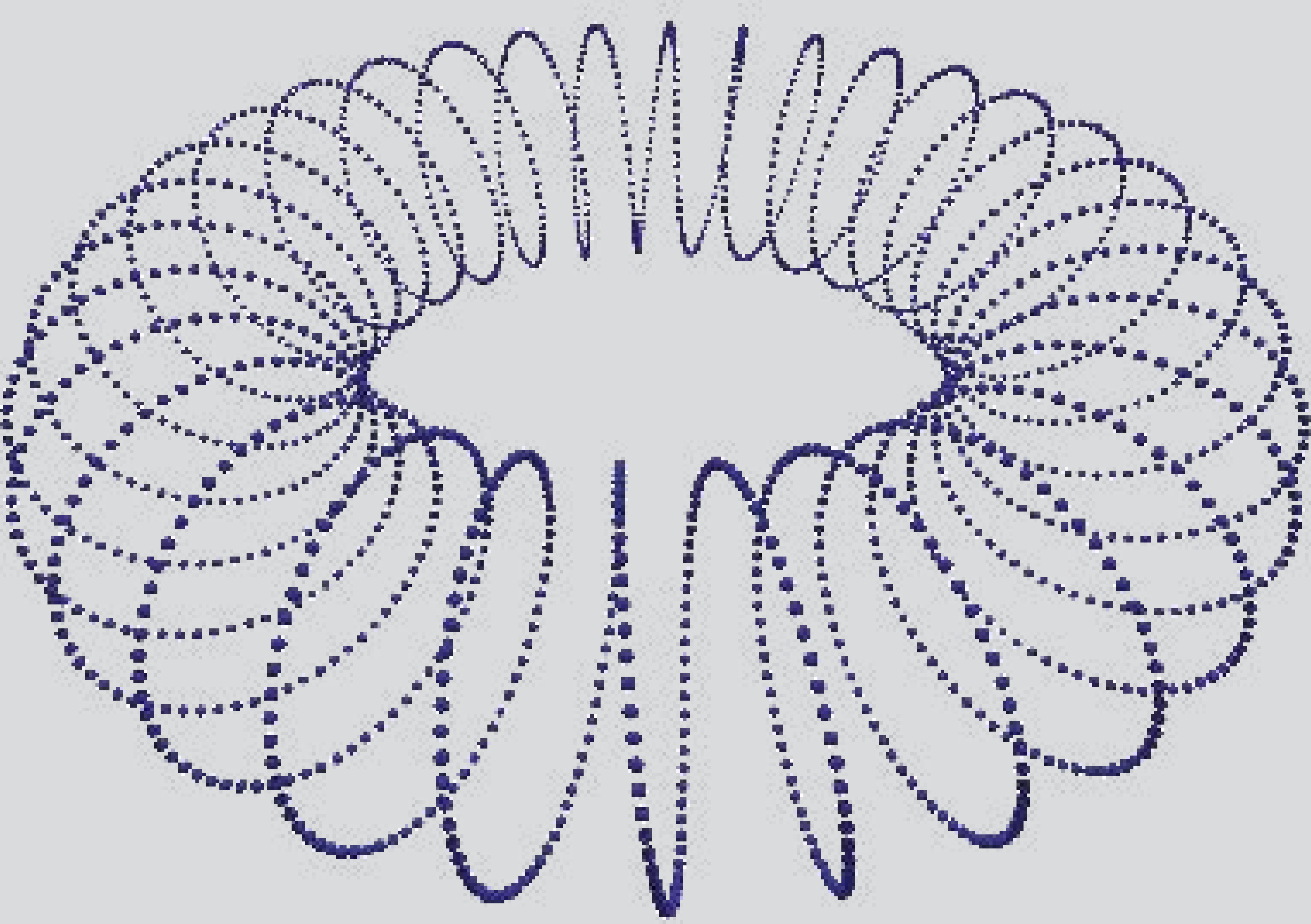
Spires

23–24 Sept 2019

CENTRE FOR TOPOLOGICAL DATA ANALYSIS
Mathematical Institute
Andrew Wiles Building

ORGANISERS
Heather Harrington
Ulrike Tillmann



Cover ilustration by Dmitriy Morozov,
Lawrence Berkeley National Lab.

