The Topological "Shape" of Brexit and Functional Networks

Mason A. Porter
Department of Mathematics
UCLA

MS 91, Tues. 5/23, 2:15 PM
(@masonporter)

Outline

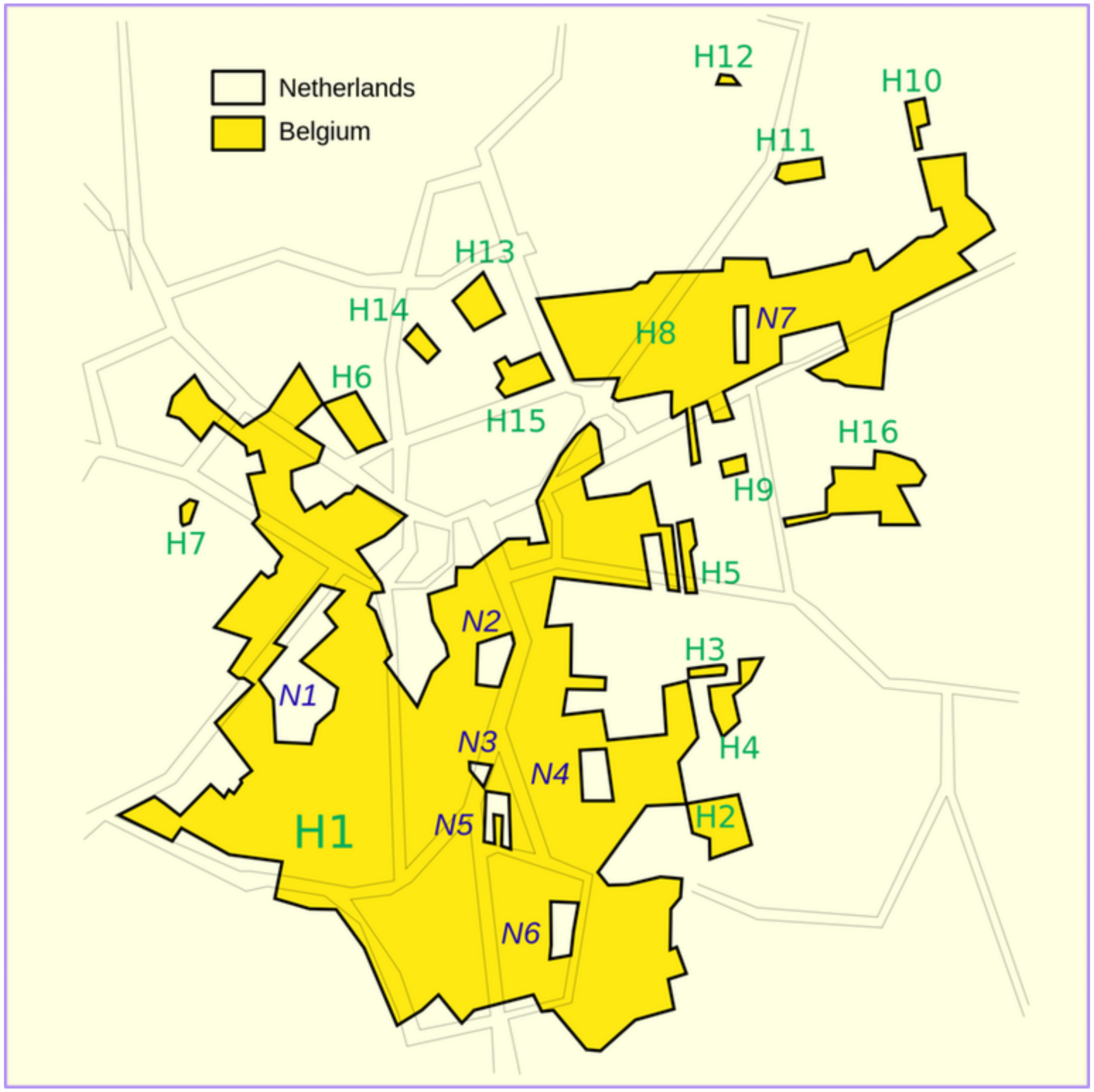
Introduction

"Brexit" Examples

"Functional Networks" in Neuroscience

Conclusions

Note: Various slides adapted or taken from slides of Bernadette J. Stolz (the first author on most of "my" papers I'll discuss). Most of the original research I am presenting is her work.



Border between Belgium & the Netherlands at Baarle-Nassau

A Few Introductory Resources for Topological Data Analysis

Chad Topaz’s awesome introductory article in DSWeb
- https://dsweb.siam.org/The-Magazine/Article/topological-data-analysis
- The most-read *DSWeb* article of all time

Book: Robert Ghrist, *Elementary Applied Topology*
- https://www.math.upenn.edu/~ghrist/notes.html

Nina Otter, MAP, Ulrike Tillmann, Peter Grindrod, and Heather A. Harrington, “A Roadmap for the Computation of Persistent Homology”, submitted, arXiv:1506.08903

Bernadette Stolz [2014], Masters Thesis, University of Oxford, *Computational Topology in Neuroscience*
- http://www.math.ucla.edu/~mason/research/Dissertation-stolz2014-Corr.pdf

Chad Giusti, Robert Ghrist, & Danielle S. Bassett [2016], “Two’s Company, Three (or More) is a Simplex”, *Journal of Computational Neuroscience*, Vol. 41, No. 1: 1–14

Links to various resources on my Quora answer on TDA
- https://www.quora.com/Applied-Mathematics-What-is-the-background-required-to-study-and-understand-topological-data-analysis

Introduction and Motivation

Algorithmic methods to study high-dimensional data (from point clouds, networks, etc.) in a quantitative manner

Examine "shape" of data

Persistent homology

- Mathematical formalism for studying topological invariants
- Fast algorithms
- Persistent structures: a way to cope with noise in data
- Allows examination of "higher-order" interactions (beyond pairwise) in data
  - A major reason for my interest in these methods (e.g., for networks)

Topological Data Analysis and Networks

Typically for weighted networks

In real-world networks, it is hard to extra significant structures (signal) from insignificant ones (noise).

