

# Topological data analysis as a bridge to geometric deep learning

Patrizio Frosini

Department of Mathematics, ARCES, Alma Human-AI,  $AM^2$  - University of Bologna  
`patrizio.frosini@unibo.it`

Department of Mathematics, University of Pisa, 9 May 2023

## Summary

---

- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?
- 4 Group equivariant non-expansive operators
- 5 Building linear and nonlinear GNEOs
- 6 How can we use GNEOs in applications?

- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?
- 4 Group equivariant non-expansive operators
- 5 Building linear and nonlinear GENEOS
- 6 How can we use GENEOS in applications?

## About my talk

---

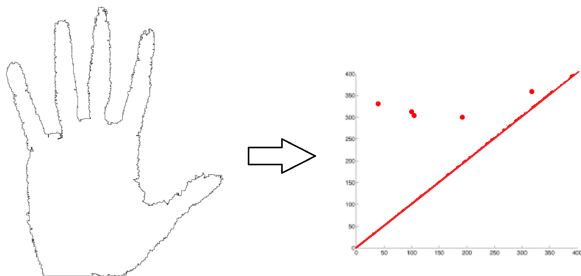
- The purpose of this talk is to illustrate my research activity in the fields of Topological Data Analysis (TDA) and Geometric Deep Learning (GDL), also describing some recent theoretical results and their applicability in concrete cases.
- I will limit the technical details to the bare minimum, while taking care to present formally precise results after a few introductory slides.
- I will also show some examples that will hopefully clarify the concepts presented.



## What is TDA, in practice?

---

Informally speaking, TDA can be seen as a particular **stable topological operator** that transforms objects belonging to a high-dimensional space into simpler descriptors. Stability is with respect to **appropriate metrics: these metrics are a key ingredient in the theory**. Let us see how TDA can transform a closed curve into a simpler object.



## Persistence diagrams have good properties

---

The operator that computes the persistence diagrams thus allows the projecting of high-dimensional topological spaces onto low-dimensional topological spaces. In other words, **TDA helps us to select the information that is considered most important in our application.**

The usefulness of this operator depends on the following properties:

- TDA is invariant under homeomorphisms of the domain of the functions we consider (i.e., our data). In other words, **it does not depend on the parameterization we are using.**
- **It can be approximated easily** and quite quickly using algorithmic procedures.
- **It admits stability results** with respect to noise.

## When was TDA born?

---

The history of TDA begins (ante litteram) with these papers:

- P. Frosini, *A distance for similarity classes of submanifolds of a Euclidean space*, Bulletin of the Australian Mathematical Society, 42, 3 (1990), 407-416.
- P. Frosini, *Measuring shapes by size functions*, Proc. of SPIE, Intelligent Robots and Computer Vision X: Algorithms and Techniques, Boston, MA 1607 (1992), 122-133.

For more details, see

- H. Edelsbrunner and D. Morozov, *Persistent homology: theory and practice*, In: *European Congress of Mathematics*, pages 31—50, European Mathematical Society, Zürich (2013).
- [https://en.wikipedia.org/wiki/Topological\\_data\\_analysis](https://en.wikipedia.org/wiki/Topological_data_analysis)

- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?
- 4 Group equivariant non-expansive operators
- 5 Building linear and nonlinear GENEOS
- 6 How can we use GENEOS in applications?

## The definition of $d_G$

---

Let  $X$  and  $G$  be a **compact metric space** and a **subgroup of the group**  $\text{Homeo}(X)$  of all homeomorphisms from  $X$  to  $X$ , respectively. If  $\varphi_1, \varphi_2$  are two continuous and bounded functions from  $X$  to  $\mathbb{R}$  we can consider the value  $\inf_{g \in G} \|\varphi_1 - \varphi_2 \circ g\|_\infty$ . This value is called the **natural pseudo-distance**  $d_G(\varphi_1, \varphi_2)$  between  $\varphi_1$  and  $\varphi_2$  with respect to the group  $G$ .

We endow both  $C^0(X, \mathbb{R})$  and  $G$  with the topology of uniform convergence, so that  $G$  becomes a **topological group acting continuously on  $C^0(X, \mathbb{R})$  by composition on the right**.

## The definition of $d_G$

---

If  $G$  is the trivial group  $\text{Id}$ , then  $d_G$  is the max-norm distance  $\|\varphi_1 - \varphi_2\|_\infty$ . Moreover, if  $G_1$  and  $G_2$  are subgroups of  $\text{Homeo}(X)$  and  $G_1 \subseteq G_2$ , then

$$d_{\text{Homeo}(X)}(\varphi_1, \varphi_2) \leq d_{G_2}(\varphi_1, \varphi_2) \leq d_{G_1}(\varphi_1, \varphi_2) \leq \|\varphi_1 - \varphi_2\|_\infty$$

for every  $\varphi_1, \varphi_2 \in C^0(X, \mathbb{R})$ .

We usually restrict  $d_G$  to  $\Phi \times \Phi$ , where  $\Phi$  is a bounded subset of  $C^0(X, \mathbb{R})$ .

## Our ground truth: the natural pseudo-distance $d_G$

---

**The natural pseudo-distance  $d_G$  is our ground truth:** it describes the differences that the observer can perceive between the measurements in  $\Phi$  with respect to the equivalence expressed by the group  $G$ .

**A possible objection:** *“The use of the concept of homeomorphism makes the natural pseudo-distance  $d_G$  difficult to apply. For example, in shape comparison two similar objects can be non-homeomorphic, hence this pseudo-metric cannot be applied to real problems.”*

## A possible objection

---

**Answer:** the homeomorphisms do not concern the “objects” but the space  $X$  where the measurements are made.

- For example, if we are interested in grey level images, the domain of our measurements can be modelled as the real plane and each image can be represented as a function from  $\mathbb{R}^2$  to  $\mathbb{R}$ . Therefore, the space  $X$  is not given by the (possibly non-homeomorphic) objects displayed in the pictures, but by the topological space  $\mathbb{R}^2$ .
- If we make two CAT scans, the topological space  $X$  is always given by a helix turning many times around a body, and no requirement is made about the topology of such a body.

In other words, it is usually legitimate to assume that the topological space  $X$  is determined only by the measuring instrument we are using to get our measurements.



