Neighborhoods for classifying binary dynamics in directed networks with machine learning

LU Interdisciplinary Applied Mathematics seminar

5 February 2021

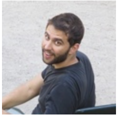
Jānis Lazovskis

Dejan Govc
University of Ljubljana

Jānis Lazovskis
University of Latvia

Ran Levi
University of Aberdeen

Henri Riihimäki
University of Aberdeen



Pedro Rodrigues da Conceição
University of Aberdeen

Jason Smith
Nottingham Trent University

Slides at: `jlazovskis.com/talks`

Overview

About.

- Based on work in preparation (2019-2021)
- In collaboration with the Blue Brain Project (EPFL)
- Funded by EPSRC grant "Topological Analysis of Neural Systems"

Goals.

- Neurological: Make a *in silico* model of a brain based on *in vivo* models.
- Mathematical: Distinguish neurological activity by its topological features.

Plan.

1. Neuroscience
   - Structure of the network
   - Experiments on the network
2. Mathematics
   - Topology in neuroscience
   - Neighborhoods in a digraph
3. Implementation and results
   - Classification by machine learning
   - Computational requirements

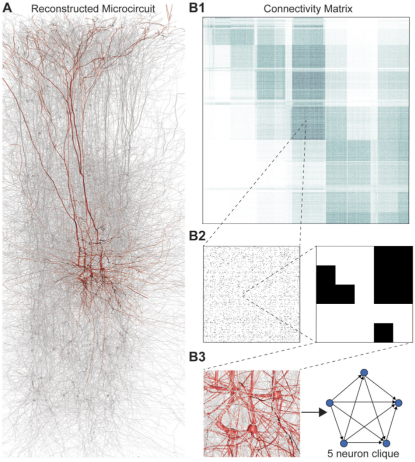
Neuroscience: Structure

Microcircuit of a rat somatosensory column, average over 6 instances:

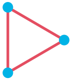
	neuroscience	*mathematics*
31 346	neurons	vertices
7 803 528	synapses	directed edges

This graph is:

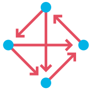
- Biologically modeled ("grown")
- Very sparse (0.8% density)
- Physically small (0.29mm^3 of brain)
- Relatively small (newest version has 4.2 million neurons and 4.8 billion synapses)
- Open source (available at `bbp.epfl.ch/nmc-portal`)
- Not a random graph (more topological features)



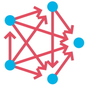
The complexity of a network can be considered via the number of *cliques*.



undirected 3-clique



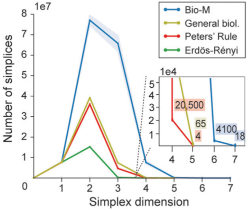
directed unordered 4-clique

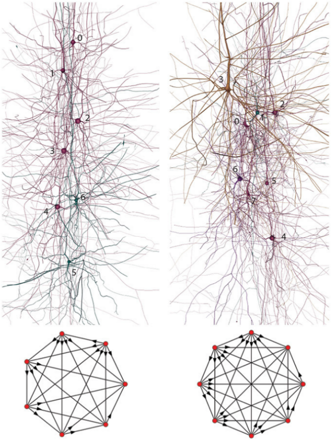


directed ordered 5-clique

Large cliques unlikely, and among those unordered cliques dominate. On n vertices:

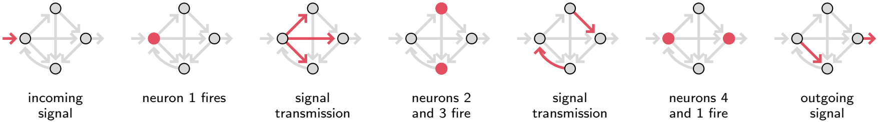
- $3^{n(n-1)/2}$ possible configurations
- $2^{n(n-1)/2}$ possible configurations as directed n-cliques
- $n!$ possible configurations as directed ordered n-cliques





Activity.

- Each neuron has an *electric potential* and a *threhsold*.
- If the potential passes the threshold, the neuron *fires* and sends its potential along outgoing synapses.
- Each synapse has a probability of signal *transmission*.
- Modern models encode *plasticity*, or the change in future transmission probability based on past transmission.