Contents

About us.................5
Location..................6
Schedule.................8
Participants.........10
Abstracts.............12

About us

One of the biggest challenges of the information age is the understanding of complex data.

Welcome to our distinguished guests. Welcome to our multidisciplinary team of mathematicians, statisticians, and computer scientists from Oxford, Swansea and Liverpool. We have the necessary breadth and depth of experience and expertise to drive the development and application of Topological Data Analysis (TDA) to the solution of real world problems.
Our vision is to build a two-way bridge between data users and scientists on which topological ideas and tools go to and fro testing & applications and research & development.

One of the biggest challenges of the information age is the understanding of complex data. Mathematics plays a crucial part in the development of essential tools for this. In the last 15 years topological methods have matured into a field of its own right, Topological Data Analysis.

The purpose of this meeting is to bring together international researchers working in topological data analysis and to discuss new developments and challenges that will shape the research in the future.

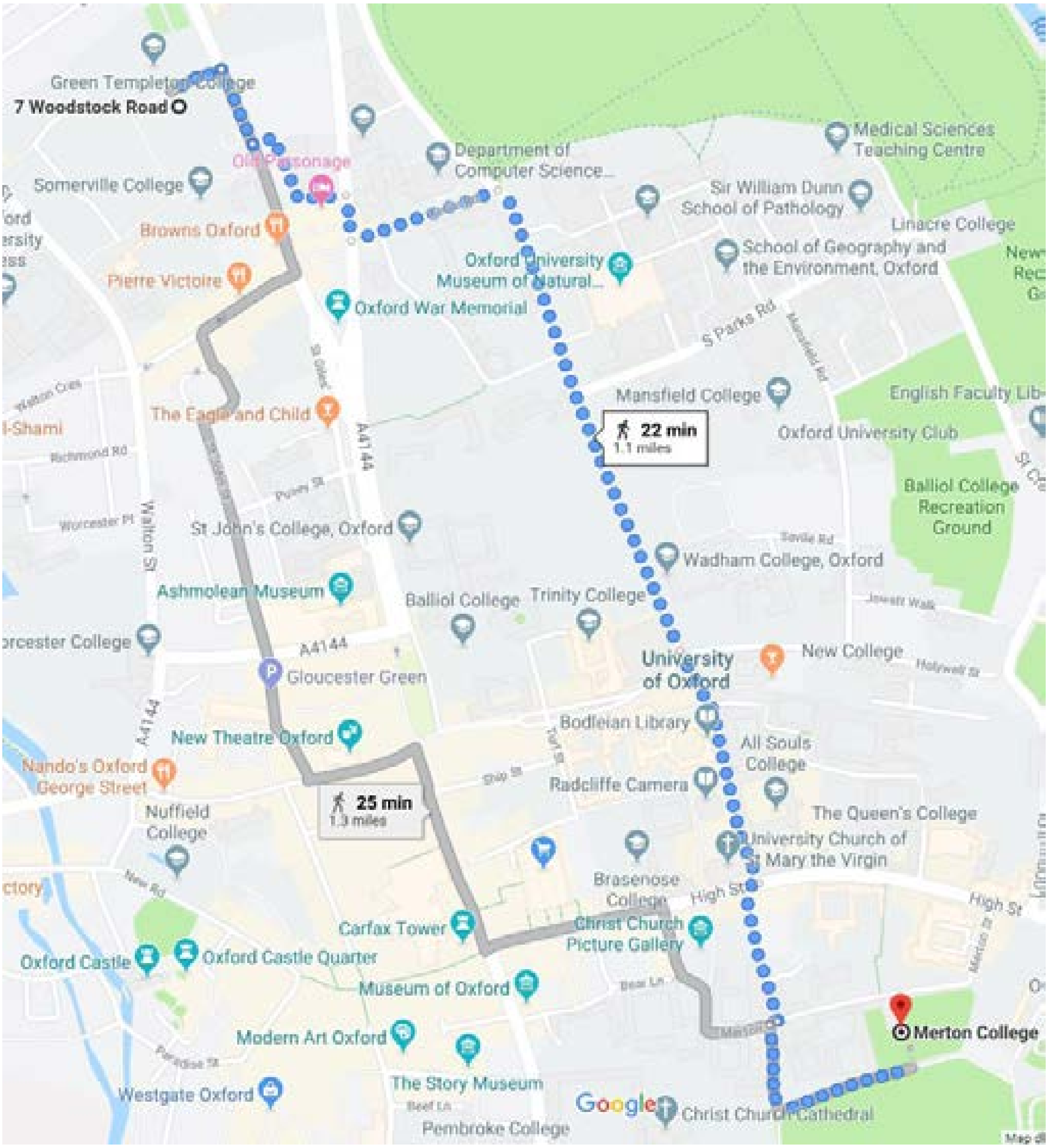
Location

SPIRES takes place in the mezzanine level of the Mathematical Institute, Andrew Wiles Building, Woodstock Road OX2 6GG.

Talks take place in lecture theatre **L2**. Refreshment breaks happen in the area just outside of L2.

Lunch will be served in the **Common Room, Level 1**.

Dinner on Monday evening is hosted by **Merton College**, which is a 20 minute walk from Mathematics at Merton St, OX1 4JD.


7 Woodstock Road
Merton College
22 min
1.1 miles
25 min
1.3 miles
Green Templeton College
Somerville College
Old Parsonage
Browns Oxford