## $d_G$ and critical values: manifolds

---

### Theorem

Assume that  $\mathcal{M}$  is a closed manifold of class  $C^1$  and that  $\varphi_1, \varphi_2 : \mathcal{M} \rightarrow \mathbb{R}$  are  $C^1$ . Set  $d := d_{\text{Homeo}(\mathcal{M})}(\varphi_1, \varphi_2)$ . Then a positive integer  $k$  exists such that one of the following properties holds:

- 1)  $k$  is odd and  $kd$  is the distance between a critical value of  $\varphi_1$  and a critical value of  $\varphi_2$ ;
- 2)  $k$  is even and  $kd$  is either the distance between two critical values of  $\varphi_1$  or the distance between two critical values of  $\varphi_2$ .

- P. Donatini, P. Frosini, *Natural pseudodistances between closed manifolds*, Forum Mathematicum, vol. 16 (2004), n. 5, 695-715.

## $d_G$ and critical values: surfaces

### Theorem

Assume that  $\mathcal{S}$  is a closed surface of class  $C^1$  and that  $\varphi_1, \varphi_2 : \mathcal{S} \rightarrow \mathbb{R}$  are  $C^1$ . Set  $d := d_{\text{Homeo}(\mathcal{S})}(\varphi_1, \varphi_2)$ . Then at least one of the following properties holds:

- 1)  $d$  is the distance between a crit. value of  $\varphi_1$  and a crit. value of  $\varphi_2$ ;
- 2)  $d$  is half the distance between two critical values of  $\varphi_1$ ;
- 3)  $d$  is half the distance between two critical values of  $\varphi_2$ ;
- 4)  $d$  is one third of the distance between a critical value of  $\varphi_1$  and a critical value of  $\varphi_2$ .

- P. Donatini, P. Frosini, *Natural pseudodistances between closed surfaces*, Journal of the European Mathematical Society, vol. 9 (2007), n. 2, 331–353.

## $d_G$ and critical values: curves

### Theorem

Assume that  $\mathcal{C}$  is a closed curve of class  $C^1$  and that  $\varphi_1, \varphi_2 : \mathcal{C} \rightarrow \mathbb{R}$  are  $C^1$ . Set  $d := d_{\text{Homeo}(\mathcal{C})}(\varphi_1, \varphi_2)$ . Then at least one of the following properties holds:

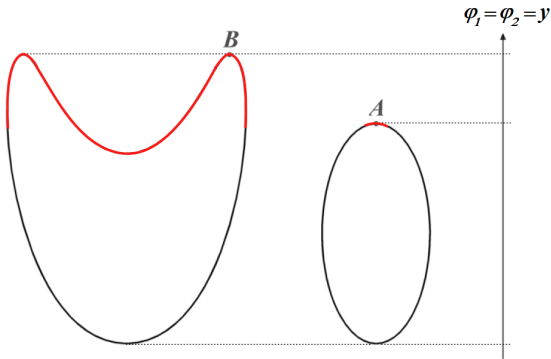
- a)  $d$  is the distance between a crit. value of  $\varphi_1$  and a crit. value of  $\varphi_2$ ;
- b)  $d$  is half the distance between two critical values of  $\varphi_1$ ;
- c)  $d$  is half the distance between two critical values of  $\varphi_2$ .

- P. Donatini, P. Frosini, *Natural pseudo-distances between closed curves*, Forum Mathematicum, vol. 21 (2009), Issue 6, 981–999.

The last theorem is sharp, as shown by the following examples.

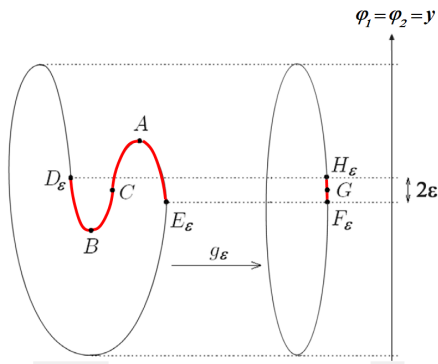
## $d_G$ and critical values: curves

Let us consider the two embeddings of  $S^1$  in  $\mathbb{R}^2$  represented in the following figure. The ordinate  $y$  defines two filtering functions  $\varphi_1, \varphi_2$  on  $S^1$ . In this case  $d_{\text{Homeo}(S^1)}(\varphi_1, \varphi_2) = |\varphi_1(A) - \varphi_2(B)|$  is the distance between a critical value of  $\varphi_1$  and a critical value of  $\varphi_2$ .



## $d_G$ and critical values: curves

Let us consider the two embeddings of  $S^1$  in  $\mathbb{R}^2$  represented in the following figure. The ordinate  $y$  defines two filtering functions  $\varphi_1, \varphi_2$  on  $S^1$ . In this case  $d_{\text{Homeo}(S^1)}(\varphi_1, \varphi_2) = \frac{1}{2}|\varphi_1(A) - \varphi_1(B)|$  is half the distance between two critical values of  $\varphi_1$ .



## Optimal homeomorphisms

Assume that  $X$  is a compact topological space and  $\varphi_1, \varphi_2 : X \rightarrow \mathbb{R}$  are continuous functions. Let  $G$  be a subgroup of  $\text{Homeo}(X)$ . We say that a homeomorphism  $g \in G$  is **optimal** in  $G$  for  $(\varphi_1, \varphi_2)$  if  $\|\varphi_1 - \varphi_2 \circ g\|_\infty = d_G(\varphi_1, \varphi_2)$ . The following result holds for optimal homeomorphisms.

### Theorem

*Assume that  $\mathcal{M}$  is a  $C^1$  closed manifold and that  $\varphi_1, \varphi_2 : \mathcal{M} \rightarrow \mathbb{R}$  are of class  $C^1$ . If an optimal homeomorphism  $g \in \text{Homeo}(\mathcal{M})$  for  $(\varphi_1, \varphi_2)$  exists, then  $d_{\text{Homeo}(\mathcal{M})}(\varphi_1, \varphi_2)$  is the distance between a critical value of  $\varphi_1$  and a critical value of  $\varphi_2$ .*

## A KEY OBSERVATION

---

The natural pseudo-distance can be seen as a ground truth in our model but, unfortunately, computing this pseudo-metric is difficult, since the group  $G$  is often too large.

How can we get information about the natural pseudo-distance  $d_G$ ?

We need a new idea: **persistent homology**.

- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?**
- 4 Group equivariant non-expansive operators
- 5 Building linear and nonlinear GENEOS
- 6 How can we use GENEOS in applications?