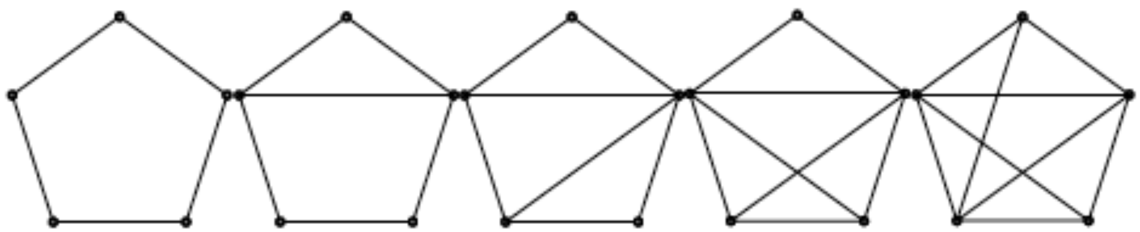
Sometimes convenient to threshold weights, binarize remaining values, and study the resulting graph

- Loss of important properties of original graph

Study global network characteristics

- Large-scale network structure, but of a different type from common ones like community structure
  - Useful: compare results of TDA approaches to “traditional” network approaches

Persistent Homology: Underlying Idea

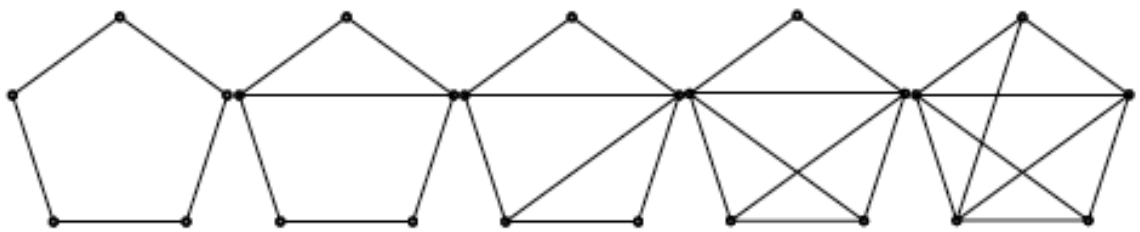


Idea: Consider a filtration

- For example: filter by the threshold for going from a weighted network to a binary network, and only keep (binarized) edges of at least that threshold.

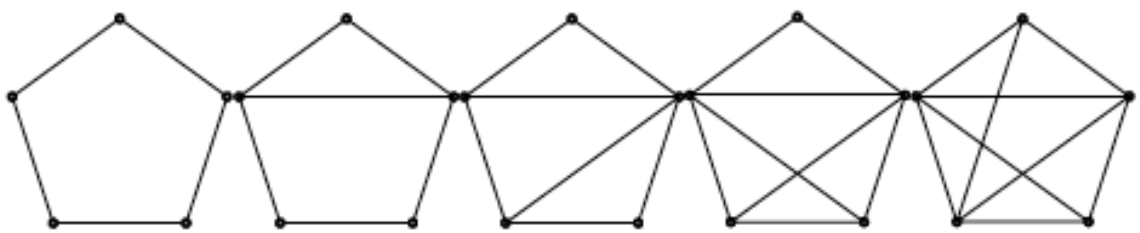
Study changes in topological structure along filtration by calculating topological invariants such as Betti numbers

Persistent Homology: Underlying Idea



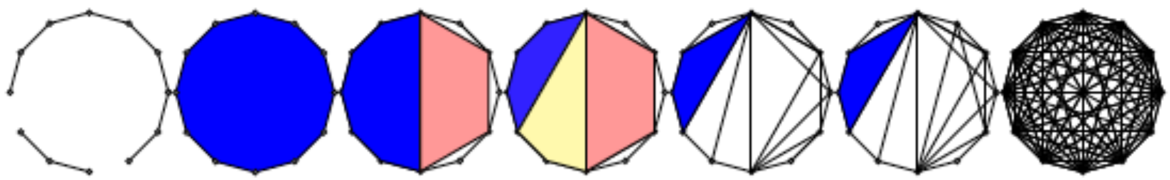
1. Construct a sequence of embedded graphs from a weighted network.
2. Define simplicial complexes.

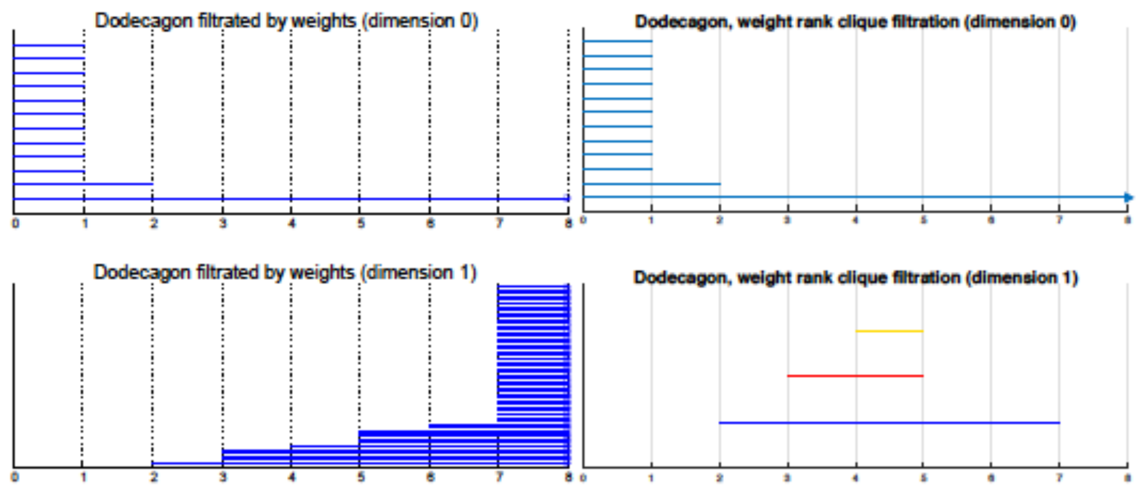
Persistent Homology: Weight Rank Clique Filtration (WRCF)
(e.g. G. Petri et al., *PLOS ONE*, 2013)