Pierre Victoire
The Eagle and Child
St John's College, Oxford
Ashmolean Museum
Gloucester Green
New Theatre Oxford
Nuffield College
Oxford Castle
Oxford Castle Quarter
Museum of Oxford
Modern Art Oxford
The Story Museum
Westgate Oxford
Pembroke College
Carfax Tower
Christ Church Picture Gallery
Christ Church Cathedral
Brasenose College
Radcliffe Camera
Bodleian Library
University of Oxford
All Souls College
University Church of St Mary the Virgin
The Queen's College
New College
Wadham College, Oxford
Trinity College
Balliol College
Mansfield College
Oxford University Museum of Natural...
Oxford War Memorial
Department of Computer Science...
Sir William Dunn School of Pathology
School of Geography and the Environment, Oxford
Medical Sciences Teaching Centre
Linacre College
English Faculty Lib
Oxford University Club
Balliol College Recreation Ground
Nando's Oxford George Street
Google

Schedule

Monday

L2

10:00	Registration (Mezzanine)
10:30-11:15	Shmuel Weinberger *University of Chicago* **Complexity, testability, the Euclidean hypothesis, and the blessings and curses of dimension**
11:15	Tea & Coffee (Mezzanine)
11:45-12:15	Jacek Brodzki *University of Southampton* **Geometry, Topology and he structure of data**
12:15-12:45	Florian Pausinger, *Queen's University Belfast* **Persistent Betti numbers of random Cech complexes**
12:45	Lunch (Common Room)
14:30-15:00	Pawel Dlotko *Swansea University* **From Ball Mapper to Dynamical Systems and back**
15:00-15:30	Ruben Sanchez Garcia *University of Southampton* **Morse clustering an impossibility theorem**
15:30	Tea & Coffee (Mezzanine)
16:00-17:00	Gunnar Carlsson *Stanford* **Topology for deep learning**
19:00	Drinks Reception (Merton)
19:30	Dinner (Merton)

Tuesday

L2

09:00-09:30 Steve Oudot *École Polytechnique*
A mathematica framework for the differentiability of persistent Homology?