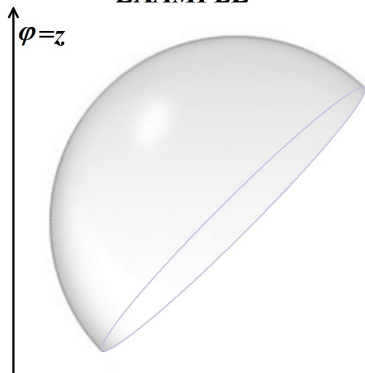


## What is persistent homology?

---

If  $\varphi : X \rightarrow \mathbb{R}$  is a continuous function, we can consider the sublevel sets  $X_t := \{x \in X : \varphi(x) \leq t\}$ . When  $t$  varies we see the birth and death of  $k$ -dimensional holes.

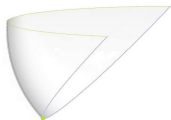
### *EXAMPLE*



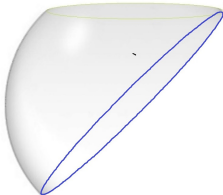
## What is persistent homology?

If  $\varphi : X \rightarrow \mathbb{R}$  is a continuous function, we can consider the sublevel sets  $X_t := \{x \in X : \varphi(x) \leq t\}$ . When  $t$  varies we see the birth and death of  $k$ -dimensional holes.

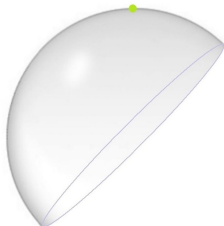
### *EXAMPLE*



*No 1-dimensional hole*



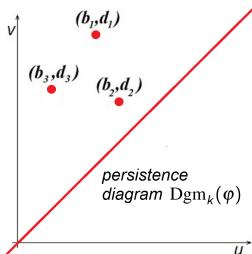
*Birth of a 1-dimensional hole*



*Death of the 1-dimensional hole*

## What is persistent homology?

In plain words, the **persistence diagram**  $\text{Dgm}_k(\varphi)$  in degree  $k$  of  $\varphi$  is the collection of the pairs  $(b_i, d_i)$  where  $b_i$  and  $d_i$  are the times of birth and death of the  $i$ -th hole of dimension  $k$ .



The points of the persistence diagram are endowed with multiplicities.

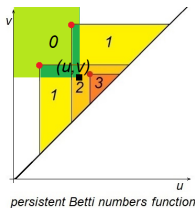
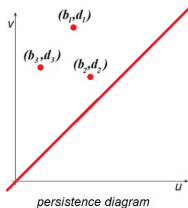
Each point of the diagonal  $u = v$  is assumed to be a point of the persistence diagram, endowed with infinite multiplicity.

# What are the Persistent Betti Numbers Functions?

Persistence diagrams are not quite suitable for statistical purposes, because no **good** definition of **average** of persistence diagrams exists. Persistent Betti numbers functions are more suitable for statistics.

## Definition

The  $k$ -th **persistent Betti numbers function**  $\beta_k(u, v)$  is the number of holes of dimension  $k$  whose time of birth is smaller than  $u$  and whose time of death is greater than  $v$ .



*The Persistent Betti Numbers Function at  $(u, v)$  is the number of points of the persistence diagram that are placed in the green region.*

# What are the Persistent Betti Numbers Functions?

More precisely:

## Definition

Let  $\varphi : X \rightarrow \mathbb{R}$  be a continuous function. If  $u, v \in \mathbb{R}$  and  $u < v$ , we can consider the inclusion  $i$  of  $X_u$  into  $X_v$ . Such an inclusion induces a homomorphism  $i^* : H_k(X_u) \rightarrow H_k(X_v)$  between the homology groups of  $X_u$  and  $X_v$  in degree  $k$ . The group  $PH_k^\varphi(u, v) := i^*(H_k(X_u))$  is called the  **$k$ -th persistent homology group** with respect to the function  $\varphi : X \rightarrow \mathbb{R}$ , computed at the point  $(u, v)$ . The rank  $r_k(\varphi)(u, v)$  of this group is called the  **$k$ -th persistent Betti numbers function** with respect to the function  $\varphi : X \rightarrow \mathbb{R}$ , computed at the point  $(u, v)$ .

**The average of persistent Betti numbers functions can be trivially defined as the usual average of real-valued functions.**

## What are persistent Betti numbers functions?

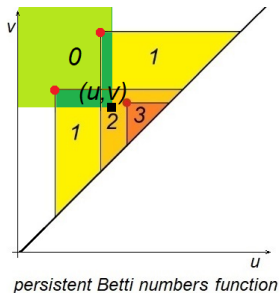
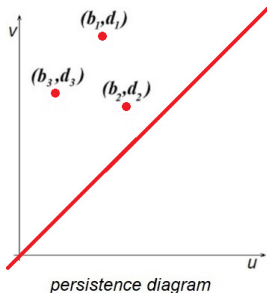
---

The use of averages of persistent Betti numbers functions in degree 0 firstly appeared in the papers

- P. Donatini, P. Frosini, A. Lovato, *Size functions for signature recognition*, Proceedings of SPIE, Vision Geometry VII, vol. 3454 (1998), 178–183.
- M. Ferri, P. Frosini, A. Lovato, C. Zambelli, *Point selection: A new comparison scheme for size functions (With an application to monogram recognition)*, Proceedings Third Asian Conference on Computer Vision, Lecture Notes in Computer Science 1351, vol. 1, R. Chin, T. Pong (editors) Springer-Verlag, Berlin Heidelberg (1998), 329–337.

These papers also introduce the first vectorization method for TDA: each point of the considered persistence diagram is replaced with a suitable function (usually a Gaussian function centered at that point).

# What are persistent Betti numbers functions?

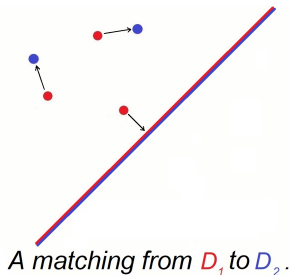


*The Persistent Betti Numbers Function at  $(u, v)$  is the number of points of the persistence diagram that are placed in the green region.*

If we use Čech homology, persistence diagrams are equivalent to persistent Betti numbers functions:

- Patrizio Frosini, Claudia Landi, *Size theory as a topological tool for computer vision*, Pattern Recognition And Image Analysis, vol. 9 (4) (1999), 596-603.

