Masters Thesis

Random Graph Models

of a neocortical column in a rat's brain and their topological statistical distributions

Kieran Barber

Advisor: Raazesh Sainudiin

UPPSALA
UNIVERSITET

- Blue Brain Project (BBP) and the brain
- Models
- Results

THE BLUE BRAIN PROJECT (BBP)

The mission for the BBP is to build a realistic model of the Human Brain

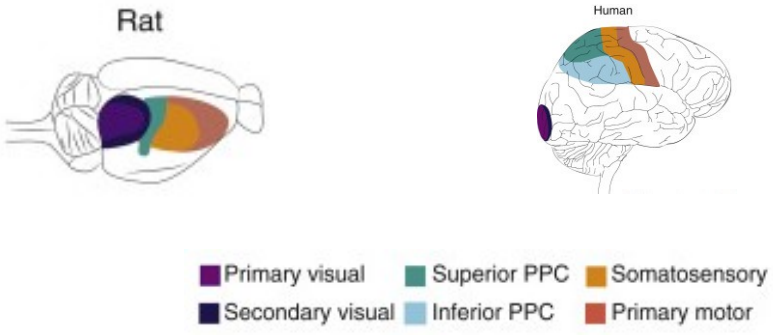


Figure: https://doi.org/10.1016/j.cub.2017.06.007

UPPSALA UNIVERSITET

DEVELOPMENT OF THE BRAIN

- Billions of years of development.
- Neocortex was developed to cope with complex cognitive functions among other things.
- Made up of tiny columns, called neocortical columns.
- Began replicating themselves.
- Neocortex expanded to point where we had no more room in skull for them so folded in onto itself.

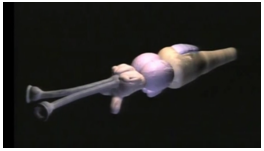


Figure: Early form of brain

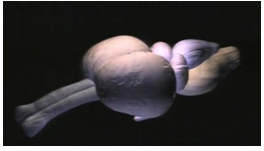


Figure: Development of neocortex

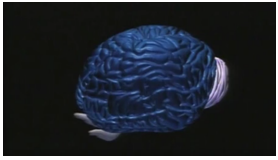


Figure: Developed brain of today

UPPSALA UNIVERSITET

ZOOM INTO THE BRAIN

- Zoomed out region
- Zoom to surface
- Neocortical columns
- A million of these in the human brain
- General structural idea of column in final picture
- Most accepted naming convention is the neocortical column

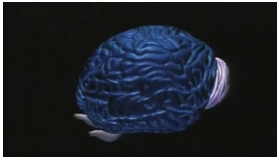


Figure: Full view of brain

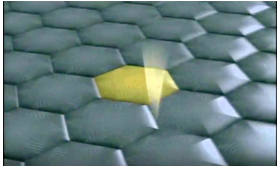


Figure: Surface view

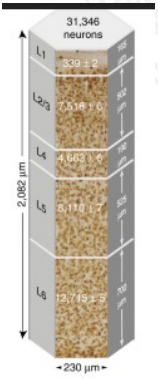
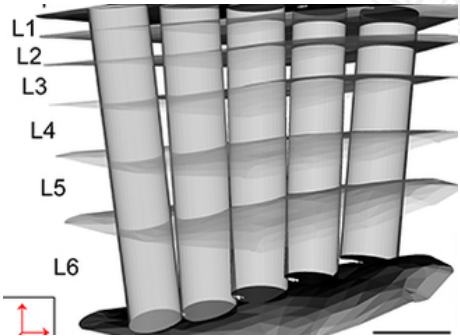


Figure: https://doi.org/10.1038/nn.4576

NEOCORTICAL COLUMN

A rat's brain has many of these neocortical columns

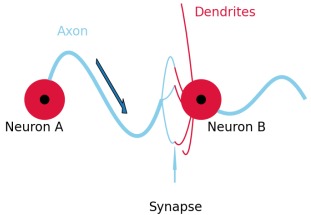


- Contain thousands of neurons
- Millions of synaptic Connections
- Forms a **Directed Graph**

UPPSALA UNIVERSITET

CONNECTIONS

- Each neuron has axons and dendrites
- These combine to create a synapse
- In relation to the synapse, we have a pre-synaptic neuron and a post-synaptic neuron
- A neuron can be both pre- and post-synaptic neurons
- Analogy to graphs is to say they have a source and a target respectively.

