On the Role of Group Equivariant Non-Expansive Operators as a Bridge between TDA and Machine Learning

Patrizio Frosini

Department of Computer Science, University of Pisa
patrizio.frosini@unipi.it

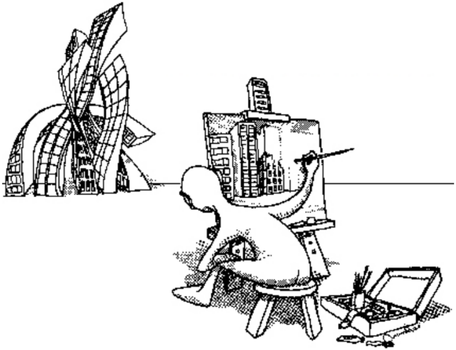
Seminar GEOTOP-A: Applications of geometry and topology,
30 January 2026

Outline

Some epistemological assumptions

Assumption 1: Data are processed by observers

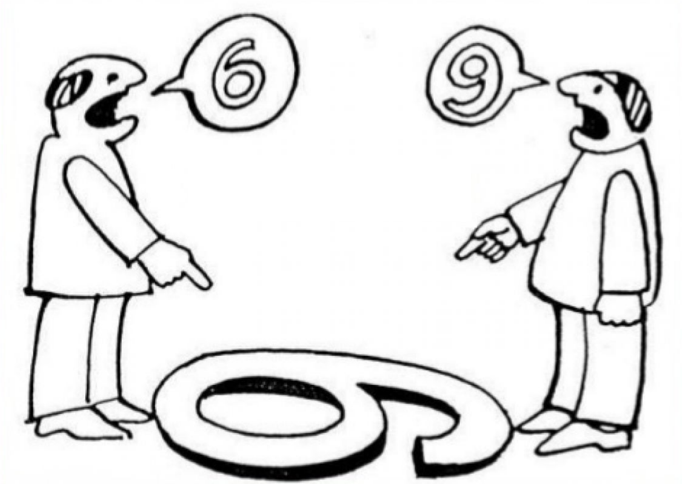
Data have no meaning without an observer to interpret them.



An observer is an agent that transforms data while preserving their symmetries.

Assumption 2: Observers are variables

Data interpretation strongly depends on the chosen observer.



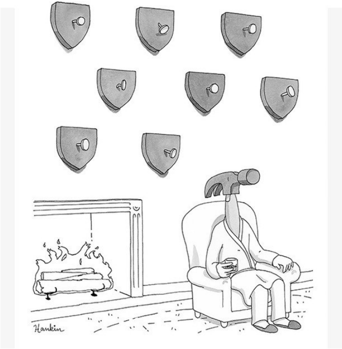
Assumption 3: Observers are important

We are rarely directly interested in the data, but rather in how observers react to their presence.

Consequently, we should focus more on the properties of the observers than on the properties of the data.

Assumption 4: There is no structure in the data

Generally speaking, data lack inherent structure. Instead, the structure of data reflects the observer's own structure.



The shape is not in the data but in the eyes of the observer.

How can we translate these ideas into mathematics?

? ?

Perception spaces and GENEOs

PERCEPTION
SPACE

GROUP
EQUIVARIANT
NON-EXPANSIVE
OPERATOR
(GENEO)

What is a GENEO?

- A Group Equivariant Non-Expansive Operator (GENEO) is a mathematical tool used to approximate observers that act on data.
- The theory of GENEOs is based on the idea that the geometric characteristics of observers significantly influence the interpretation of data.
- In this talk, we will explore the core properties of GENEOs, examine their role in machine learning, and discuss their promising applications in explainable artificial intelligence.

Perception spaces

Recall that a pseudometric is a distance function d satisfying nonnegativity, symmetry, and the triangle inequality, but not necessarily the property $d(x_1, x_2) = 0 \implies x_1 = x_2$.

Definition

Let us consider:

1. A nonempty set Φ endowed with a pseudometric D_Φ.
2. A group $(G, \circ)$ acting on Φ on the left, denoted by $*$. We assume the action is by isometries, i.e., for every $\varphi_1, \varphi_2 \in \Phi$ and every $g \in G$, $D_\Phi(g * \varphi_1, g * \varphi_2) = D_\Phi(\varphi_1, \varphi_2)$.

We call (Φ, G) a perception space.

Perception spaces

The set Φ represents the data we may obtain from our measuring tools (functions, graphs, point clouds, ...). The group G represents the possible transformations of the data that the observer may be interested in.

Some examples:

- $\Phi =$ the set of grey-level images viewed as functions from $\mathbb{R}^2$ to $[0,1]$, endowed with the sup-norm distance;
  $G =$ the group of isometries of the real plane.
- $\Phi =$ the set of all subgraphs of a given graph Γ, endowed with the distance $d_{iso}(\Gamma_1, \Gamma_2) = 0$ if $\Gamma_1 \equiv \Gamma_2$ and 1 otherwise;
  $G =$ the group of graph isomorphisms of Γ.
- $\Phi =$ the set of nonempty compact subsets of the real plane, endowed with the Hausdorff distance;
  $G =$ the group of isometries of the real plane.