09:30-10:00 Nina Otter *University of Oxford*
Magnitude meets persistence. What happens after?

10:00 Tea & Coffee (Mezzanine)

10:30-10:45 Oliver Vipond *University of Oxford*
TDA for histology data

10:45-11:00 Ambrose Yim *University of Oxford*
Persistent Homology and graph classification

11:00-11:15 Agnese Barbensi *University of Oxford*
Grid diagrams as tools to investigate knot spaces and to poisomerase-mediated simplification of DNA topology*

11:15-11:30 Phil Smith *University of Liverpool*
Skeletonisation Algorithms with Theoretical Guarantees for Unorganised Point Clouds with High Levels of Noise

11:30-11:45 Marco Mosca *University of Liverpool*
Grid diagrams with their relatively Geometric invariants for clustering datasets of crystal structure

11:45 Break (Mezzanine)

12:00-12:30 Bernadette Stolz *University of Oxford*
Outlier-robust subsampling techniques for persistent homology

12:30-13:00 Kelly Spendlove *Rutgers*
Computational Conley theory

Participants

ARAS ASAAD	OXFORD DRUG DESIGN
AGNESE BARBENSI	UNIVERSITY OF OXFORD
GABRIELE BELTRAMO	QUEEN MARY UNIVERSITY
MATTHEW BURFITT	UNIVERSITY OF SOUTHAMPTON
JACEK BRODZKI	UNIVERSITY OF SOUTHAMPTON
ADAM BROWN	IST AUSTRIA
GUNNAR CARLSSON	STANFORD
ALVARO TORRAS CASAS	CARDIFF UNIVERSITY
PAWEL DLOTKO	SWANSEA UNIVERSITY
RUBEN SANCHEZ GARCIA	UNIVERSITY OF SOUTHAMPTON
JEHAN GHAFURI	UNIVERSITY OF BUCKINGHAM
DEJAN GOVC	ABERDEEN UNIVERSITY
HEATHER HARRINGTON	UNIVERSITY OF OXFORD
JOHN HARVEY	SWANSEA UNIVERSITY
TAHIR HASSAN	UNIVERSITY OF BUCKINGHAM
SABAH JASSIM	UNIVERSITY OF BUCKINGHAM
KELVIN KILLEEN	NATIONAL UNIVERSITY OF IRELAND- GALWAY
SIDDHARTH KUMAR	UNIVERSITY OF OXFORD
VITALIY KURLIN	UNIVERSITY OF LIVERPOOL
JANIS LAZOVSKIS	UNIVERSITY OF ABERDEEN
JACOB LEYGONIE	UNIVERSITY OF OXFORD
SUNG HYUN LIM	KOREA INSTITUTE FOR ADVANCED STUDY
BARBARA MAHLER	UNIVERSITY OF OXFORD
LEWIS MARSH	UNIVERSITY OF OXFORD
INGRID MEMBRILLO-SOLIS	UNIVERSITY OF SOUTHAMPTON