1. Construct a sequence of embedded graphs from a weighted network.
2. Define *k*-simplices to be the *k*-cliques present in the graph.

Example: Dodecagon





Persistence Landscapes

Introduced by P. Bubenik (2015)

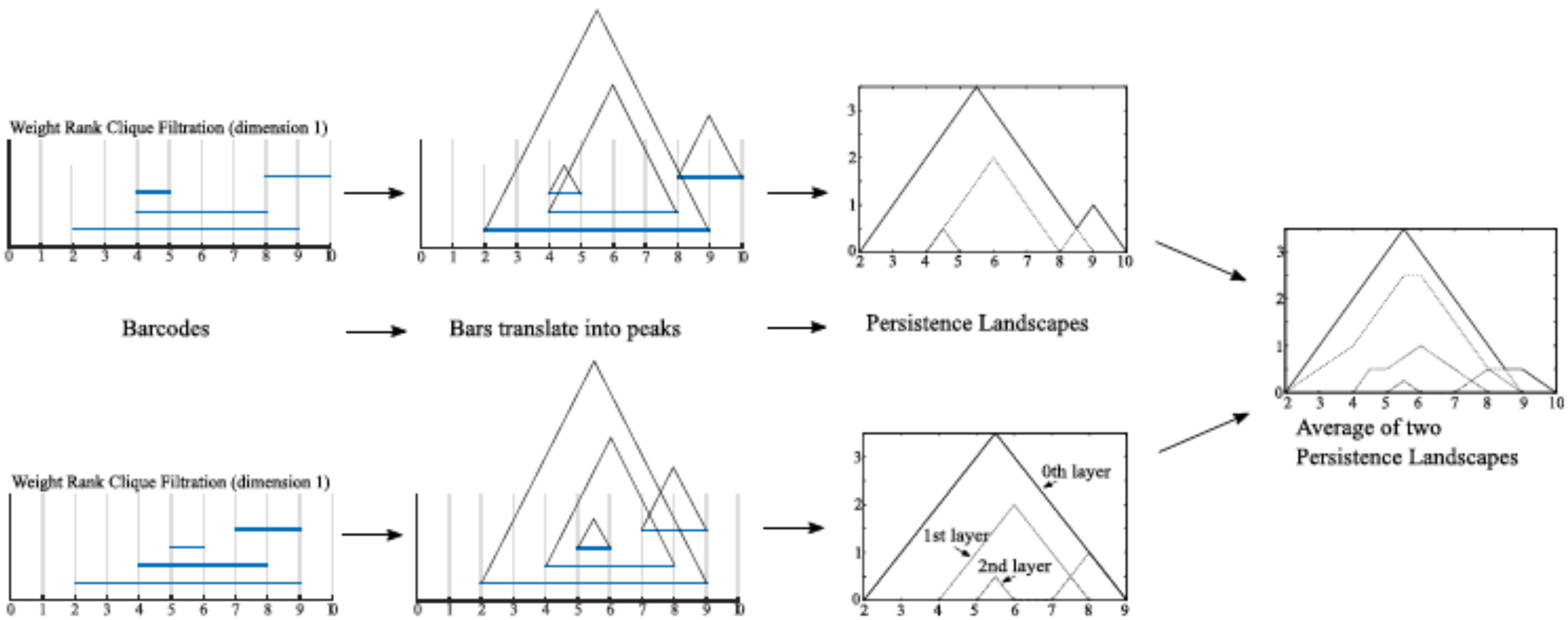


FIG. 5. Visualization of the relationship between barcodes and an average persistence landscape. To obtain a landscape from a barcode, one replaces every bar of the barcode by a peak, whose height is proportional to the persistence of the bar. In the landscape, we translate all peaks so that they touch the horizontal axis. The persistence landscape consists of different layers, where the qth layer corresponds to the qth-largest function values in the collection of peak functions. One creates an average of two landscapes by taking the mean over the function values in every layer.