A pseudometric on G, induced by D_Φ

If a perception space (Φ, G) is given, then G can be endowed with the pseudometric D_G defined by setting
$D_G(g_1, g_2) := \sup_{\varphi \in \Phi} D_\Phi(g_1 * \varphi, g_2 * \varphi)$ for any $g_1, g_2 \in G$.

Proposition

Let (Φ, G) be a perception space. The followings hold.

a) *G is a topological group.*

b) *The action of G on Φ is continuous.*

GEOs and GENEOs

Definition

- Let (Φ, G), (Ψ, K) be two perception spaces. If a map $F : \Phi \to \Psi$ and a group homomorphism $T : G \to K$ are given, such that $F(g * \varphi) = T(g) * F(\varphi)$ for every $\varphi \in \Phi$, $g \in G$, we say that (F, T) is an (extended) *group equivariant operator* (GEO).
- If F is non-expansive (i.e., $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$), we say that (F, T) is an (extended) *group equivariant non-expansive operator* (GENEO).

Remark

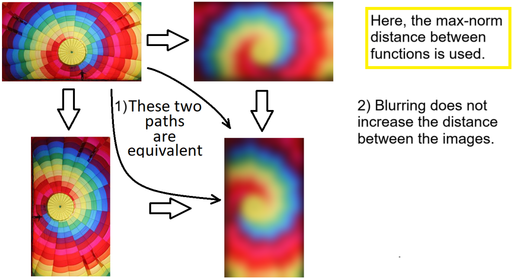
Let $(F, T) : (\Phi, G) \to (\Psi, K)$ be a GEO and assume that F is **surjective**. If $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$, then $D_K(T(g_1), T(g_2)) \leq D_G(g_1, g_2)$ for every $g_1, g_2 \in G$.

GENEO Example: Image Blurring

$\mathscr{I}$ = the collection of all color images, viewed as $C_c\left(\mathbb{R}^2, [0,1]^3\right)$, endowed with the max-norm distance;

- $F : \mathscr{I} \to \mathscr{I}$, $F(\varphi) = \varphi * k$, where k is rotationally symmetric and $\|k\|_1 = 1$;
- $T = \mathrm{id} : \mathrm{Isom}(\mathbb{R}^2) \to \mathrm{Isom}(\mathbb{R}^2)$.

Then (F, T) is a GENEO.

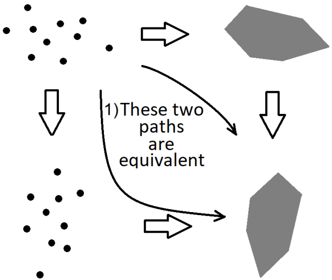


GENEO Example: Computing Convex Hulls

$\mathscr{S}$ = the collection of all finite nonempty subsets of the real plane;
$\mathscr{P}$ = the collection of all polyhedra in the real plane.

- $F : \mathscr{S} \to \mathscr{P}$, $F(S) =$ *convex hull of* S;
- $T = \mathrm{id} : \mathrm{Isom}(\mathbb{R}^2) \to \mathrm{Isom}(\mathbb{R}^2)$.

Then (F, T) is a GENEO.



Here, the Hausdorff distance between compact sets is used.

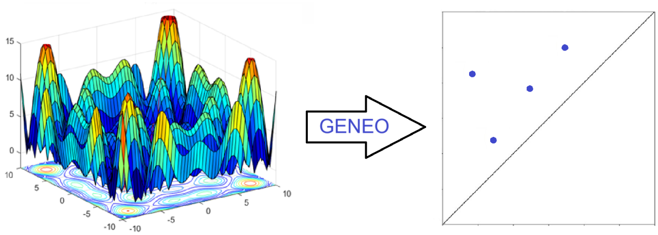
2) The operation of taking the convex hull does not increase the Hausdorff distance between sets.

GENEO Example: Computing Persistence Diagrams

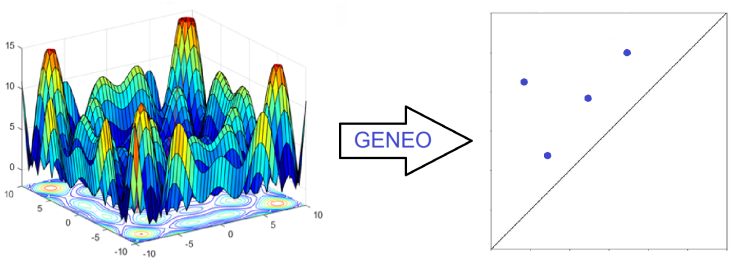
DGM = the collection of all persistence diagrams of real-valued continuous functions defined on a topological space X, where DGM is endowed with the bottleneck distance;

- $F : C_c(X, \mathbb{R}) \to \mathrm{DGM}$, $F(\varphi) = \mathrm{Dgm}_k(\varphi)$;
- $T : \mathrm{Homeo}(X) \to \{\mathrm{id}_{\mathrm{DGM}}\}$ is the trivial homomorphism.

Then (F, T) is a GENEO.