DARVIN MERTSCH	UNIVERSITY GREIFSWALD
MARCO MICHELE MOSCA	UNIVERSITY OF LIVERPOOL
SLAWOMIR NASUTO	UNIVERSITY OF READING
NINA OTTER	UNIVERSITY OF OXFORD
STEVE OUDOT	ÉCOLE POLYTECHNIQUE
MEGAN PALSER	UNIVERSITY OF SOUTHAMPTON
MARIAM PIRASHVILI	UNIVERSITY OF SOUTHAMPTON
FLORIAN PAUSINGER	QUEEN'S UNIVERSITY BELFAST
IGOR POTAPOV	UNIVERSITY OF LIVERPOOL
ANA LUCIA GARCIA PULIDO	UNIVERSITY OF LIVERPOOL
GESINE REINERT	UNIVERSITY OF OXFORD
HENRI RIIHIMAKI	UNIVERSITY OF ABERDEEN
MICHAEL SHAUB	UNIVERSITY OF OXFORD
PHIL SMITH	UNIVERSITY OF LIVERPOOL
JASON SMITH	UNIVERSITY OF ABERDEEN
DIOGO SORIANO	FEDERAL UNIVERSITY OF ABC (UFABC)
KELLY SPENDLOVE	RUTGERS
BERNADETTE STOLZ-PRETZER	UNIVERSITY OF OXFORD
ULRIKE TILLMANN	UNIVERSITY OF OXFORD
BRUNELLA TORRICELLI	UNIVERSITY OF CAMBRIDGE
OLIVER VIPOND	UNIVERSITY OF OXFORD
NANXIN WEI	IMPERIAL COLLEGE LONDON
SHMUEL WEINBERGER	UNIVERSITY OF CHICAGO
NAYA YEROLEMOU	UNIVERSITY OF OXFORD
AMBROSE YIM	UNIVERSITY OF OXFORD
KA MAN YIM	UNIVERSITY OF OXFORD

Abstracts

AGNESE BARBENSI
University of Oxford

11:00 Tuesday

Grid diagrams as tools to investigate knot spaces and topoisomerase-mediated simplification of DNA topology*

Abstract: Grid diagrams with their relatively simple mathematical formalism provide a convenient way to generate and model projections of various knots. It has been an open question whether these 2D diagrams can be used to model a complex 3D process such as the topoisomerase-mediated preferential unknotting of DNA molecules. We model here topoisomerase-mediated passages of double-stranded DNA segments through each other using the formalism of grid diagrams. We show that this grid diagram-based modelling approach captures the essence of the preferential unknotting mechanism, based on topoisomerase selectivity of hooked DNA juxtapositions as the sites of intersegmental passages. We show that grid diagram-based approach provide an important, new and computationally convenient framework for investigating entanglement in biopolymers.

Joint work with: D.Celoria, H.A.Harrington, A.Stasiak, D.Buck

JACEK BRODZKI
University of Southampton

11:45 Monday

Geometry, Topology, and the structure of data

Modern data is astonishing in its variety, and is far removed from anything that could be held in a spreadsheet or analysed using traditional methods. Through intense research effort, topology has emerged as a source of novel methodology to provide insight into the structure of very complex, high dimensional data. It provided us with tools like persistent homology, which are used to compute numerical topological characteristics of the data. More recently, these methods have been augmented by geometric insights, which are valuable in capturing the structure of and relationships between complex shapes.

In this talk, I will provide an overview of new techniques from topology and geometry and illustrate them on particular examples. One set of data was created through the study of CT scans of human lungs, and another addresses the problem of classification of three-dimensional shapes.

GUNNAR CARLSSON
Stanford

16:00 Monday

Topology for deep learning

Deep learning refers to a family of methods for modeling large and complex data sets. They have demonstrated remarkable capabilities on various kinds of data. It is also true, however, that the workings of the algorithms are not well understood, and that therefore the internal states of the "learner" provides an opportunity for data analysis to gain understanding. I will discuss the methods, as well as the applications of topology to them.

PAWEL DLOTKO
Swansea University

14:30 Monday
From Ball Mapper to Dynamical Systems and back

I will start this talk by introducing the Ball Mapper, an algorithm that provides a low dimensional layout of higher dimensional data sets. I will present a recent R package implementing it and show how it can be used to analyze some real life high dimensional datasets. Later I will explain how the ideas behind Ball Mapper can be used to reconstruct dynamical systems from time series of observations. We will conclude by building graph-based topological models of such a reconstructed dynamics.

RUBEN SANCHEZ GARCIA
University of Southampton

15:00 Monday
Morse culstering an impossibility theorem*

Given a set of objects and a pairwise similarity function, a clustering algorithm is a computational procedure that groups together objects which are similar and separate the ones which are not. Although simple and useful, standard clustering methods fail to incorporate node annotation, typically in the form of vertex labels or weights.