This figure in B. J. Stolz et al., *Chaos*, 2017

Part I: The Topological "Shape" of Brexit

B. J. Stolz, H. A. Harrington, & MAP [2016], "The Topological "Shape" of Brexit", arXiv:1610.00752

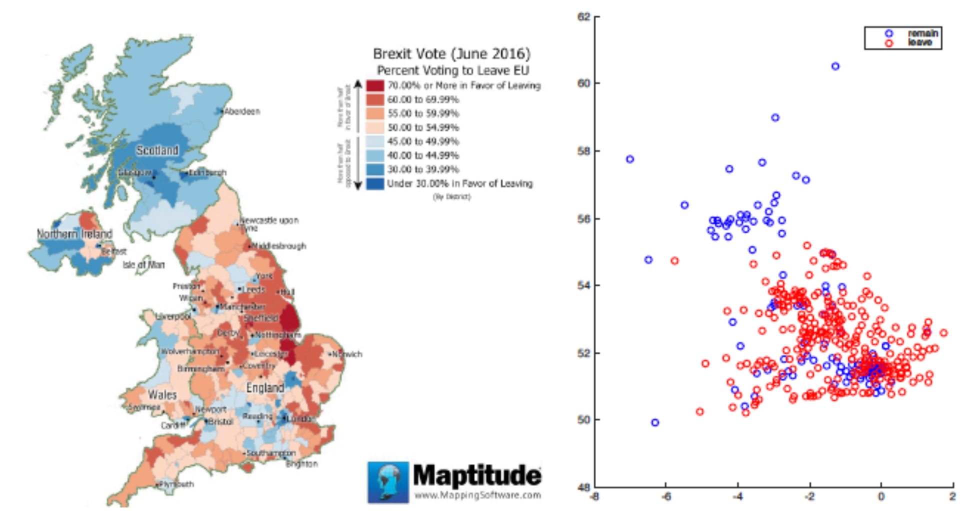


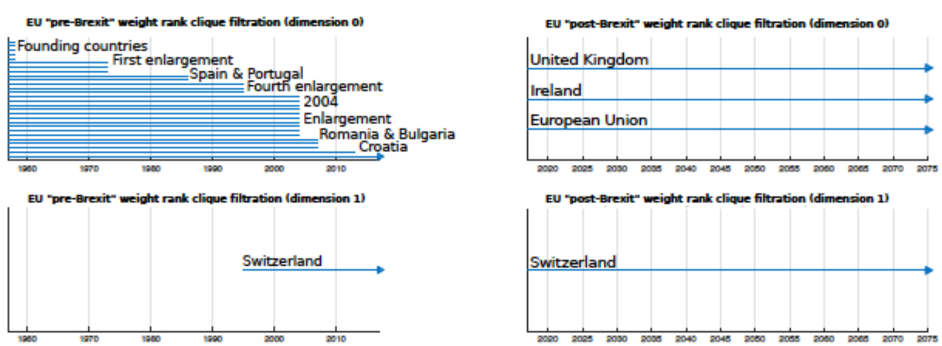
Figure 3: Point clouds based on the 2016 EU referendum voting result in the UK. We show the coordinates of "leave" districts in blue and the coordinates of "remain" districts in red.

Warmup: Network of EU Countries

Connect two countries with an edge if they are considered neighbors via a border (either in Europe or abroad), a bridge or a tunnel.

The edge weight is the later of the two years that the two countries joined the EU.

Consider WRCF



Example 2: Referendum Voting Data

Construct 2 point clouds

- 'Remain' point cloud: coordinates of cities in voting districts that voted to remain in the EU
  - Gibraltar omitted for simplicity
- 'Leave' point cloud: coordinates of cities in voting districts that voted to leave the EU

Construct a Vietoris–Rips filtration

- Choose a sequence $\{r_1, \ldots, r_n\}$ of increasing distances
- In the ith filtration step, we have k-simplices from unordered $(k+1)$-tuples with pairwise distance at most r_i

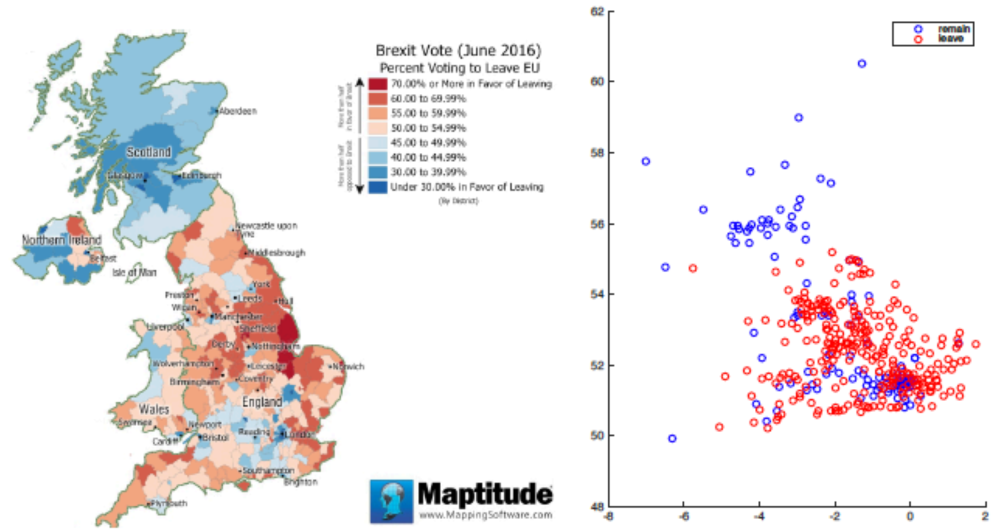


Figure 3: Point clouds based on the 2016 EU referendum voting result in the UK. We show the coordinates of "leave" districts in blue and the coordinates of "remain" districts in red.

