# CD02

## *Topological Data Analysis: Theory and Applications*

## Análisis Topológico de Datos: Teoría y Aplicaciones

*Organizers* **Organizadores** **Antolatzaileak**

**Manuel M Cuerno**
(CUNEF Universidad)

**Inés García-Redondo**
(Imperial College )

*Description* **Descripción** **Deskribapena**

***Topological Data Analysis (TDA) leverages algebraic topology to address high-dimensional, complex data problems, where data present non-linear interactions that are hard to capture using traditional statistical and data analytic techniques. Since its introduction in the early 2000s, TDA has been instrumental in advancing both theoretical and applied research. In this special session, we invite you to embark on a journey through both sides of TDA: the rich mathematical framework that underpins TDA, which has led to robust methodologies for analyzing complex data sets, and the successful application of this theory to cutting-edge, real-world problems across diverse scientific domains.***

El Análisis de Datos Topológico (TDA, por sus siglas en inglés) utiliza la topología algebraica para abordar problemas de datos complejos y de alta dimensión, los cuales presentan interacciones no lineales difíciles de capturar con técnicas estadísticas y analíticas tradicionales. Desde su introducción a principios de los 2000, el TDA ha sido fundamental para avanzar en investigación teórica y aplicada. En esta sesión especial, invitamos a un recorrido por ambas facetas del TDA: el rico marco matemático que lo sustenta, el cual ha dado lugar a metodologías robustas para analizar problemas de datos complejos, y la exitosa aplicación de esta teoría a problemas de vanguardia en el mundo real en diversos dominios científicos.

*MSC Codes* **Códigos MSC** **MSC Kodeak**

55N31
(primary)

62R40
(secondary)

*Slots* **Bloques** **Blokeak**

2.A (Aula 0.13); 2.B (Aula 0.13); 2.C (Aula 0.13)

*QR Code* **Código QR** **QR Kodea**

*Session Schedule* **Horario de la Sesión** **Saioaren Ordutegia**

J16 | 11:00-11:20 | 0.13
*On the Limitations of Fractal Dimension as a Measure of Generalization*
**Inés García-Redondo** (Imperial College)

J16 | 11:30-11:50 | 0.13
*Simplicial-based neural networks*
**Eduardo Paluzo Hidalgo** (Universidad de Loyola)

J16 | 12:00-12:20 | 0.13
*Topological evolution of the layers of a neural network during training*
**Jon Ander Alonso** (Universidad de La Rioja)

J16 | 12:30-12:50 | 0.13
*Shedding light on black-box models through topological data analysis*
**Clara Isabel López González** (Universidad Complutense de Madrid)

J16 | 16:30-16:50 | 0.13
*The Mayer-Vietoris spectral sequence for Alpha Complexes*
**Álvaro Torras** (CRESS UMR 1153 Inserm)

J16 | 17:00-17:20 | 0.13
*Reach of Embeddings of Metric Spaces inside the Persistence Diagram Space*
**Javier Casado** (Universidad Autónoma de Madrid)

J16 | 17:30-17:50 | 0.13
*The Depth Poset: A Channel Between Topology and Dynamics*
**Manuel Soriano-Trigueros** (Institute of Science and Technology Austria)

J16 | 18:00-18:20 | 0.13
*New Betti numbers for finite simplicial complexes*
**Pablo Hernández** (Universidad de Salamanca)

V17 | 9:30-9:50 | 0.13
*Topology and data reduction*
**Javier Perera Lago** (Universidad de Sevilla)

V17 | 10:00-10:20 | 0.13
*Topology across scales on multiplexed data*
**María Torras Pérez** (University of Oxford)

V17 | 10:30-10:50 | 0.13
*The Extended Pareto Grid in Multiparameter Persistent Homology*
**Eloy Mosig** (Università di Pisa)

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *11:00-11:20* | **11:00-11:20** | **11:00-11:20** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

***On the Limitations of Fractal Dimension as a Measure of Generalization***

**Inés García-Redondo**
(Imperial College)

Bounding and predicting the generalization gap of neural networks remains a central open problem in machine learning. Neural network optimization trajectories have been proposed to possess fractal structure, leading to bounds based on notions of fractal dimension on these trajectories. In this talk, I will present an extended statistical evaluation of these topological generalization measures, as well as two counterexamples where the predicted theory fails.

Joint work with Charlie Tan, Qiquan Wang, Michael Bronstein and Anthea Monod.

arXiv:2406.02234

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *11:30-11:50* | **11:30-11:50** | **11:30-11:50** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

***Simplicial-based neural networks***

**Eduardo Paluzo Hidalgo**
(Universidad de Loyola)

This talk explores the use of simplicial complexes and maps in designing neural networks, introducing a novel simplicial layer and advancements in training algorithms for these topologically inspired structures. These methods enhance network robustness and universal approximation capabilities. Theoretical insights and practical implementations will be presented, showing how this approach integrates with classical neural networks and its potential impact on machine learning.

arXiv:2403.15083

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *12:00-12:20* | **12:00-12:20** | **12:00-12:20** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

***Topological evolution of the layers of a neural network during training***

**Jon Ander Alonso**
(Universidad de La Rioja)

TDA is increasingly used in DL model analysis, offering insights into issues like generalization and explainability. We conducted an experimental study of the topological evolution of a small neural network during training. At each epoch, we applied persistent homology to the activations of layers as they processed training and test samples. Our findings indicate relationships between the learning curve and topological statistics, as well as variations in topological complexity across layers.