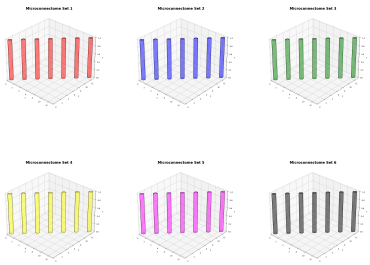


Information travel between neurons

UPPSALA UNIVERSITET

BLUE BRAIN PROJECT (BBP)

- Focus on somatosensory region
- Data obtained from experiments
- Reconstructions simulated from experimental data and refined until giving a likeness to the rats cortical column
- Multiple Instantiations of column - Microconnectome (MC)
- Sets based on different heights of the layers in neocortex
- Contain different neuron types (morphological)
- Set containing averages (we use final MC)

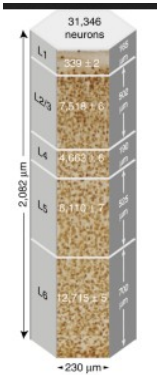


Instantiations

UPPSALA UNIVERSITET

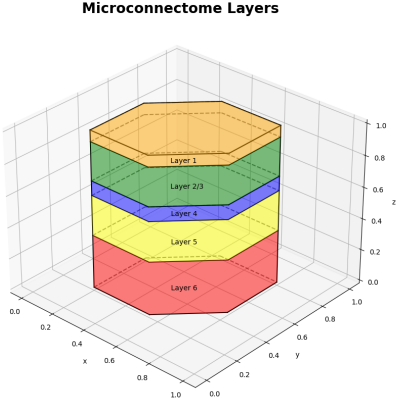
BIOM MICROCONNECTOME (MC) - OVERVIEW

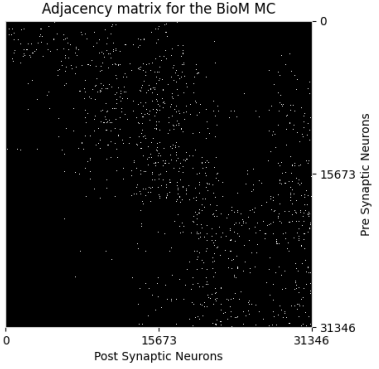
- 6 Layers
- 55 morphological neuron types
- Excitatory and inhibitory neurons
- 31,000 neurons
- 8 million synaptic connections



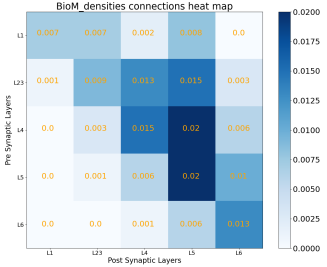
UPPSALA UNIVERSITET

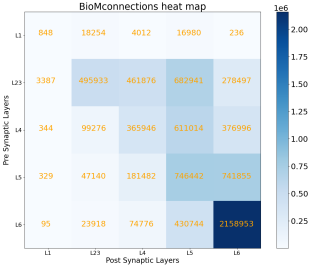
BIOM MC - DISTRIBUTION OF LAYERS AND NEURONS





BIOM MC - STATISTICS OF THE BIOM MC

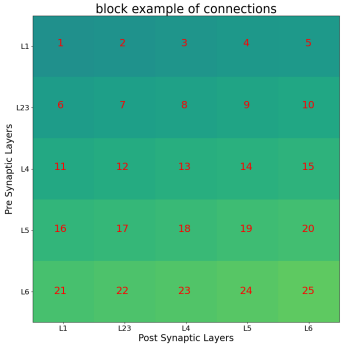




- $\frac{No.ofEdges}{No.ofNodes^2}$
- Most densely populated blocks in terms of connections are in layers 4 and 5
- Large portion of connections occurring in block 25, or layer 6 to layer 6 (the deeper cortical layer)