GENEO Example: Computing Persistence Diagrams



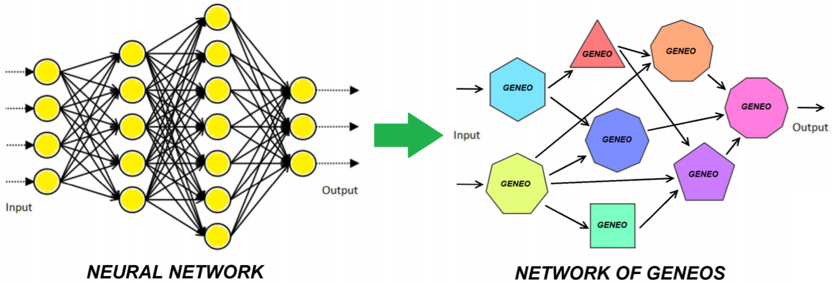
- Equivariance of (F, T) = invariance of persistence diagrams under reparameterization of the domain.
- Non-expansiveness of (F, T) = stability of persistence diagrams.

Why are GENEOs interesting?

- GENEOs rest on a **rigorous topological/geometric framework** (in what follows we outline several results).
- GENEOs **encode prior knowledge** about the chosen observer.
- The non-expansiveness property of GENEOs imposes a strong constraint that enables **meaningful data simplification**.
- GENEOs enable a **compositional approach to deep learning**.
- Analyzing the **geometry of the observer space** (as represented by GENEOs) is often more informative than analyzing the geometry of the data space.

The main point in the approach based on GENEOs

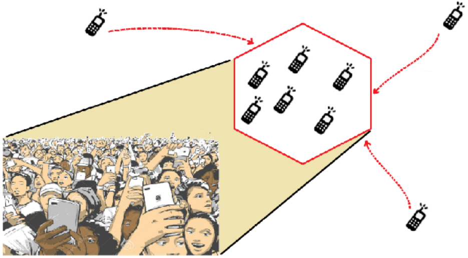
In perspective, we are looking for a good compositional theory for building efficient and transparent networks of GENEOs.
Some preliminary experiments suggest that replacing neurons with GENEOs could make some applications in deep learning more transparent and interpretable and speed up the learning process.



Some research projects concerning GENEOs (I)

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

Project on GENEOs for 6G



WILAB HUAWEI

Some research projects concerning GENEOs (II)

PANDORA

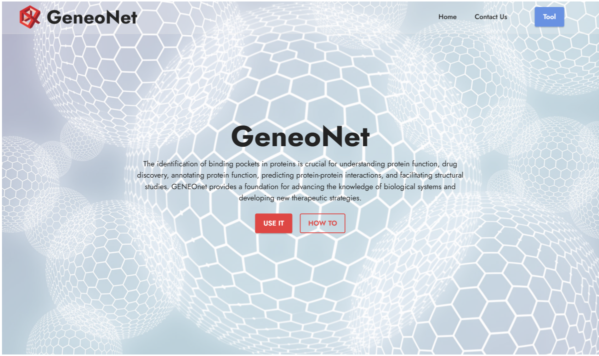
Horizon Europe (HORIZON)
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Project: 101135775-PANDORA
Funding: approximately 9 million euros.

Task 3.3 - Leveraging domain knowledge for explainable learning:
This task aims to investigate the use of domain knowledge in the development of explainable AI models. Tools like GENEOs for applications in TDA and ML and new theoretical methods of GENEOs for explainable AI will be used.

The project has received funding form the European Union's Horizon Europe Framework Programme (Horizon) under grant agreement No 101135775

https://pandora-heu.eu/consortium/

Some research projects concerning GENEOs (III)



The GENEOnet webservice represents the outcome of a partnership with Italian Pharmaceutical Company Dompé Farmaceutici S.p.A.: https://geneonet.exscalate.eu/

Compactness of the space of GENEOs

From now on, let $\mathscr{F}_T^{\mathrm{all}}$ denote the set of all GENEOs (F, T) from a perception space (Φ, G) to a perception space (Ψ, K), **with fixed homomorphism** T. We equip $\mathscr{F}_T^{\mathrm{all}}$ with the following distance:

$$D_{\mathrm{GENEO}}((F, T), (F', T)) := \sup_{\varphi \in \Phi} D_\Psi\big(F(\varphi), F'(\varphi)\big),$$

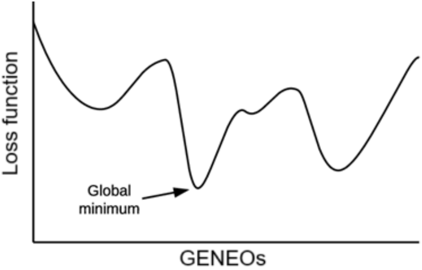
where D_Ψ is the metric chosen on Ψ.

Theorem

If (Φ, D_Φ), (Ψ, D_Ψ) *are* **compact**, *then the space* $(\mathscr{F}_T^{\mathrm{all}}, D_{\mathrm{GENEO}})$ *is* **compact**.

Compactness of the space of GENEOs

The compactness of the space of GENEOs is also an important property from an applied standpoint. It implies that the space of GENEOs can be approximated arbitrarily well by a finite set of GENEOs, and that every continuous loss function defined on this space attains an absolute minimum (i.e., there exists an optimal GENEO with respect to the loss function).