## Comparison of persistent Betti numbers functions



Persistence diagrams (and hence persistent Betti numbers functions) can be compared by means of the **bottleneck distance**  $d_{\text{match}}$ . The bottleneck distance  $d_{\text{match}}(D_1, D_2)$  between two persistence diagrams  $D_1, D_2$  is the minimum cost of moving the points of  $D_1$  to the points of  $D_2$ , where the cost of moving each point is given by the **max-norm distance** in  $\mathbb{R}^2$ . Moving a point to the diagonal is equivalent to delete it.



## The PBNFs give lower bounds for $d_G$

A fundamental property of the metric  $d_{\text{match}}$  is its **stability**:

### Theorem

If  $k$  is a natural number and  $\varphi_1, \varphi_2 \in C^0(X, \mathbb{R})$ , then

$$d_{\text{match}}(\text{Dgm}_k(\varphi_1), \text{Dgm}_k(\varphi_2)) \leq d_{\text{Homeo}(X)}(\varphi_1, \varphi_2) \leq \|\varphi_1 - \varphi_2\|_{\infty}.$$

- D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, *Stability of persistence diagrams*, *Discr. Comput. Geom.*, 37:103–120, 2007.
- M. d'Amico, P. Frosini and C. Landi (2005), *Natural pseudo-distance and optimal matching between reduced size functions*, Technical Report no. 66, DISMI, University of Modena and Reggio Emilia, Italy. (case  $k = 0$ )

## Limitations of TDA (1)

---

So the theory does what it promises: **TDA gives us a lower bound for the natural pseudo-distance  $d_G$  when  $G = \text{Homeo}(X)$ :**

$$d_{\text{match}}(\text{Dgm}_k(\varphi_1), \text{Dgm}_k(\varphi_2)) \leq d_{\text{Homeo}(X)}(\varphi_1, \varphi_2).$$

By definition, this is also a lower bound for  $d_G$  when  $G \subset \text{Homeo}(X)$ , but it is not a good lower bound, in general.

**How can we get good lower bounds for  $d_G$  in the general case?**

## Limitations of TDA (2)

---

When TDA was born, it was data-centric, but in many cases, data analysis is highly dependent on the observer (think, for example, of judging a film).

**How can we adapt TDA according to this remark?**

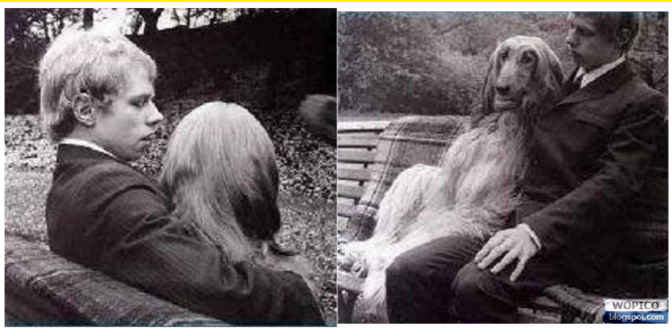


GUIDO MORETTI CON UNA SUA SCULTURA (LA STESSA, DA TRE PUNTI DI VISTA DIVERSI)

## A topological theory for spaces of observers

---

We need to develop a **topological** theory for spaces of **observers**. The main question is not *“What is the shape of data?”* but *“What is the shape of observers?”*



- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?
- 4 Group equivariant non-expansive operators**
- 5 Building linear and nonlinear GENEOS
- 6 How can we use GENEOS in applications?

## Representing observers as equivariant operators

---

Observers are structures able to change data into other data, and usually do that by respecting some data equivalences, i.e., by commuting with some transformations.

As a first approximation, observers can be represented as group equivariant operators (GEOs).

In this talk we will illustrate some results on the theory of **Group Equivariant Non-Expansive Operators (GENEOs)**.

Why “non-expansive”?

Because

- 1 **observers are often assumed to simplify the metric structure of data in order to produce meaningful interpretations;**
- 2 **non-expansiveness guarantees good topological properties.**

# How could we represent observers?

nature  
machine intelligence

ARTICLES

<https://doi.org/10.1038/s42256-019-0087-3>

## Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning

Mattia G. Bergomi<sup>1</sup>, Patrizio Frosini<sup>2,3\*</sup>, Daniela Giorgi<sup>4</sup> and Nicola Quercioli<sup>2,3</sup>

We provide a general mathematical framework for group and set equivariance in machine learning. We define group equivariant non-expansive operators (GENEOs) as maps between function spaces associated with groups of transformations. We study the topological and metric properties of the space of GENEOs to evaluate their approximating power and set the basis for general strategies to initialize and compose operators. We define suitable pseudo-metrics for the function spaces, the equivariance groups and the set of non-expansive operators. We prove that, under suitable assumptions, the space of GENEOs is compact and convex. These results provide fundamental guarantees in a machine learning perspective. By considering isometry-equivariant non-expansive operators, we describe a simple strategy to select and sample operators. Thereafter, we show how selected and sampled operators can be used both to perform classical metric learning and to inject knowledge in artificial neural networks.

<https://rdcu.be/bP6HV>

## All begins with the space of admissible functions

---

Let  $X$  be a nonempty set. Let  $\Phi$  be a topological subspace of the set  $\mathbb{R}_b^X$  of all bounded functions  $\varphi$  from  $X$  to  $\mathbb{R}$ , endowed with the topology induced by the metric

$$D_\Phi(\varphi_1, \varphi_2) := \|\varphi_1 - \varphi_2\|_\infty.$$

We can see  $X$  as the space where we can make our measurements, and  $\Phi$  as the space of all possible measurements. We will say that  $\Phi$  is the **set of admissible functions**. In other words,  **$\Phi$  is the set of all functions from  $X$  to  $\mathbb{R}$  that can be produced by our measuring instruments (or by other observers)**. For example, a gray-level image can be represented as a function from the real plane to the interval  $[0,1]$  (in this case  $X = \mathbb{R}^2$ ).



## Perception pairs

---

Let us consider a group  $G$  of bijections  $g : X \rightarrow X$  such that  $\varphi \in \Phi \implies \varphi \circ g \in \Phi$  for every  $\varphi \in \Phi$ . We say that  $(\Phi, G)$  is a **perception pair**.

The choice of a perception pair states which data can be considered as legitimate measurements (the functions in  $\Phi$ ) and which group represents the equivalence between data (the group  $G$ ).