We describe a family of clustering algorithms for annotated networks, inspired by Morse Theory in Differential Topology, which detect the most influential nodes (local maxima) as well as their associated basins of attraction and flows. We test our algorithm against the LFR benchmark, and explain how we combined it with TDA methods to identify asthma phenotype from blood gene expression profiles.

We also study Morse clustering from an axiomatic point of view. In a famous impossibility result, Kleinberg introduced three natural clustering properties, or axioms, and showed they cannot be simultaneously satisfied by any clustering algorithm. We present a new clustering property, Monotonic Consistency, which avoids the well-known problematic behaviour of Kleinberg's Consistency axiom, and the impossibility result: we show that Morse Clustering satisfies Kleinberg's original axioms with Consistency replaced by Monotonic Consistency.

*Joint work with Fabio Strazzeri.

MARCO MOSCA
University of Liverpool

11:30 Tuesday
Grid diagrams with their relatively Geometric invariants for clustering datasets of crystal structure

The study of similarity between crystal structures is the aim of my Phd project. My presentation will focus on geometric invariants which are properties that may allow us to distinguish and cluster different crystal structures. Voronoi Cell of a lattice and distances between molecules are the main subjects of this talk.

Abstracts

NINA OTTER

University of Oxford

09:30 Tuesday

Magnitude meets persistence. What happens after?

The magnitude is an isometric invariant of metric spaces that was introduced by Tom Leinster in 2010, and is currently the object of intense research. Magnitude encodes many invariants of a metric space such as volume, dimension, capacity, etc. In this talk I will give an overview of existing results and current research in this area, explain how magnitude is related to persistent homology, and then discuss some open questions and work in progress. In particular, I will discuss how this relationship might help in defining an Euler characteristic for multiparameter persistent homology.

This talk is partly based on the preprint: https://arxiv.org/abs/1807.01540.

STEVE OUDOT

École Polytechnique

09:00 Tuesday

A mathematica framework for the differentiability of persistent Homology?*

: In this talk I will define notions of differentiability for maps from and to the space of persistence barcodes. Inspired from diffeology theory, the proposed framework uses lifts to the space of ordered barcodes, from which derivatives can be computed. The two derived notions of differentiability (respectively from and to the space of barcodes) combine together naturally to produce a chain rule that enables the use of gradient descent for objective functions factoring through the space of barcodes. I will illustrate the versatility of this framework by showing how it can be used to analyze the smoothness of various parametrized families of filtrations arising in the TDA literature.

*This is joint work with Jacob Leygonie (University of Oxford), with some contributions by Ulrike Tillmann.

FLORIAN PAUSINGER

Queen's University Belfast

12:15 Monday

Persistent Betti numbers of random Cech complexes

We study the persistent homology of random Cech complexes.*

Generalizing a method of Penrose for studying random geometric graphs, we first describe an appropriate theoretical framework in which we can state and address our main questions. Then we define the Kth persistent Betti number of a random Cech complex and determine its asymptotic order in the subcritical regime. This extends a result of Kahle on the asymptotic order of the ordinary kth Betti number of such complexes to the persistent setting.

*Joint work with Ulrich Bauer (TU Munich).

PHIL SMITH
University of Liverpool

11:15 Tuesday

Skeletonisation algorithms with theoretical guarantees for unorganised point clouds with high levels of noise

The skeletonisation problem is to find a 1-dimensional skeleton that correctly represents the shape of the cloud. For example, given a noisy point sample of an unknown underlying graph, a reconstructed skeleton should be geometrically close and homotopy equivalent to the underlying graph. We compare several algorithms that attempt to solve the above skeletonisation problem. One of these algorithms produces a Homologically Persistent Skeleton (HoPeS) for any cloud without extra parameters. This universal skeleton contains subgraphs that provably represent the 1-dimensional shape of the cloud at any scale. The experiments on synthetic and real data suggest HoPeS generally outperforms the other algorithms.