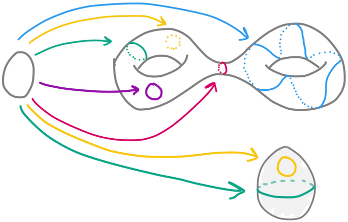


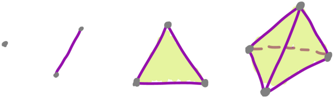
Experiments.

- External input is connected by *thalamic fibers* to receptory neurons.
- Different *stimuli* are sent to the brain (flick of a whisker)
- The propogation of activity is recorded over 250ms after the input of the signal
- 8 stimuli, 557 repetitions of each
- Data recorded as a list of values (n_i, t_i) of the neuron index and the time it fires

Mathematics: Topology

- The *topology* of a space X is reflected by classes of maps $S^n \to X$.
- The more (homotopy) classes in C_n, the more complex the space.
- The *flag complex* of a directed graph comes from associating to every directed n-clique an $(n-1)$-dimensional *simplex*.





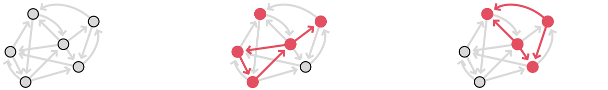
---

homology groups:	$H_n(X) = \ker(C_n \to C_{n-1})/\mathrm{im}(C_{n+1} \to C_n)$
Betti numbers:	$\beta_n(X) = \lvert H_n(X) \rvert$
Euler characteristic:	$\chi(X) = \beta_0 - \beta_1 + \beta_2 - \beta_3 + \cdots$
normalized Betti coefficient:	$\mathfrak{B}(X) = \dfrac{\beta_0(X)}{\#\text{ of vertices}} + \dfrac{2\beta_1(X)}{\#\text{ of edges}} + \cdots$

Mathematics: Neighborhoods in a digraph

Neuroscience uses the *firing rate* of a neuron or region for classification.
We use different parameters based on graph neighborhoods.

- Fundamental: firing rate, in degree, out degree, recioprocal connection count
- Algebraic: clustering coefficient
- Topological: Euler characteristic, Betti coefficient, density coefficient
- Spectral: adjacency, Laplacian, transition probability



structural | *active subgraph on* $[t_0, t_1]$ | *active subgraph on* $[t_1, t_2]$

In total 30 different parameters.

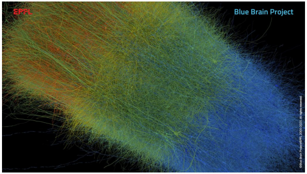
Implementation: Technical requirements

Neuroscience: EPFL

- Circuit built and experiments run on EPFL Blue Brain Project supercomputers
- 42k cores, 94TB of RAM
- Takes about 10 hours to simulate 250ms

Mathematics: University of Aberdeen

- Analysis run on UoA Maxwell HPC
- 1.2k cores, 12TB of RAM
- Takes about 2 hours to featurize each parameter, 1 minute to classify
  - Requests 40 cores, 150GB of RAM
  - Topological computations in parallel



End

References.

- Michael Riemann, Max Nolte, Martina Scolamiero, Katharine Truner, Rodrigo Perin, Giuseppe Chindemi, Pawel Dlotko, Ran Levi, Kathryn Hess and Henry Markram. *Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function,* 2017.
- Michael W. Reimann, Henri Riihimäki, Jason P. Smith, Jānis Lazovskis, Christoph Pokorny, Ran Levi. *Topology of synaptic connectivity constrains neuronal stimulus representation, predicting two complementary coding strategies,* 2020.
- Dejan Govc, Jānis Lazovskis, Ran Levi, Henri Riihimäki, Pedro Rodrigues da Conceição, Jason Smith. *An application of neighbourhoods in digraphs to the classification of binary dynamics,* in preparation.
- Fan Chung. *Laplacians and the Cheeger inequality for directed graphs,* 2005.

Acknowledgements.
The authors are partially funded by EP/P025072/1