To proceed, we need to introduce suitable topologies on  $X$  and  $G$ . Before doing that, we recall that the **initial topology**  $\tau_{\text{in}}$  on  $X$  with respect to  $\Phi$  is the coarsest topology on  $X$  such that every function  $\varphi$  in  $\Phi$  is continuous.

## A pseudo-metric on $X$

---

Let us define on  $X$  the pseudo-metric

$$D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|.$$

$D_X$  induces a topology  $\tau_{D_X}$  on  $X$ .

The use of  $D_X$  implies that we can distinguish two points only if a measurement exists, taking those points to different values.

### Proposition

*The topology  $\tau_{D_X}$  is finer than the initial topology  $\tau_{\text{in}}$  on  $X$  with respect to  $\Phi$ . If  $\Phi$  is totally bounded, then  $\tau_{D_X}$  coincides with  $\tau_{\text{in}}$ .*

## A pseudo-metric on $X$

---

The following properties are of use in our model.

### Proposition

*Every function in  $\Phi$  is non-expansive, and hence continuous.*

### Proposition

*If  $\Phi$  is compact and  $X$  is complete, then  $X$  is compact.*

In the following, we will usually assume that  $\Phi$  is compact and  $X$  is complete (and hence compact).

## An interesting outcome: each bijection is an isometry

- $\text{Bij}_\Phi(X) = \{\text{bijections } g: X \rightarrow X \text{ s.t. } \Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\};$
- $\text{Homeo}_\Phi(X) = \{\text{homeomorphisms } g: X \rightarrow X \text{ s.t. } \Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\};$
- $\text{Iso}_\Phi(X) = \{\text{isometries } g: X \rightarrow X \text{ s.t. } \Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\}.$

### Proposition

$$\text{Bij}_\Phi(X) = \text{Homeo}_\Phi(X) = \text{Iso}_\Phi(X).$$

## A pseudo-metric on $G$

---

Let us now focus our attention on a subgroup  $G$  of  $\text{Homeo}_\Phi(X)$ .

We can define a pseudo-metric  $D_G$  on  $G$  by setting

$$D_G(g_1, g_2) := \sup_{\varphi \in \Phi} D_\Phi(\varphi \circ g_1, \varphi \circ g_2).$$

### Proposition

*$G$  is a topological group with respect to  $D_G$  and the action of  $G$  on  $\Phi$  by right composition is continuous.*

### Proposition

*If  $\Phi$  is compact and  $G$  is complete, then  $G$  is compact.*

## GEOs and GENEOS

---

Each pair  $(\Phi, G)$  with  $G \subseteq \text{Homeo}_\Phi(X)$  is called a **perception pair**.

Let us assume that two perception pairs  $(\Phi, G)$ ,  $(\Psi, H)$  are given, and fix a group homomorphism  $T : G \rightarrow H$ .

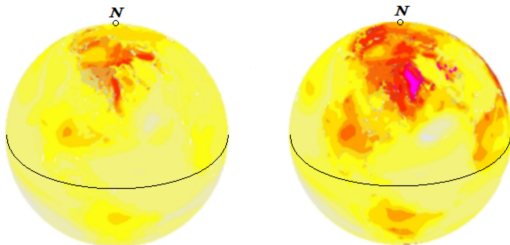
Each function  $F : \Phi \rightarrow \Psi$  such that  $F(\varphi \circ g) = F(\varphi) \circ T(g)$  for every  $\varphi \in \Phi, g \in G$  is called a **Group Equivariant Operator (GEO)** associated with the homomorphism  $T$ .

If  $F$  is also non-expansive (i.e.,  $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$  for every  $\varphi_1, \varphi_2 \in \Phi$ ), then  $F$  is called a **Group Equivariant Non-Expansive Operator (GENEO)** associated with the homomorphism  $T$ .

## An example of GENE0

---

Let us assume to be interested in the comparison of the **distributions of temperatures** on a sphere, taken at two different times:



Let us also assume that only two opposite points  $N, S$  can be localized on the sphere.

## An example of GENEIO

---

Let us introduce two perception pairs  $(\Phi, G), (\Psi, H)$  by setting

- $X = S^2$
- $\Phi =$  set of 1-Lipschitz functions from  $S^2$  to a fixed interval  $[a, b]$
- $G =$  group of rotations of  $S^2$  around the axis  $N - S$

and

- $Y =$  the equator  $S^1$  of  $S^2$
- $\Psi =$  set of 1-Lipschitz functions from  $S^1$  to  $[a, b]$
- $H =$  group of rotations of  $S^1$



## An example of GENEIO

---

This is a simple example of GENEIO from  $(\Phi, G)$  to  $(\Psi, H)$ :

- $T(g)$  is the rotation  $h \in H$  of the equator  $S^1$  that is induced by the rotation  $g$  of  $S^2$ , for every  $g \in G$ .
- $F(\varphi)$  is the function  $\psi$  that takes each point  $y$  belonging to the equator  $S^1$  to the average of the temperatures along the meridian containing  $y$ , for every  $\varphi \in \Phi$ ;

We can easily check that  $F$  verifies the properties defining the concept of group equivariant non-expansive operator with respect to the isomorphism  $T : G \rightarrow H$ .

In plain words, our GENEIO simplifies the data by transforming “temperature distributions on the earth” into “temperature distributions on the equator”.

## Two key results (and two good news for applications)

Let us assume that a homomorphism  $T : G \rightarrow H$  has been fixed.

Let us define a metric  $D_{\text{GENEO}}$  on  $\text{GENEO}((\Phi, G), (\Psi, H))$  by setting

$$D_{\text{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} D_{\Psi}(F_1(\varphi), F_2(\varphi)).$$

### Theorem

*If  $\Phi$  and  $\Psi$  are **compact**, then  $\text{GENEO}((\Phi, G), (\Psi, H))$  is **compact** with respect to  $D_{\text{GENEO}}$ .*

### Theorem

*If  $\Psi$  is **convex**, then  $\text{GENEO}((\Phi, G), (\Psi, H))$  is **convex**.*

## Two key observations (1)

---

- While the space of data is often non-convex (and hence averaging data does not make sense), the assumption of convexity of  $\Psi$  implies the convexity of the space of observers and allows us to consider the **“average of observers”**.