KELLY SPENDLOVE
Rutgers

12:30 Tuesday

Computational Conley theory

Algebraic topology and dynamical systems are intimately related: the algebra may constrain or force the existence of certain dynamics. Morse homology is the prototypical theory grounded in this observation. Conley theory is a far-reaching topological generalization of Morse theory and a great deal of effort over the last few decades has established a computational version of the Conley theory. The computational Conley theory is a blend of combinatorics, order theory and algebraic topology and has proven effective in tackling problems within dynamical systems. Within the Conley theory the connection matrix is the mathematical object which transforms the approach into a truly homological theory; it is the Conley-theoretic generalization of the Morse boundary operator. We'll discuss a new formulation of connection matrix theory that makes it clear the connection matrix is a new, important tool in topological data analysis. We'll demonstrate our techniques with an application of our theory and software to the setting of a Morse theory on spaces of braid diagrams.

BERNADETTE STOLZ-PRETZER
University of Oxfordr

12:00 Tuesday

Outlier-robust subsampling techniques for persistent homology*

The amount and complexity of biological data has increased rapidly in recent years with the availability of improved biological tools. Topological data analysis and more specifically persistent homology have been successfully applied to biological settings. When attempting to study large data sets however, many of the currently available algorithms fail due to computational complexity preventing many interesting biological applications. De Silva and Carlsson (2004) introduced the so called Witness Complex that reduces computational complexity by building simplicial complexes on a small subset of landmark points selected from the original data set. The landmark points are chosen from the data either at random or using the so called maxmin algorithm. These approaches are not ideal as the random selection tends to favour dense areas of the point cloud while the maxmin algorithm often selects outliers as landmarks. Both of these problems need to be addressed in order to make the method more applicable to data. Chawla (2013) developed a version of k- means that detects outliers while clustering data points. We show how this method can be used to select landmarks for persistent homology and also propose another new method specifically for the use in persistent homology that can detect outliers based on the local persistent homology around data points. We show how both of these methods outperform the existing subsampling methods for persistent homology. We further illustrate how local persistent homology can be applied to detect geometric anomalies in data.

* Joint work with: Heather Harrington, Vidit Nanda and Jared Tanner

Abstracts

OLIVER VIPOND
University of Oxford

10:30 Tuesday

TDA for Histology Data

In this talk we shall explore the insights that topological data analysis can bring to the analysis of histology data. We shall begin with simulation data and familiar single parameter persistence techniques, progressing to present novel multiparameter persistence techniques, developed in order to cope with the noisier real world histology samples.

i

SHMUEL WEINBERGER
University of Chicago

10:30 Monday

Complexity, testability, the Euclidean hypothesis, and the blessings and curses of dimension*

I would like to advertise a collection of ideas that I believe are interesting for theoretical reasons, and should have implications on the way different data sets are analyzed. One of the key insights of probability theory is that large populations frequently have more regular behavior than individuals. This is critical (i.e. a blessing) for data compression and some models of deep learning. It poses a strong challenge to TDA (one curse among many!) because small measure events can have very large topological implications. I shall explain that idea of a testable invariant which is essentially an invariant that can be approximated through some sampling method and discuss (non)examples. I will show how assuming that one's data lies in a Euclidean space helps with testability, but this fact can give one pause about the (widely implicitly assumed) Euclidean hypothesis.

*This talk is largely based on other people's ideas.

.

AMBROSE YIM
University of Oxford

10:45 Tuesday

Persistent Homology and Graph Classification

We discuss how graphs can be optimally filtered such that their persistence diagrams are optimal for graph classification. With the invention of barcode derivatives, we can now compute the gradient of a classifier's loss function with respect to the parameters of the filter function. This allows the filter to be tuned such that the persistence diagram is optimised for the classification task. We discuss the results of some numerical experiments.

Notes