Convexity of the space of GENEOs

Now assume that

- Φ, Ψ **are normed real vector spaces**, with distances D_Φ, D_Ψ induced by the norms $\|\cdot\|_\Phi$, $\|\cdot\|_\Psi$ (i.e., $D_\Phi(\varphi,\varphi') = \|\varphi - \varphi'\|_\Phi$ and $D_\Psi(\psi,\psi') = \|\psi - \psi'\|_\Phi$);
- **the group actions are linear**, i.e.,
  $g * (a\varphi + b\varphi') = a(g * \varphi) + b(g * \varphi')$,
  $k * (a\psi + b\psi') = a(k * \psi) + b(k * \psi')$,
  for all $a, b \in \mathbb{R}$, $g \in G$, $k \in K$, $\varphi, \varphi' \in \Phi$, $\psi, \psi' \in \Psi$.

Let $(F_1, T), (F_2, T), \ldots, (F_n, T)$ be GENEOs from (Φ, G) to (Ψ, K).
If $(a_1, \ldots, a_n) \in \mathbb{R}^n$ and $\sum_{i=1}^n |a_i| \leq 1$, define

$$F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi).$$

Convexity of the space of GENEOs

Proposition

(F_Σ, T) is a GENEO from (Φ, G) to (Ψ, K).

Corollary

The space $\mathscr{F}_T^{\mathrm{all}}$ 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 Ψ 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, "to approximate an observer" means to look for a GENEO F that minimizes a suitable "cost function" $c(F)$. The cost function quantifies the error that is committed by taking the GENEO F instead of the ideal observer. Since the space of GENEOs 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 GENEO that best approximates the ideal observer.

The functional case of GENEOs

A particularly important case of GENEO is the one in which the perception spaces involved consist of sets of data expressed as functions with values in $\mathbb{R}$ or $\mathbb{R}^n$.

In the remainder of this talk we shall restrict our attention to the case of data expressed as real-valued functions.

In every perception space (Φ, G) under consideration, Φ will be a set of $\mathbb{R}$-valued functions defined on a domain X (denoted by $\mathrm{Dom}(\Phi)$) and endowed with the metric $D_\Phi(\varphi, \varphi') = \|\varphi - \varphi'\|_\infty$).

The group G will be a group of permutations of X such that, if $\varphi \in \Phi$ and $g \in G$, then $\varphi \circ g^{-1} \in \Phi$ as well, and the left action of G on Φ will be given by $g * \varphi = \varphi \circ g^{-1}$. We observe that this action is **isometric**.

Elementary methods to build GENEOs

In order to use our model profitably, we need constructive methods to produce GENEOs in the presence of pre-established data and equivariance groups.

Without going into technical details, we simply observe that, under reasonable assumptions,

- the composition of GENEOs is still a GENEO;
- the maximum and the minimum of GENEOs are still GENEOs;
- the translation of a GENEO is still a GENEO;
- the convex combination of GENEOs is still a GENEO.

(But there is much more than that...)

Generalized permutant measures

Let us consider the set $\Phi = \mathbb{R}^X \cong \mathbb{R}^n$ of all functions from a finite set $X = \{x_1, \ldots, x_n\}$ to $\mathbb{R}$, and a subgroup G of the group $\mathrm{Aut}(X)$ of all bijections from X to X. Similarly, let us consider the set $\Psi = \mathbb{R}^Y \cong \mathbb{R}^m$ of all functions from a finite set $Y = \{y_1, \ldots, y_m\}$ to $\mathbb{R}$, and a subgroup K of the group $\mathrm{Aut}(Y)$ of all bijections from Y to Y. Fix a homomorphism $T \colon G \to K$.

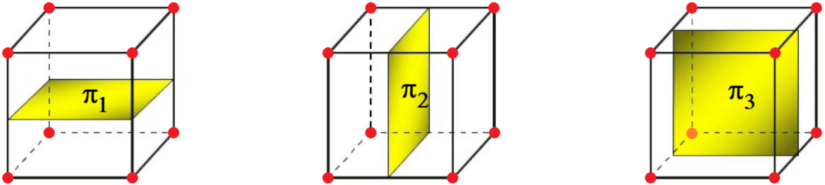
Definition

A finite signed measure μ on the power set of X^Y is called a **(generalized)** T**-permutant measure** if each subset H of X^Y is measurable and $\mu(\{h\}) = \mu(\{ghT(g^{-1})\})$ for every $g \in G$ and every $h \in X^Y$.

[F. Ahmad, M. Ferri, P. Frosini, *Generalized permutants and graph GENEOs*, **Machine Learning and Knowledge Extraction**, 5(4) (2023), 1905-1920.]

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. Set $Y = X$, $K = G$ and $T = \mathrm{id}_G$. Let π_1, π_2, π_3 be the three planes that contain the center of mass of X and are parallel to a face of the cube. Let $h_i : X \to X$ be the orthogonal symmetry with respect to π_i, for $i \in \{1,2,3\}$. We can now define a permutant measure μ on the power set of $X^Y = X^X$ 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 $h \in X^Y$ with $h \notin \{h_1, h_2, h_3\}$.



