directions for further research include more complicated models, but also different data, for example

the whole population rates. Secondly, they suggest that it is possible to consider other topological features, not necessarily more complicated (than Betti-numbers used), such as total count of simplices of different dimensions, but assume that this would not carry enough information about the regimes. Lastly, some practical applications may include analysis of fMRI and EEG data, for example for automated seizure detection.

## References


[Bru00] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Physiology Paris*, 2000.

[BSH19] Jean-Baptiste Bardin, Gard Spreemann, and Kathryn Hess. Topological exploration of artificial neuronal network dynamics. *Network Neuroscience*, 2019.

[KMB15] T. Kreuz, M. Mulansky, and N. Bozanic. Spiky: a graphical user interface for monitoring spike train synchrony. *Journal of Neurophysiology*, 2015.

[NGP09] Eilen Nordlie, Marc-Oliver Gewaltig, and Hans Ekkehard Plesser. Towards reproducible descriptions of neuronal network models. *PLoS Computational Biology*, 2009.