## Two key observations (2)

---

- Our main goal is to develop a good geometric and compositional theory to approximate an ideal observer. In our model, “approximate an observer” means to look for a GENEIO  $F$  that minimizes a suitable “cost function”  $c(F)$ . The cost function quantifies the error that is committed by taking the GENEIO  $F$  instead of the ideal observer. Since the space of GENEIOs is compact and convex (under the assumption that the data spaces are compact and convex), if the cost function  $c(F)$  is strictly convex we have that there is one and only one GENEIO that best approximates the ideal observer.

- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?
- 4 Group equivariant non-expansive operators
- 5 Building linear and nonlinear GNEOs**
- 6 How can we use GNEOs in applications?

# How can we build linear and nonlinear GENEOS?



## On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions

Francesco Conti<sup>1,2</sup>, Patrizio Frosini<sup>3,4,5,6\*</sup> and Nicola Quercioli<sup>3,7</sup>

<sup>1</sup> Department of Mathematics, University of Pisa, Pisa, Italy, <sup>2</sup> Institute of Information Science and Technologies "A. Faedo", National Research Council of Italy (CNR), Pisa, Italy, <sup>3</sup> Department of Mathematics, University of Bologna, Bologna, Italy, <sup>4</sup> Alma Mater Research Center on Applied Mathematics, University of Bologna, Bologna, Italy, <sup>5</sup> Alma Mater Research Institute for Human-Centered Artificial Intelligence, University of Bologna, Bologna, Italy, <sup>6</sup> Research Centre on Electronic Systems for the Information and Communication Technology, University of Bologna, Bologna, Italy, <sup>7</sup> ENEA Centro Ricerche Bologna, Bologna, Italy

<https://www.frontiersin.org/articles/10.3389/frai.2022.786091/full>

## Elementary methods to build GNEOs

### Proposition (Composition)

*If  $F_1 \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T_1 : G \rightarrow H$  and  $F_2 \in \text{GENEO}((\Psi, H), (\chi, K))$  w.r.t.  $T_2 : H \rightarrow K$  then  $F_2 \circ F_1 \in \text{GENEO}((\Phi, G), (\chi, K))$  w.r.t.  $T_2 \circ T_1 : G \rightarrow K$ .*

### Proposition (Image by a 1-Lipschitz function)

*If  $F_1, \dots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T : G \rightarrow H$ ,  $L$  is a 1-Lipschitz map from  $\mathbb{R}^n$  to  $\mathbb{R}$ , and  $L^*(F_1, \dots, F_n)(\Phi) \subseteq \Psi$  (where  $L^*$  is the map induced by  $L$ ), then  $L^*(F_1, \dots, F_n) \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T$ .*

The next three statements follow from the last proposition.

## Elementary methods to build GNEOs

### Proposition (LATTICE OF GNEOS)

If  $F_1, \dots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T : G \rightarrow H$  and  $\max(F_1, \dots, F_n)(\Phi), \min(F_1, \dots, F_n)(\Phi) \subseteq \Psi$ , then  $\max(F_1, \dots, F_n), \min(F_1, \dots, F_n) \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T$ .

### Proposition (Translation)

If  $F \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T : G \rightarrow H$ , and  $F_b(\Phi) \subseteq \Psi$  for  $F_b(\varphi) := F(\varphi) - b$ , then  $F_b \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T$ .

### Proposition (Convex combination)

If  $F_1, \dots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T : G \rightarrow H$ ,  $(a_1, \dots, a_n) \in \mathbb{R}^n$  con  $\sum_{i=1}^n |a_i| \leq 1$  and  $F_\Sigma(\Phi) \subseteq \Psi$  for  $F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi)$ , then  $F_\Sigma \in \text{GENEO}((\Phi, G), (\Psi, H))$  w.r.t.  $T$ .



## Permutant measures

Let us consider the set  $\Phi = \mathbb{R}^X \cong \mathbb{R}^n$  of all functions from a **finite** set  $X = \{x_1, \dots, x_n\}$  to  $\mathbb{R}$ , and a subgroup  $G$  of the group  $\text{Bij}(X)$  of all permutations of  $X$ .

### Definition

A finite (signed) measure  $\mu$  on  $\text{Bij}(X)$  is called a **permutant measure** with respect to  $G$  if every subset  $H$  of  $\text{Bij}(X)$  is measurable and  $\mu$  is invariant under the conjugation action of  $G$  (i.e.,  $\mu(H) = \mu(gHg^{-1})$  for every  $g \in G$ ).

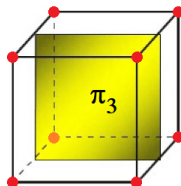
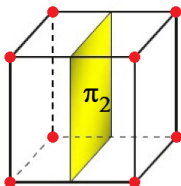
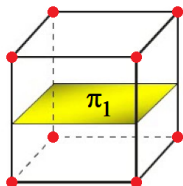
### Proposition

If  $\mu$  is a permutant measure with respect to  $G$ , then the map  $F_\mu : \mathbb{R}^X \rightarrow \mathbb{R}^X$  defined by setting  $F_\mu(\varphi) := \sum_{h \in \text{Bij}(X)} \varphi h^{-1} \mu(h)$  is a linear GEO. If  $\sum_{h \in \text{Bij}(X)} |\mu(h)| \leq 1$ , then  $F_\mu(\varphi)$  is a GENEIO.

## An example of permutant measure

Let us consider the set  $X$  of the vertices of a cube in  $\mathbb{R}^3$ , and the group  $G$  of the orientation-preserving isometries of  $\mathbb{R}^3$  that take  $X$  to  $X$ . Let  $\pi_1, \pi_2, \pi_3$  be the three planes that contain the center of mass of  $X$  and are parallel to a face of the cube. Let  $h_i : X \rightarrow X$  be the orthogonal symmetry with respect to  $\pi_i$ , for  $i \in \{1, 2, 3\}$ .

We can now define a permutant measure  $\mu$  on the group  $\text{Bij}(X)$  by setting  $\mu(h_1) = \mu(h_2) = \mu(h_3) = c$ , where  $c$  is a positive real number, and  $\mu(h) = 0$  for any  $h \in \text{Bij}(X)$  with  $h \notin \{h_1, h_2, h_3\}$ .