Representation Theorem for linear GENEOs

The following theorem explains the importance of the concept of a permutant measure.

Theorem (Representation Theorem for linear GENEOs)

If $T(G) \subseteq K$ acts transitively on the finite set Y and F is a map from $\mathbb{R}^X$ to $\mathbb{R}^Y$, then the pair (F,T) is a linear GENEO from $(\mathbb{R}^X, G)$ to $(\mathbb{R}^Y, K)$ if and only if there exists a generalized T-permutant measure μ such that

$$F(\varphi) = \sum_{h \in X^Y} \varphi \circ h \, \mu(h)$$

for every $\varphi \in \mathbb{R}^X$, and

$$\sum_{h \in X^Y} |\mu(h)| \leq 1.$$

Representation Theorem for linear GENEOs

For more details, please see:

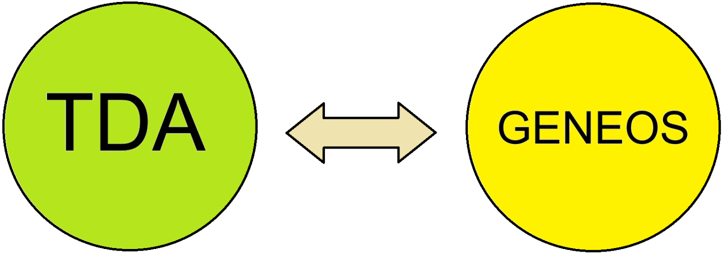
- G. Bocchi, S. Botteghi, M. Brasini, P. Frosini, N. Quercioli, *On the finite representation of linear group equivariant operators via permutant measures*, **Annals of Mathematics and Artificial Intelligence**, vol. 91 (2023), n. 4, 465 487.
- F. Conti, P. Frosini, N. Quercioli, *An algebraic representation theorem for linear GENEOs in Geometric Machine Learning*, arxiv.org/abs/2601.03910 (2026).

GENEOs and TDA

We have already highlighted that the computation of persistence diagrams can be formalized as a GENEO.

However, there are also other interesting connections between the concept of GENEO and Topological Data Analysis.

In what follows, we will illustrate some of them.



GENEOs restrict the invariance of TDA

The use of GENEOs allows us **to restrict** the invariance of TDA to subgroups of the group $\mathrm{Homeo}_\Phi(X)$ of all Φ-preserving homeomorphisms of X. To show this, let us assume that a set $\mathscr{F}$ of GENEOs from a perception space (Φ, G) to a perception space (Ψ, K) is given, with respect to a fixed homomorphism $T\colon G \to K$. Let X denote the domain of the functions in Φ. Then, for every degree k, we can define a new pseudometric $\mathscr{D}_{\mathrm{match}}^{\mathscr{F},\Phi}$ on Φ by

$$\mathscr{D}_{\mathrm{match}}^{\mathscr{F},\Phi}(\varphi_1,\varphi_2) = \sup_{F\in\mathscr{F}} d_{\mathrm{match}}\big(\mathrm{Dgm}_k(F(\varphi_1)), \mathrm{Dgm}_k(F(\varphi_2))\big).$$

If $\varphi \in \Phi$, then $\mathscr{D}_{\mathrm{match}}^{\mathscr{F},\Phi}(\varphi \circ g, \varphi) = 0$ for every $g \in G$, but, in general, not for every $g \in \mathrm{Homeo}_\Phi(X)$.
Since GENEOs are non-expansive, the pseudometric $\mathscr{D}_{\mathrm{match}}^{\mathscr{F},\Phi}$ is stable:

$$\mathscr{D}_{\mathrm{match}}^{\mathscr{F},\Phi}(\varphi_1,\varphi_2) \le \|\varphi_1 - \varphi_2\|_\infty.$$

GENEOs restrict the invariance of TDA

If Φ is compact, the pseudometric $\mathscr{D}_{\text{match}}^{\mathscr{F},\Phi}$ can be approximated arbitrarily well by a finite subset of operators.

Proposition

Assume that Φ is compact. Let $\mathscr{F}$ be a nonempty set of GENEOs $(F,T)\colon(\Phi,G)\to(\Psi,K)$, where the homomorphism $T\colon G\to K$ is fixed. For every $\varepsilon>0$, there exists a finite subset $\mathscr{F}^*\subset\mathscr{F}$ such that

$$\left|\mathscr{D}_{\text{match}}^{\mathscr{F}^*,\Phi}(\varphi_1,\varphi_2)-\mathscr{D}_{\text{match}}^{\mathscr{F},\Phi}(\varphi_1,\varphi_2)\right|\le\varepsilon$$

for every $\varphi_1,\varphi_2\in\Phi$.

The previous statement follows from the compactness theorem for the space of GENEOs.