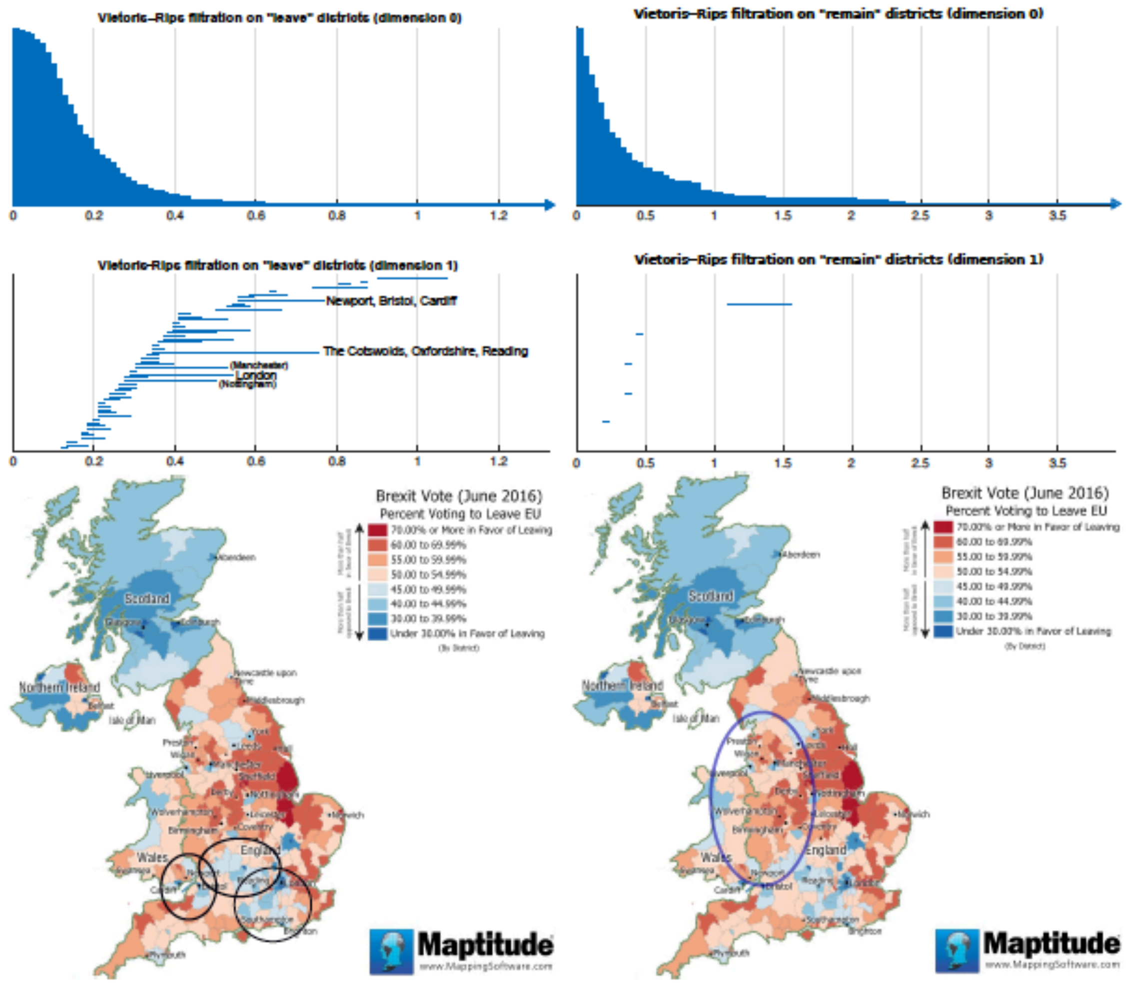


Figure 4: (Top two rows) Barcodes for dimensions 0 and 1 from a Vietoris–Rips filtration on (left) the leave point cloud and (right) the remain point cloud. (Bottom row) UK referendum voting map indicating the location of the most persistent loops in the dimension 1 barcodes of the Vietoris–Rips filtration.

Part II: Functional Networks

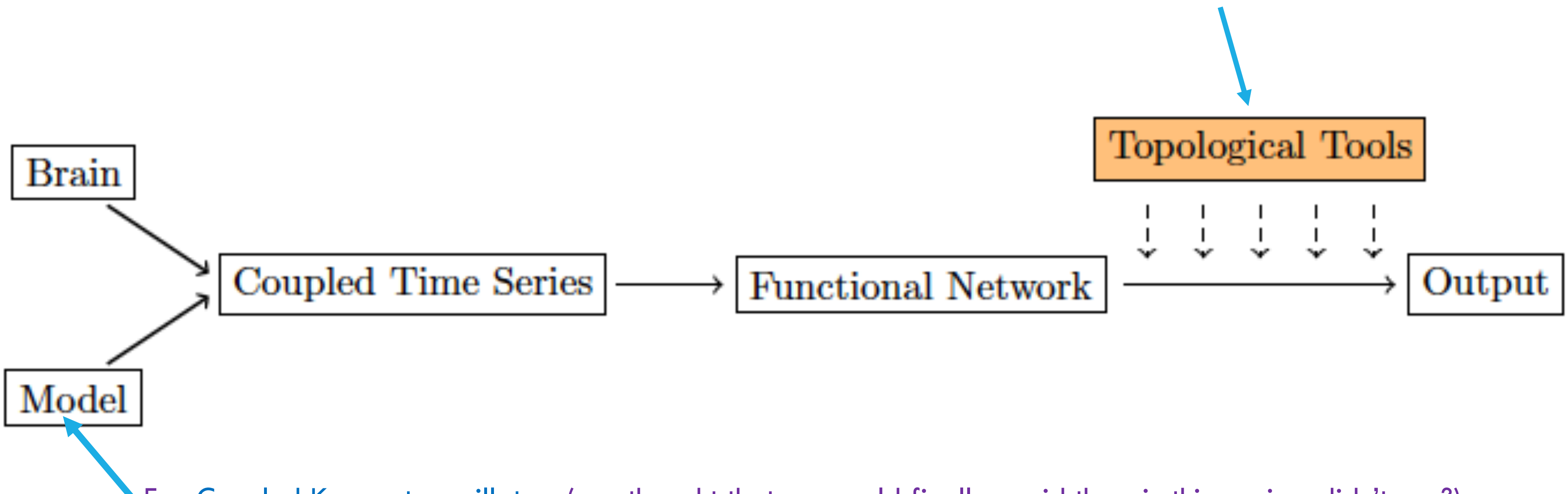
B. J. Stolz [2014], Masters Thesis, University of Oxford, *Computational Topology in Neuroscience*

- http://www.math.ucla.edu/~mason/research/Dissertation-stolz2014-Corr.pdf

B. J. Stolz, H. A. Harrington, & MAP [2017], "Persistent Homology of Time-Dependent Functional Networks Constructed from Coupled Time Series", *Chaos*, Vol. 27, No. 4: 047410

Pipeline

In contrast to using "traditional" methods for studying weighted networks



E.g. Coupled Kuramoto oscillators (you thought that you could finally avoid them in this session, didn't you?)