## Building GENEOS by permutant measures

---

It is interesting to observe that the set  $\text{PM}(G)$  of permutant measures with respect to  $G$  is a lattice. Indeed, if  $\mu_1, \mu_2 \in \text{PM}(G)$ , then the measures  $\mu', \mu''$  on  $\text{Bij}(X)$ , respectively defined by setting  $\mu'(h) := \min\{\mu_1(h), \mu_2(h)\}$  and  $\mu''(h) := \max\{\mu_1(h), \mu_2(h)\}$ , still belong to  $\text{PM}(G)$ . Moreover, if  $\mu \in \text{PM}(G)$  then  $|\mu| \in \text{PM}(G)$ . Furthermore,  $\text{PM}(G)$  is closed under linear combination. Therefore,  $\text{PM}(G)$  has a natural structure of real vector space.

We stress that when the group  $G$  becomes larger and larger the lattice  $\text{PM}(G)$  becomes smaller and smaller.

In other words, the theory of permutant measures becomes more and more useful as the groups get bigger and bigger.

## Building GENEOS by permutant measures

**The method for building GENEOS based on permutant measures can be generalized by replacing the arithmetic mean with another symmetric function.** We can indeed show that when a symmetric function and a **permutant** for the equivariance group  $G$  are available, we can easily build a (non-linear) GENEOS with respect to  $G$ . First of all, let us recall the concept of *permutant*, which is equivalent to the one of *permutant measure uniformly distributed on its support*.

### Definition

We say that a subset  $H \subseteq \text{Bij}(X)$  is a **permutant** for  $G$  if either  $H = \emptyset$  or  $gHg^{-1} = H$  for every  $g \in G$ .

Note that a subset  $H$  of  $\text{Bij}(X)$  is a permutant for  $G$  if and only if  $H$  is a union of orbits for the conjugation action of  $G$  on  $\text{Bij}(X)$ .

## Building GENEOS by permutant measures

Let  $\mathcal{S} : \mathbb{R}^n \rightarrow \mathbb{R}$  be a symmetric function. If  $H = \{h_i\}_{i=1}^n$  is a non-empty permutant for  $G \subseteq \text{Bij}_\Phi(X)$ , then we can define an operator  $\mathcal{S}_H : \Phi \rightarrow \mathbb{R}_b^X$  by setting, for any  $\varphi \in \Phi$ ,

$$\mathcal{S}_H(\varphi) := \mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n),$$

where  $\mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) := \mathcal{S}((\varphi \circ h_1)(x), \dots, (\varphi \circ h_n)(x))$  for every  $x \in X$ .

### Proposition

$\mathcal{S}_H$  is a GEO from  $(\Phi, G)$  to  $(\mathbb{R}_b^X, G)$  with respect to the identity homomorphism  $\text{id}_G : G \rightarrow G$ . If the restriction of  $\mathcal{S}$  to  $\text{Im}(\Phi)^n$  is non-expansive, then  $\mathcal{S}_H$  is a **GENEO** from  $(\Phi, G)$  to  $(\mathbb{R}_b^X, G)$  with respect to  $\text{id}_G$ .

# How can we represent linear GENEOS?


---

Annals of Mathematics and Artificial Intelligence

<https://doi.org/10.1007/s10472-022-09830-1>

---

## On the finite representation of linear group equivariant operators via permutant measures

Giovanni Bocchi<sup>1</sup> · Stefano Botteghi<sup>2</sup> · Martina Brasini<sup>2</sup> · Patrizio Frosini<sup>2</sup>  · Nicola Quercioli<sup>3</sup>

Accepted: 26 December 2022

© The Author(s) 2023

<https://rdcu.be/c5Obw>

## Representation Theorem for linear GENEOS

The following theorem strengthens our previous result about building linear GENEOS via permutant measures.

### Theorem (Representation Theorem for linear GENEOS)

*Let us assume that  $G \subseteq \text{Bij}(X)$  transitively acts on the finite set  $X$  and that  $F$  is a map from  $\mathbb{R}^X$  to  $\mathbb{R}^X$ . The map  $F$  is a linear GENEOS from  $\mathbb{R}^X$  to  $\mathbb{R}^X$  with respect to the identical homomorphism*

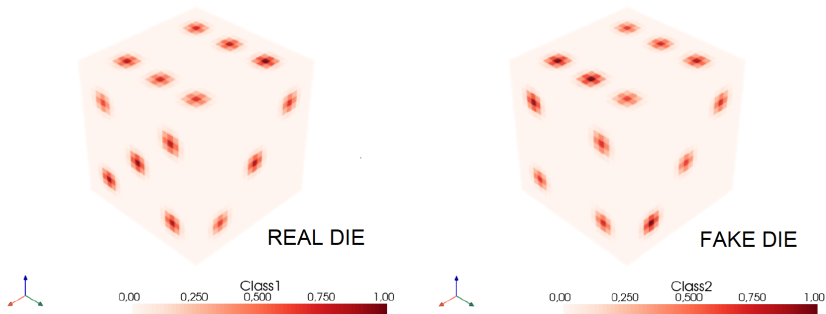
*$\text{Id}_G : g \mapsto g$  **if and only if** a permutant measure  $\mu$  with respect to  $G$  exists, such that  $F(\varphi) = \sum_{h \in \text{Bij}(X)} \varphi h^{-1} \mu(h)$  for every  $\varphi \in \mathbb{R}^X$ , and  $\sum_{h \in \text{Bij}(X)} |\mu(h)| \leq 1$ .*

- 1 Informal introduction to the topic
- 2 Ground truth in TDA: the natural pseudo-distance
- 3 What is persistent homology?
- 4 Group equivariant non-expansive operators
- 5 Building linear and nonlinear GNEOs
- 6 How can we use GNEOs in applications?**



# What happens when we apply GENEOS to our data?

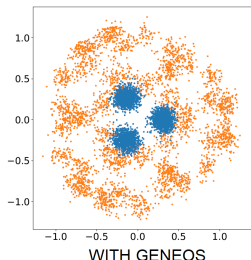
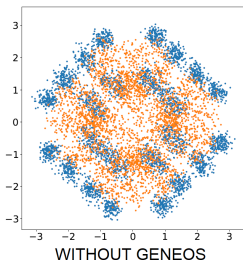
An example of use: comparison between real dice and fake dice.



(Experiment and computations by Giovanni Bocchi)

## What happens to data when we apply GENEOSs?

We produced 10000 dice (a training set of size 7000 and a test set of size 3000), then we applied PCA to the test set and to the test set transformed by a suitable GENEIO, optimized on the training set:



For each die the first two principal components are plotted. Blue points are associated with **real dice**, while orange ones with **fake dice**. The GENEIO we use was built by a convex combination of 3 GENEIOs defined by permutant measures.