GENEOs and biparameter persistent homology

Assume that $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ and $\boldsymbol{\psi} = (\psi_1, \psi_2)$ belong to $C^0(X, \mathbb{R}^2)$. For each $(a, b) \in]0, 1[\times \mathbb{R}$, let us define the map $F_{a,b}\colon C^0(X, \mathbb{R}^2) \to C^0(X, \mathbb{R})$ by setting

$$F_{a,b}(\boldsymbol{\varphi}) = \max \left\{ \frac{\min\{a, 1-a\}}{a} (\varphi_1 - b), \frac{\min\{a, 1-a\}}{1-a} (\varphi_2 + b) \right\}.$$

Interestingly, $F_{a,b}$ is a GENEO.

For each degree k, we can consider the distance

$$\mathbb{D}_{\mathrm{match},k}(\boldsymbol{\varphi}, \boldsymbol{\psi}) := \sup_{(a,b) \in]0,1[\times \mathbb{R}} d_{\mathrm{match}}(\mathrm{Dgm}_k(F_{a,b}(\boldsymbol{\varphi})), \mathrm{Dgm}_k(F_{a,b}(\boldsymbol{\psi}))).$$

It is called the two-dimensional **matching distance**.

GENEOs and biparameter persistent homology

It is worth noting that, through the use of GENEOs, one can obtain many different distances for biparametric persistence. For example, if we are interested in distances based on one-parameter families of GENEOs, we can define pseudometrics in biparametric TDA using

$$\bar{F}_t(\boldsymbol{\varphi}) := (1-t)\varphi_1 + t\varphi_2, \quad t \in [0,1],$$

or

$$\hat{F}_t(\boldsymbol{\varphi}) := (1-t)\max(\varphi_1,\varphi_2) + t\min(\varphi_1,\varphi_2), \quad t \in [0,1],$$

or

$$\tilde{F}_t(\boldsymbol{\varphi}) := \left(\tfrac{1}{2}|\varphi_1|^t + \tfrac{1}{2}|\varphi_2|^t\right)^{\frac{1}{t}}, \quad t \geq 1.$$

[M. G. Bergomi, P. Frosini, D. Giorgi, N. Quercioli, *Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning*, **Nature Machine Intelligence**, vol. 1, n. 9, 423–433 (2019).]

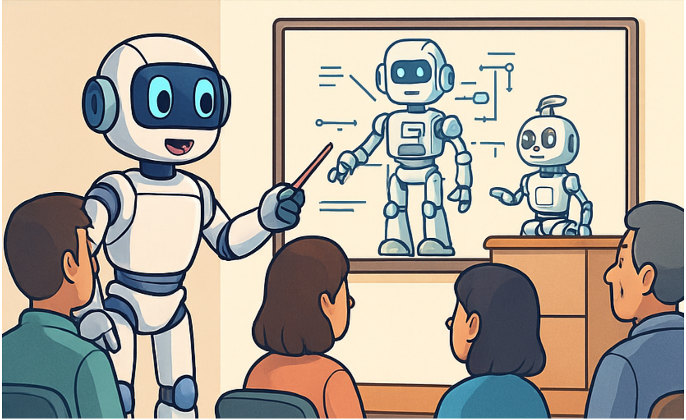
[N. Quercioli, *Some new methods to build group equivariant non-expansive operators in TDA*. In: Devaney, R.L., Chan, K.C., Vinod Kumar, P. (eds) Topological Dynamics and Topological Data Analysis. IWCTA 2018. **Springer Proceedings in Mathematics & Statistics**, vol 350. Springer, Singapore (2021).]

[P. Frosini, U. Fugacci, E. Mosig García, N. Quercioli, S. Scaramuccia, F. Tombari, *The convex matching distance in multiparameter persistence*, arxiv.org/abs/2512.02944 (2025).]

[N. Berkouk, F. Petit, *Projected distances for multi-parameter persistence modules*, **Annales de l'Institut Fourier**, Online first (2026), 62 p.] (a different but related approach to our research)

GENEOs and XAI

We can apply GENEOs to explainable AI.



Collaborators in this research

- Filippo Bonchi (University of Pisa)
- Jacopo Joy Colombini (Scuola Normale Superiore, Pisa)
- Francesco Giannini (University of Pisa)
- Fosca Giannotti (Scuola Normale Superiore, Pisa)
- Roberto Pellungrini (Scuola Normale Superiore, Pisa)

[J. J. Colombini, F. Bonchi, F. Giannini, F. Giannotti, R. Pellungrini and P. Frosini, *Mathematical foundation of interpretable equivariant surrogate models*, 3rd World Conference on Explainable Artificial Intelligence (XAI-2025), Novel Post-hoc & Ante-hoc XAI Approaches, 09-11 July, 2025 - Istanbul, Turkey, **Communications in Computer and Information Science**, vol 2577 (2026), 294-318. Springer, Cham.]

Basic idea

How can we mathematically and generally formalize the concept of an explanation provided by an agent, viewed as an operator?

Informal idea: We say that the action of an agent A is **explained** by another agent B from the perspective of an agent C if:

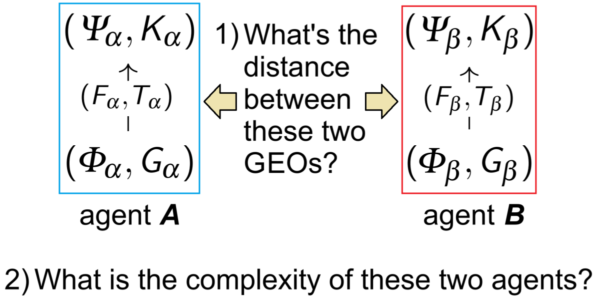
1. C perceives A and B as similar to each other;
2. C perceives B as less complex than A.