UPPSALA UNIVERSITET

BIOM MC - CONSTRAINT FOR MODELS

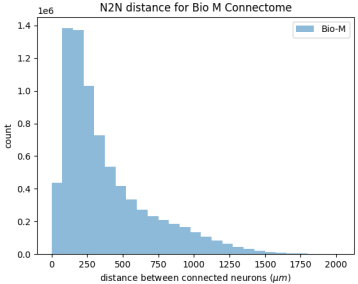


- Biological constraint where we split pre-synaptic neurons to post-synaptic neurons by layers to create blocks

UPPSALA UNIVERSITET

BIOM MC - CONSTRAINT FOR MODELS

Statistics of the construction of the BioM MC



- Distance distribution taken at 75μm as was the case of the General Biological model (discussed later on) where this preserved the soma-distances of connected pairs of neurons in all sub-matrices

UPPSALA UNIVERSITET

THE MODELS

Our Models

- Configuration Model
- Geometric Configuration Model
- Block Configuration Model
- Block Geometric Configuration Model

For Comparison

- Erdos-Renyi Model
- General Biological Model

  https://doi.org/10.3389/fncom.2017.00048

- BioM MC

UPPSALA UNIVERSITET

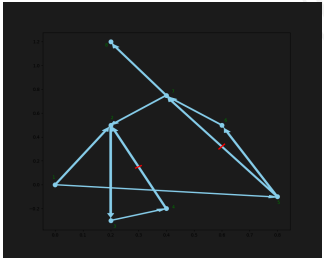
CONFIGURATION MODEL

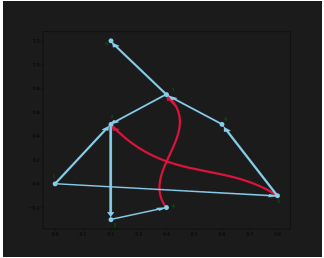
First Model

- Uses the empirical data from the BioM MC
- Matrix M of 1's (connection) and 0's(no connection)
- **Cut-Permute-Rewire** (Newman-Strogatz-Watts, 2001) algorithm is used to reconnect neurons
- Take a start and end point and make sure they are not the same, then connect
- Reassemble matrix M with permuted pre list and post list
- Matrix M gives us a representation of the directed random graph $\mathscr{G}_C$

Pre synaptic neurons	Post synaptic neurons
0	2
0	5
0	12
0	28
0	340
0	10401
0	12789
0	21992
0	25673
1	4
1	14
1	18987
1	31034
2	342
2	1945
2	12789
2	21992
2	25673
...	...

Pre synaptic neurons	Post synaptic neurons
15	2
171	5
3753	12
12908	28
56	340
738	10401
12	12789
1	21992
90	25673
254	4
12501	14
150	18987
904	31034
12908	342
1403	1945
24943	12789
2210	21992
45	25673
...	...



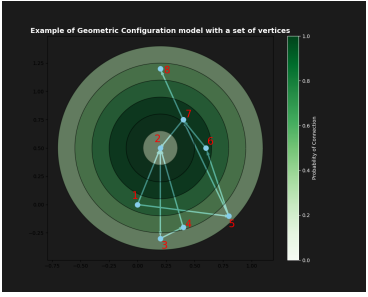


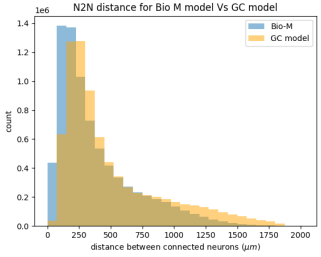
UPPSALA UNIVERSITET

GEOMETRIC CONFIGURATION MODEL

Second Model

- Uses Empirical data from MC
- Uses distance dependence obtained from BioM MC
- Apply the same **Cut-Permute-Rewire** algorithm with the constraint of distance dependence
- Return directed random graph $\mathscr{G}_{GC}$
- 84% of the mass is common regarding the distance distribution of the connected neurons in the BioM MC and this model