## A real application: finding pockets in proteins

---

### **GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection.**

Giovanni Bocchi <sup>1</sup>, Patrizio Frosini <sup>2</sup>, Alessandra Micheletti <sup>1</sup>, Alessandro Pedretti <sup>3</sup>  
Carmen Gratteri <sup>4</sup>, Filippo Lunghini <sup>5</sup>, Andrea Rosario Beccari <sup>5</sup> and Carmine Talarico <sup>5</sup>

<sup>1</sup> Department of Environmental Science and Policy, Università degli Studi di Milano

<sup>2</sup> Department of Mathematics, Università degli Studi di Bologna

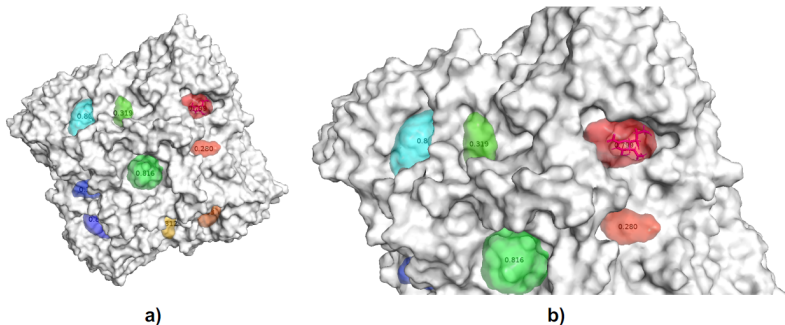
<sup>3</sup> Department of Pharmaceutical Sciences, Università degli Studi di Milano

<sup>4</sup> Dipartimento di Scienze della Salute, Università degli Studi "Magna Græcia di Catanzaro"

<sup>5</sup> Dompé Farmaceutici SpA

<https://arxiv.org/ftp/arxiv/papers/2202/2202.00451.pdf>

## A real application: finding pockets in proteins

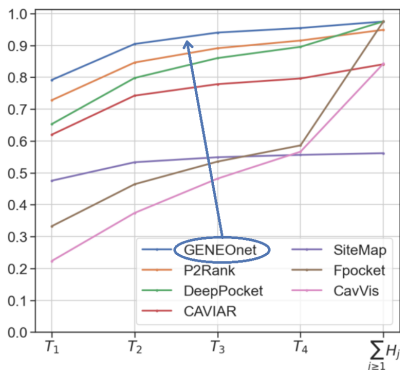


**Model predictions for protein 2QWE.** In Figure a) the global view of the prediction is shown, where different pockets are depicted in different colors and are labelled with their scores. In Figure b) the zoomed of the pocket containing the ligand is shown.

The search for the pockets was carried out by identifying an optimal GENE0 in the convex hull of 8 GENE0s (each focused on a particular property of the pockets).

## A real application: finding pockets in proteins

Here are the results of our experiments:

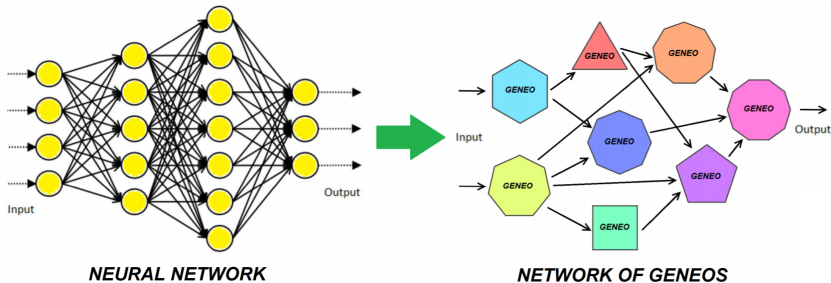


Please note that GENEOnet uses 17 parameters, while a CNN such as DeepPocket requires 665122 parameters.

## The main point in our approach

In perspective, we are looking for a good compositional theory for building **efficient** and **transparent** networks of GENEOS.

Some preliminary experiments suggest that replacing neurons with GENEOS could make deep learning more transparent and interpretable and speed up the learning process.



# GENEOs and Machine Learning

For more details about the use of GENEOS in Machine Learning:



The screenshot shows the EMS (European Mathematical Society) website. The header includes the EMS logo and navigation links: News, Magazine, Membership, Services, Activities, and Society Overview. A 'Login' button is in the top right. The main content area features an article titled 'A new paradigm for artificial intelligence based on group equivariant non-expansive operators' by Alessandra Micheletti, published online first on April 24, 2023. To the left of the article is a thumbnail image of the EMS Magazine cover, which displays the title 'EMS Magazine' and a blue abstract geometric design.

- A. Micheletti, *A new paradigm for artificial intelligence based on group equivariant non-expansive operators*, In: *EMS Magazine*, Online First, 24 April 2023.
- <https://ems.press/content/serial-article-files/27673>

## My current lines of research

---

I am presently studying these problems:

- What is the expected value of the reconstruction error when we apply some suitable denoising GENEOS to 1D signals? (P. Frosini, I. Gridelli, A. Pascucci, *A probabilistic result on impulsive noise reduction in Topological Data Analysis through Group Equivariant Non-Expansive Operators*, <https://arxiv.org/pdf/2202.14021.pdf>.)
- How can we extend the theory of GENEOS to graphs? (F. Ahmad, M. Ferri, P. Frosini, *Generalized Permutants and Graph GENEOS*, <https://arxiv.org/pdf/2206.14798.pdf>.)
- How can we extend the theory of GENEOS to probability spaces of signals? P. Cascarano, P. Frosini, N. Quercioli, A. Saki, *On the geometric and Riemannian structure of the spaces of group equivariant non-expansive operators*, <https://arxiv.org/pdf/2103.02543.pdf>

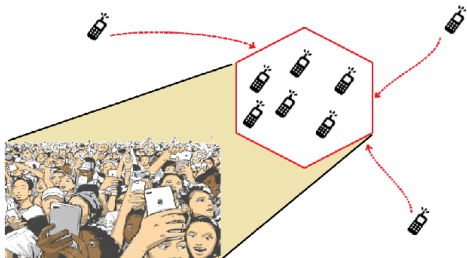


# A current research project

---

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

## Project on GENEOS for 6G



THANKS  
FOR YOUR  
ATTENTION