For example, we can consider two neural networks represented by two GEOs, A and B. Note that a GEO C can take GEOs A and B as inputs.

An extended pseudometric for *ALL* GEOs

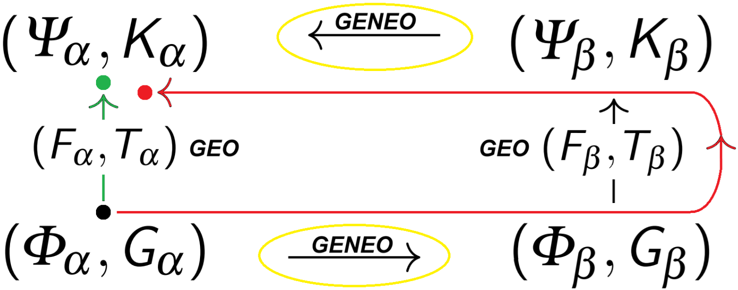
To proceed in this way, we need

- a pseudometric between GEOs that remains well-defined even when the GEOs operate on **different** domains and produce outputs in **distinct** codomains;
- a measure of complexity for GEOs.



An extended pseudometric for *ALL* GEOs

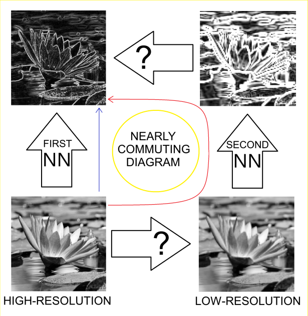
Informally speaking, two GEOs are considered similar if there exist two horizontal GENEOs that make this diagram "nearly commutative", with the same holding true in the opposite directions ($\rightleftarrows$):

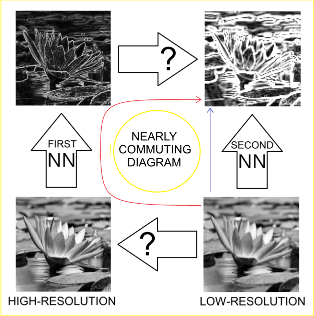


A **cost function** can quantify the non-commutativity of each diagram.

An example

Suppose we have two neural networks for edge detection in images, represented as GEOs.



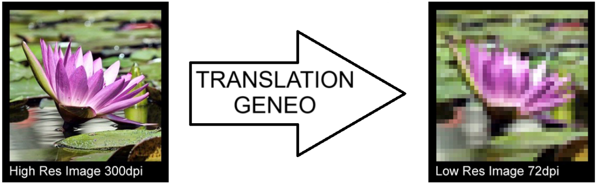


The two neural networks are considered close if there exist two pairs of horizontal GENEOs that make these diagrams "nearly commutative".

An extended pseudometric for *ALL* GEOs

To formalize our new pseudometric d_E between GEOs, we consider the category $\mathbf{S}_{all}$ whose objects are all perception spaces, and whose morphisms $(F, T) : (\Phi, G) \to (\Phi', G')$ are GENEOs.
The morphisms in $\mathbf{S}_{all}$ are called *translation GENEOs*. These morphisms describe the possible "logical correspondences" between data represented by different perception spaces.

For example, a translation GENEO might transform high-resolution images into low-resolution images.



Definition of the explainability distance d_E

We omit the technical details of the definition of the distance d_E between GEOs, and simply observe that **the non-expansiveness of translation GENEOs is a key ingredient** in its definition.

When the distance between two GEOs is small, it indicates that they act approximately in the same way on the data they process.

Complexity of GEOs

Let us assume a set $\Gamma = \{(F_i, T_i) : (\Phi_i, G_i) \to (\Psi_i, K_i)\}$ of GEOs is given. We will say that Γ is our internal library. For each GEO $(F_i, T_i) \in \Gamma$ we arbitrarily choose a value c_i representing the complexity $\mathrm{comp}((F_i, T_i))$ of (F_i, T_i).

The set Γ represents the elementary GEOs that we can use to build other more complex GEOs.

Let us now consider the closure of Γ, i.e., the minimal set $\bar{\Gamma}$ such that

- $\bar{\Gamma} \supseteq \Gamma$;
- $\bar{\Gamma}$ is closed under composition (i.e., if $(F, T), (F', T') \in \bar{\Gamma}$ are composable, then $(F', T') \circ (F, T) \in \bar{\Gamma}$);
- $\bar{\Gamma}$ is closed under direct product (i.e., if the GEOs $(F, T), (F', T') \in \bar{\Gamma}$, then $(F, T) \otimes (F', T') \in \bar{\Gamma}$).