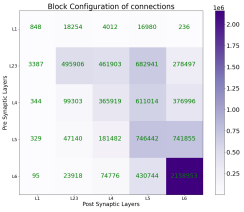


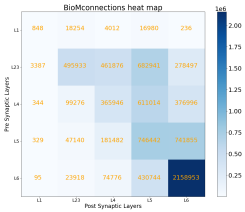
UPPSALA UNIVERSITET

BLOCK CONFIGURATION MODEL

Third Model

- Using BioM MC, split into blocks
- Use **Cut-Permute-Rewire** algorithm on each block
- Store each ”rewired” block
- Reassemble blocks to create matrix M representing directed random graph $\mathscr{G}_{BC}$
- Blocks ensure that we have the same number of connections in each block as the BioM MC



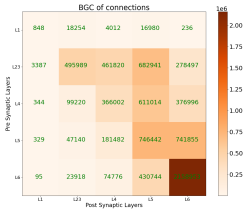


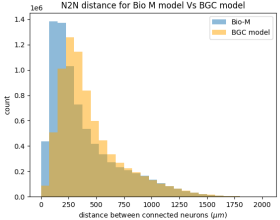
UPPSALA UNIVERSITET

BLOCK GEOMETRIC CONFIGURATION MODEL

Fourth Model

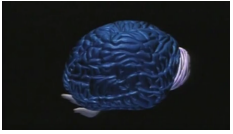
- Uses BioM MC
- Split into blocks again
- Compute pairwise distances of connected neurons in each block
- Combination of Model 2 and 3
- Use **Cut-Permute-Rewire** algorithm with the constraint of distance dependence applied for each block
- Reassemble blocks to create directed random graph $\mathcal{G}_{BGC}$

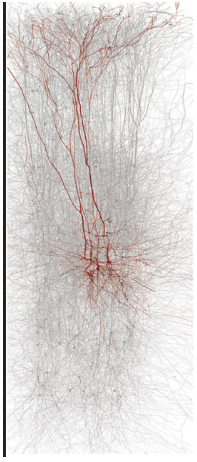


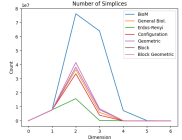


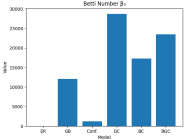
UPPSALA UNIVERSITET

RESULTS



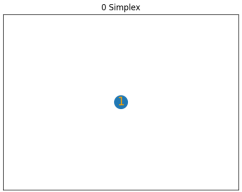


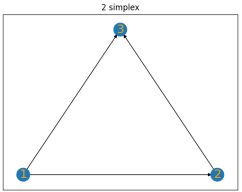


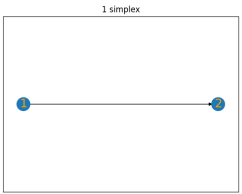


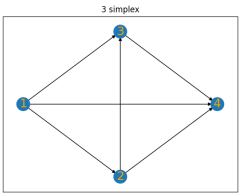
UPPSALA UNIVERSITET

SIMPLICES







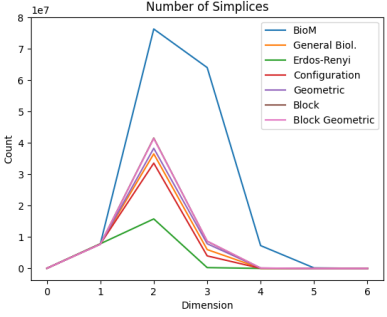


- A simplex(clique) is a generalisation of the notion of a triangle or tetrahedron to arbitrary dimensions
- It is named as such since it represents the simplest polytope in any given space
- The simplex gives a measure of the local connectivity of the connectome.
- 0 simplex a point
- 1 simplex, two vertices connected by a directed edge
- 2 simplex, 3 vertices connected by directed edges with a source and sink
- 3 simplex, 4 vertices connected by 6 directed edges, with a source and sink