Joint work with Eduardo Sáenz de Cabezón.

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *12:30-12:50* | **12:30-12:50** | **12:30-12:50** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

***Shedding light on black-box models through topological data analysis***

**Clara Isabel López González**
(Universidad Complutense de Madrid)

Convolutional neural networks excel in many computer vision tasks, but their black-box nature limits decision transparency. To address this, we introduce Latent Landscapes, a TDA-based tool for studying the topology of latent representations. It offers a global view of encoded information, its variety, and evolution, and compares these insights with existing explanatory methods.

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *16:30-16:50* | **16:30-16:50** | **16:30-16:50** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

*The Mayer-Vietoris spectral sequence for Alpha Complexes*

**Álvaro Torras**
(CRESS UMR 1153 Inserm)

The alpha filtration is used to compute persistent homology (PH) from point samples, but interpreting PH in complex or periodic datasets is challenging. I present the Mayer-Vietoris spectral sequence (MVSS) for distributed computation of alpha filtration PH, along with conditions for simplicial collapses and experiments using the Permaviss++ package.

Joint work with F. Jensen.

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *17:00-17:20* | **17:00-17:20** | **17:00-17:20** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

*Reach of Embeddings of Metric Spaces inside the Persistence Diagram Space*

**Javier Casado**
(Universidad Autónoma de Madrid)

We recall the reach of a subset of a metric space, defined as the supremum of radii for which the metric projection onto the subset is unique. The reach is 0 if no neighborhood of the subset admits a unique projection. Bubenik and Wagner present an isometric embedding of any bounded separable metric space into the space of persistence diagrams. The main result shows that if every point in X is an accumulation point, the reach of X in the persistence diagram space is 0.

Joint work with Manuel Cuerno and Jaime Santos-Rodríguez.

arXiv:2307.01051

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *17:30-17:50* | **17:30-17:50** | **17:30-17:50** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

*The Depth Poset: A Channel Between Topology and Dynamics*

**Manuel Soriano-Trigueros**
(Institute of Science and Technology Austria)

The depth poset is defined over persistence pairs of a filtered complex, with its minimal elements being shallow pairs, recently defined by U. Bauer and F. Roll. When viewed as a Morse function, shallow pairs act as vectors of the combinatorial gradient, offering insights into the topological simplification of the complex. This talk will explore the structure of the poset and its connection to classical concepts in persistence and discrete Morse theory.

Joint work with Herbert Edelsbrunner, Marian Mrozek, and Michał Lipiński.

arXiv:2311.14364

| *Thursday 16* | **Jueves 16** | **Osteguna 16** |
|---|---|---|
| *18:00-18:20* | **18:00-18:20** | **18:00-18:20** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

*New Betti numbers for finite simplicial complexes*

**Pablo Hernández**
(Universidad de Salamanca)

Simplicial homology's homotopy invariance limits Betti numbers in capturing properties of simplicial complexes. In this talk, I will introduce new invariants for finite simplicial complexes that extend traditional Betti numbers, revealing dimensions and higher-order adjacencies of simplices around holes. I'll also discuss how persistence enhances the interpretation of these invariants and explore their applications in topological data analysis.

*Friday 17*
*9:30-9:50*
*[Room 0.13]*

**Viernes 17**
**9:30-9:50**
**[Aula 0.13]**

**Ostirala 17**
**9:30-9:50**
**[Gela 0.13]**

*Topology and data reduction*

**Javier Perera Lago**
(Universidad de Sevilla)

Training machine learning models requires large datasets and can be computationally expensive. Data reduction is a family of preprocessing techniques that reduces the size or complexity of the data, speeding up the training process. In this talk, we present data reduction methods based on topological concepts and demonstrate that preserving the dataset's topological properties leads to models with strong performance.

*Friday 17*
*10:00-10:20*
*[Room 0.13]*

**Viernes 17**
**10:00-10:20**
**[Aula 0.13]**

**Ostirala 17**
**10:00-10:20**
**[Gela 0.13]**

*Topology across scales on multiplexed data*

**María Torras Pérez**
(University of Oxford)

Advancements in multiplex imaging have enabled the simultaneous visualisation of multiple cell types in a single tissue sample. In this talk, we present a data analysis pipeline based on persistence homology (PH) and its application to multiplexed lupus murine spleen data. We propose a visualisation of PH cycles which is simple and computationally inexpensive and can help interpretatability. We investigate various PH vectorisations and propose additional weightings of persistence images.

Joint work with Heather Harrington, Ulrike Tillmann, Iris Yoon, and Helen Byrne.

| *Friday 17* | **Viernes 17** | **Ostirala 17** |
|---|---|---|
| *10:30-10:50* | **10:30-10:50** | **10:30-10:50** |
| *[Room 0.13]* | **[Aula 0.13]** | **[Gela 0.13]** |

***The Extended Pareto Grid in Multiparameter Persistent Homology***

**Eloy Mosig**
(Università di Pisa)

Recent research has developed strategies for computing and approximating the matching distance in multiparameter persistent homology. This talk will focus on the extended Pareto grid, a geometric construction at the intersection of Morse theory and optimization. Associated with smooth functions from a Riem. manifold to the Eucli. plane, the Pareto grid helps track homological features of the manifold's filtration, offering new methods for computing matching distance in the multiparameter regime.

Joint work with P. Frosini, N. Quercioli and F. Tombari.

arXiv:2312.04201