What is a *Functional Network*?

Functional versus Structural Networks

- Example from neuroscience:
  - Structural network: nodes = neurons, edges = synapses
  - Functional network: nodes = cortical areas, edges = behavioral similarity (quantified by similarity of time series)
- Example from ordinary differential equations:
  - Structural network: nodes = oscillators, edges = coupling between oscillators
  - Functional network: nodes = oscillators, edges = behavioral similarity (quantified by similarity of time series)

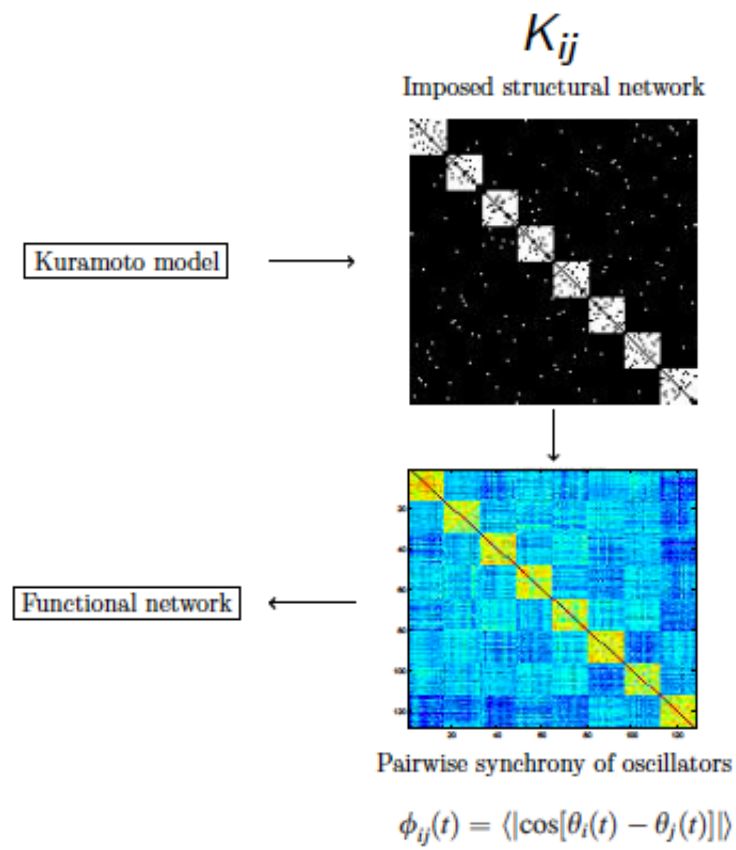
Functional networks are weighted and fully connected (or almost fully connected)

- We can study them using persistent homology
- Can compare results on large-scale structure to other approaches, such as community structure