UPPSALA UNIVERSITET

SIMPLICES

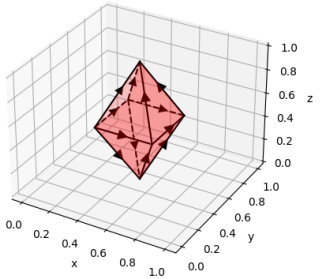
- Incrementally better results for our models (Block models gave very similar results)
- GB model gave more ordered and complex connectivity than the Configuration Model but less ordered and complex connectivity than the 3 other models in this project
- Biggest simplex count difference between the models occur in dimension 2
- Blocking the layers removed the difference between the inclusion of a geometric constraint and not including one



UPPSALA UNIVERSITET

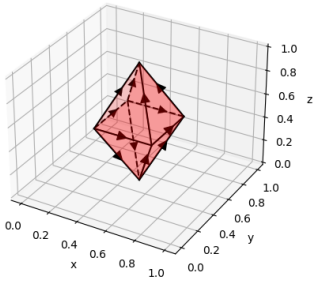
BETTI NUMBERS

- A measure of global connectivity
- β_0 is the number of connected components
- β_1 is the number of 1D or ”circular” holes
- β_2 is the number of 2D holes or voids
- $\beta_0 = 1, \beta_1 = 0, \beta_2 = 1$



BETTI NUMBERS

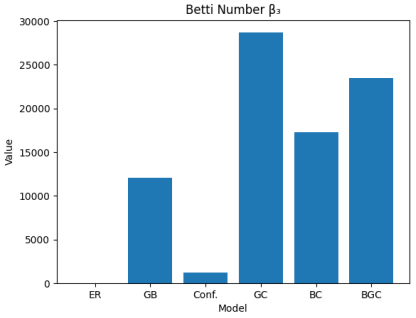
- If we add an edge connecting two vertices that are not already connected, then the void that is enclosed by our 2 dimensional cliques is no longer counted as a second Betti number but rather split into tetrahedra
- Vertices - Edges + Faces = 2
- Similarly, $\beta_0 - \beta_1 + \beta_2 = 1 - 0 + 1 = 2$
- Since our connectome is fully connected, then this is also considered as having a zeroth Betti number of 1



UPPSALA UNIVERSITET

BETTI NUMBERS

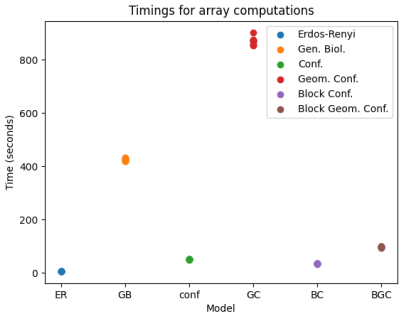
- Highest dimension for which we have a non-zero Betti Number is the third dimension
- Bigger difference between the bigger graph when the geometric constraint is applied
- The fourth Betti Number is non-zero for all models apart from the Erdos-Renyi model
- We have a fifth Betti Number for the BGC model



UPPSALA UNIVERSITET

COMPUTATION TIMES

- All computations done using same machine
- Memory: 15Gi, CPU(s): 12
- GC model would take similar time to complete until computation of remaining 20% of connections had to be computed.
- GB model took longer to compute because of reshuffling and requirement of a double for loop
- Remaining models had similar computation time since they either weren't restricted by distance distributions or they had smaller sets to work with or both.



SUMMARY

- Similar complexity of the models when compared to the GB model in terms of simplicial counts
- All models lacking complexity compared to BioM MC (biologically accurate model)
- Bigger difference with Betti Number counts in larger graphs, than when a block constraint is applied
- Geometric constraint showing high levels of global complexity
- `https://github.com/lamastex/working-manuscript-TopologicalDataAnalysisOnABrainNetwork`

UPPSALA UNIVERSITET

Thank You!!

Questions?