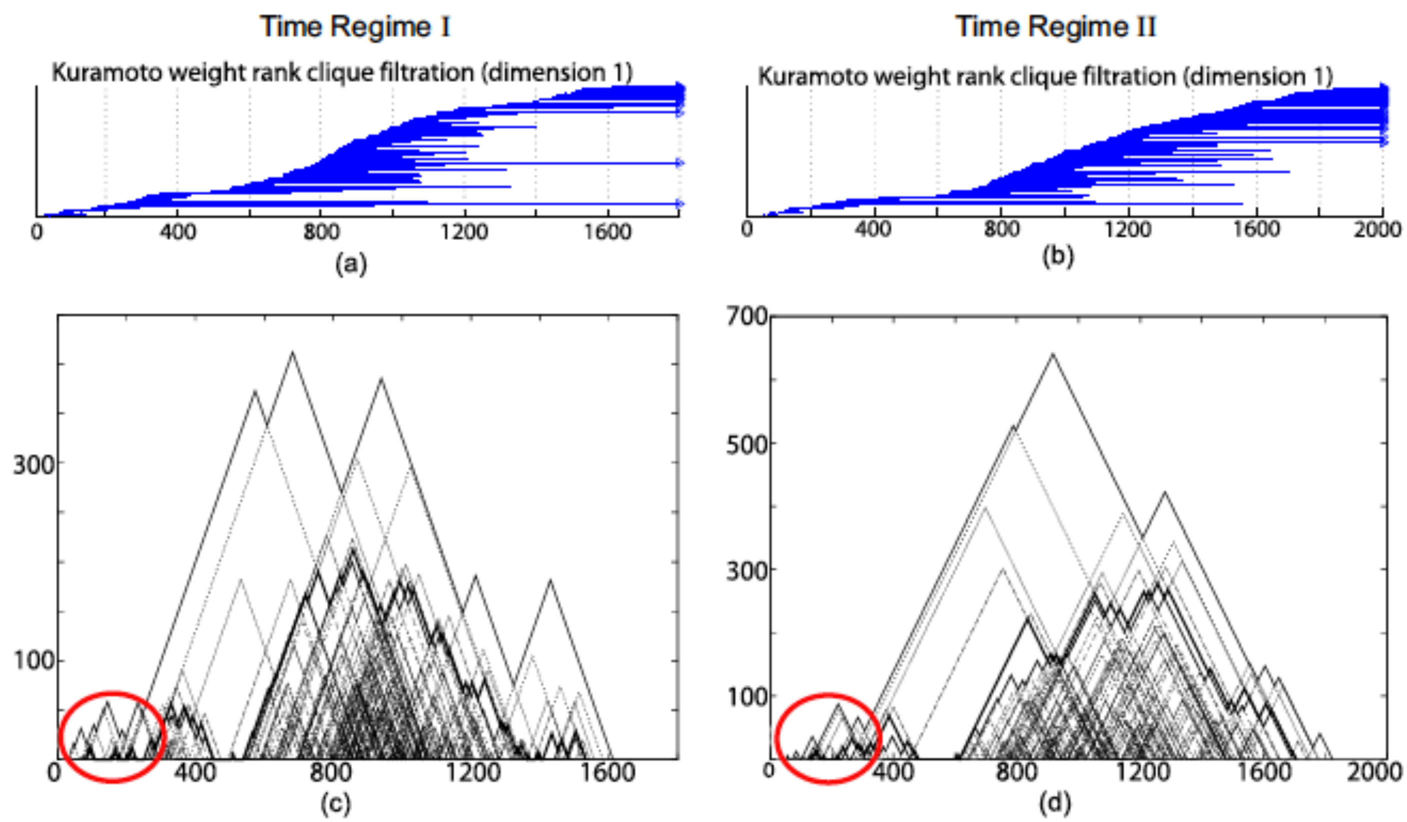
Example: Coupled Kuramoto Oscillators

$$\frac{d\theta_i}{dt} = \omega_i + \frac{1}{N}\sum_{j=1}^{N} K_{ij}\sin(\theta_j - \theta_i) \quad \text{for } i = 1, \ldots, N.$$

θ_i : phase of oscillator i,
ω_i: natural frequency
$K_{ij} \geq 0$: coupling strength,
N: number of oscillators in the model.



Barcodes and Persistence Landscapes



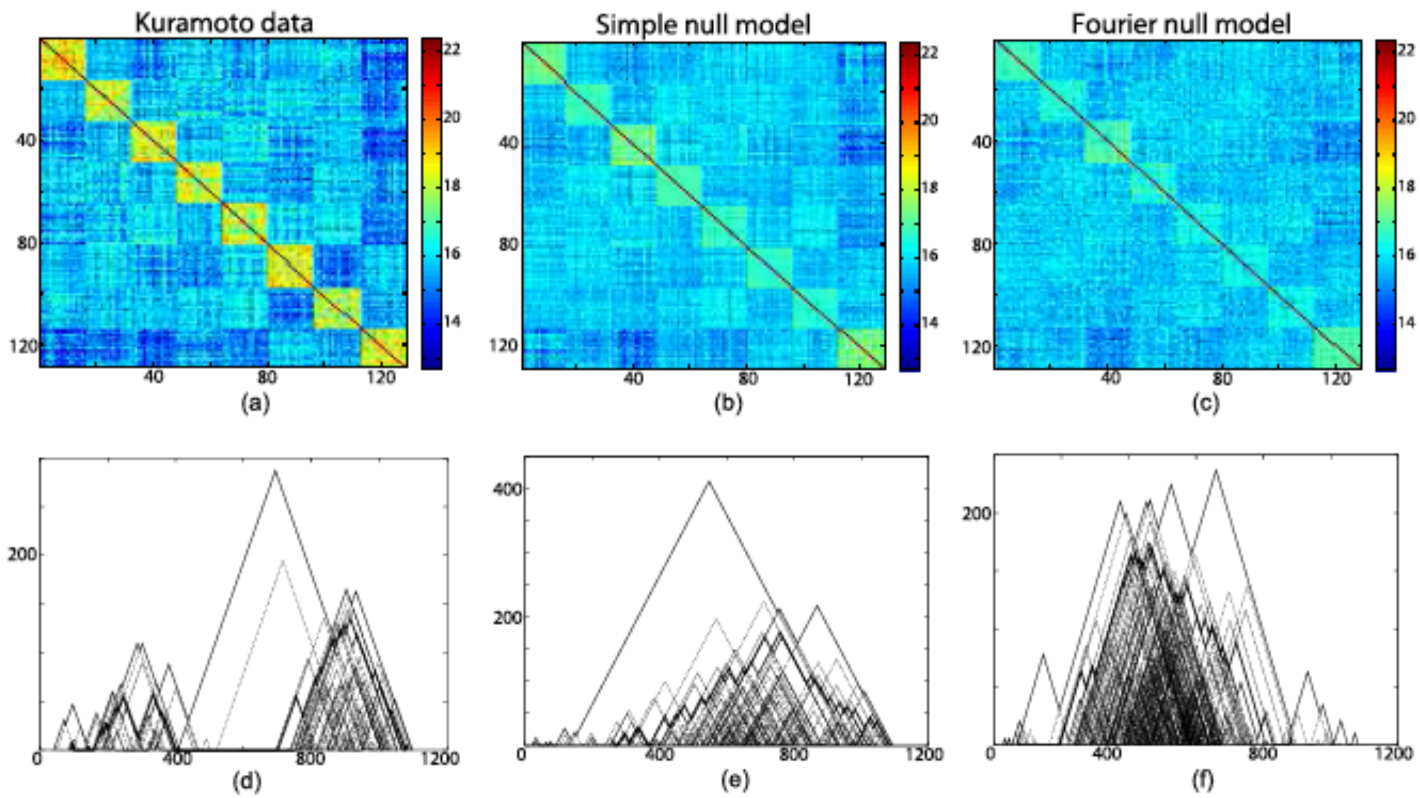
Kuramoto Data versus Null Models

Simple null model:

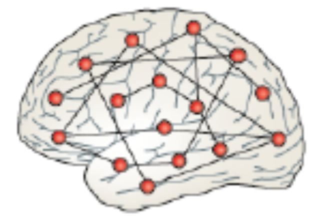
- Independently reassign the order of each oscillator's time series according to a uniform distribution (i.e., scramble time independently for each oscillator)

Fourier null model:

- Generate surrogate data by scrambling phases in Fourier space



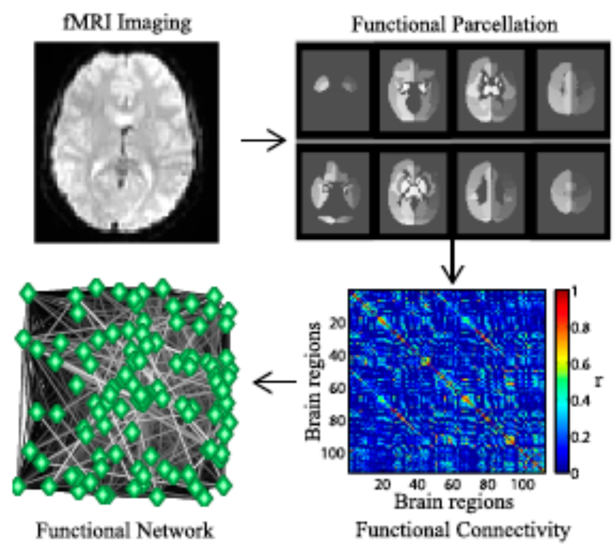
Example: fMRI Data



(Source: Bullmore and Sporns (2009), Nature Reviews: 186 – 198.)

Data from D. S. Bassett, N. F. Wymbs, MAP, P. J. Mucha, J. M. Carlson, & S. T. Grafton [2011], "Dynamic Reconfiguration of Human Brain Networks During Learning", *PNAS*, Vol. 118, No. 18: 7641–7646

- Weighted networks from time-series similarity (wavelet coherence) of neuronal activity of brain regions during performance of simple motor task
- In the above paper and follow-ups, we studied things like community structure and core–periphery structure.
- Using persistent homology gives another way to examine large-scale ("mesoscale") network structures
- These data also used in D. S. Bassett, MAP, N. F. Wymbs, S. T. Grafton, J. M. Carlson, & P. J. Mucha [2013], "Robust Detection of Dynamic Communities in Networks", *Chaos*, Vol. 23, No. 1: 013141



Differences in Different Days?

Experimental observations on 3 different days (20 participants)

Right plot: Average persistence landscapes

Landscape peak shifts to the left in later days

- I.e. they are formed by edges with higher weights, indicating that there is stronger synchronization between the associated brain regions

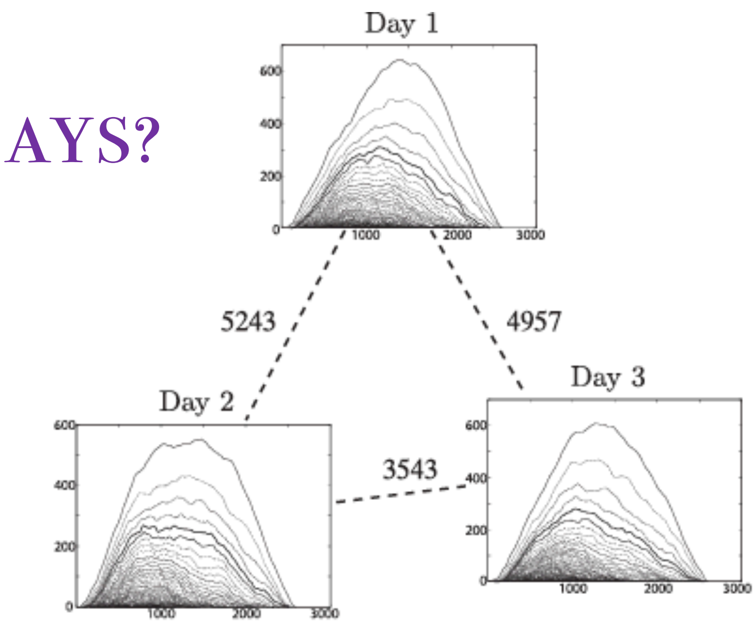


FIG. 10. Visualization of average persistence landscapes for days 1, 2, and 3 of task-based fMRI networks. The distance between the landscape for day 1 and the other two landscapes is larger than that between the landscapes for days 2 and 3. (The L^2 distances between them are 5243 between days 1 and 2, 4957 between days 1 and 3, and 3543 between days 2 and 3.) The standard deviations from the average landscapes are larger than the calculated distances, so these values need to be interpreted cautiously. We also observe a shift to the left of the landscape peak during the three days, indicating that the particularly persistent loops in these networks arise earlier in the filtration for the later days. In other words, they are formed by edges with higher edge weights, indicating that there is stronger synchronization between the associated brain regions.

Conclusions

Computing persistent homology can give insights into large-scale structure of networks

- Complements network clustering methods, such as detection of mesoscale features like community structure and core–periphery structure
- Important: going beyond pairwise interactions in networks

Observation: Sometimes relatively short features (e.g. as visualized in short barcodes) represent meaningful features. (We saw this in both Kuramoto and fMRI data.)

- E.g. strongly synchronized Kuramoto oscillators within the same community of a structure network
- Contrasts with conventional wisdom: longer (i.e. more persistent) features are supposed to be the signals, and shorter features are usually construed as noise

Our Brexit example was a toy, but it’s worth looking at that kind of data more seriously using TDA approaches.

Reminder: If you want to get started on PH, look at our “roadmap” paper: Nina Otter, MAP, Ulrike Tillmann, Peter Grindrod, and Heather A. Harrington, “A Roadmap for the Computation of Persistent Homology”, submitted, arXiv:1506.08903