Complexity of GEOs

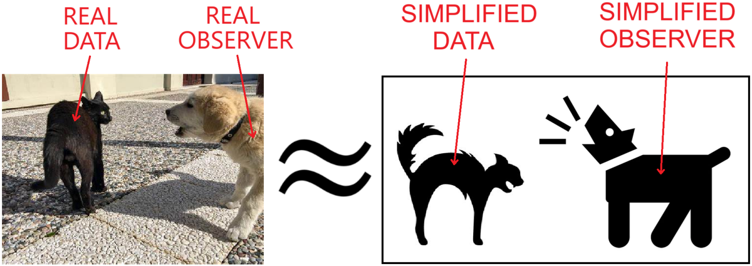
Each composition and direct product is associated with a complexity.

The complexity of each GEO $(F, T) \in \bar{\Gamma}$ is obtained by minimizing the sum of the complexities of the GEOs (F_i, T_i) that we use and the complexities of the compositions and direct products that we apply to build (F, T).

Other forms of composition of GEOs can be added to the model.

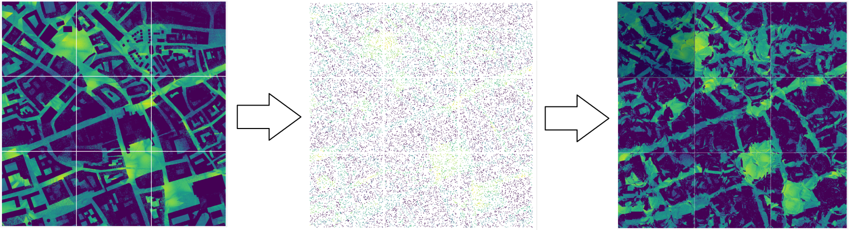
A mathematical concept of explanation

Now we can formalize our mathematical concept of **explanation**. Specifically, we can define it as follows, after choosing the values ε and k: *"The action of an agent represented by a GEO* (F_α, T_α) *is* ***explained at a level*** ε *by the action of another agent of* **complexity less than** k *represented by a GEO* $(F_\beta, T_\beta) \in \bar{\Gamma}$ *when* $d_E((F_\alpha, T_\alpha), (F_\beta, T_\beta)) \leq \varepsilon$*."*



Explainable signal reconstruction

Reconstruction of sparse urban wireless signals via GENEOs



Collaborators in this research

- Lorenzo Mario Amorosa (University of Bologna & WiLab-CNIT)
- Francesco Conti (Université Côte d'Azur)
- Yiqun Ge (Huawei Technologies Canada)
- Tayebeh Lotfi Mahyari (Huawei Technologies Canada)
- Nicola Quercioli (CINECA)
- Flavio Zabini (University of Bologna & WiLab-CNIT)

[L. M. Amorosa, F. Conti, N. Quercioli, F. Zabini, T. L. Mahyari, Y. Ge, P. Frosini, *Reconstruction of sparse urban wireless signals via group equivariant non-expansive operators*, arxiv.org/html/2507.19349v1 (2025).]

Explainable signal reconstruction

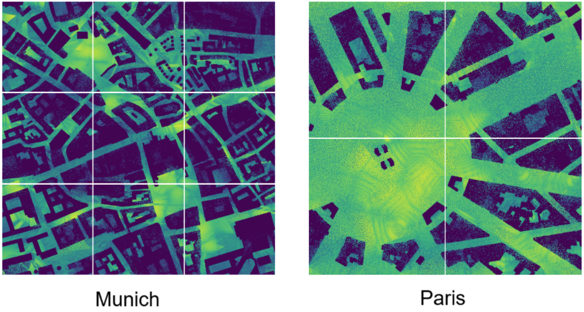
We propose a **GENEO-based approach** for reconstructing radio signals in urban wireless networks from **extremely sparse samples**.

We describe the signal to be reconstructed as a function $\varphi : \mathbb{R}^2 \to [0,1]$. The function φ represents the intensity of the signal over a 2D grid, and is often referred to as ground truth or GT. We also consider the function $\psi : \mathbb{R}^2 \to [0,1]$ taking each point p to the reliability $\psi(p)$ of the value $\varphi(p)$ measured at the point $p = (x,y) \in \mathbb{R}^2$.

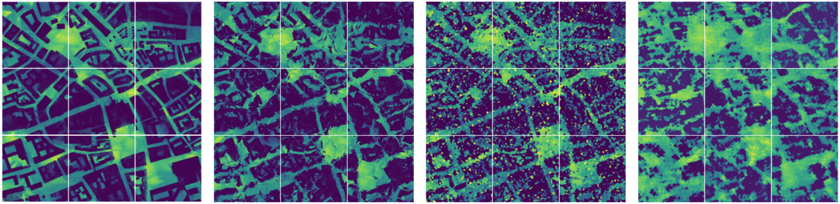
We aim at reconstructing φ by starting from a very poor sampling of φ (tipically 2% or 3% of the GT), by using suitable families of GENEOs.

Explainable signal reconstruction

In our numerical experiments, we focus on reconstructing the signal over a two-dimensional area. We generated signal measurements with the Sionna RT ray-tracing simulator, using its built-in outdoor urban scenarios for Munich and Paris. Each scenario is discretized into a $L \times L$ grid of $1\,\mathrm{m}^2$ pixels $\{p_j\}_{j=1}^{L^2}$.



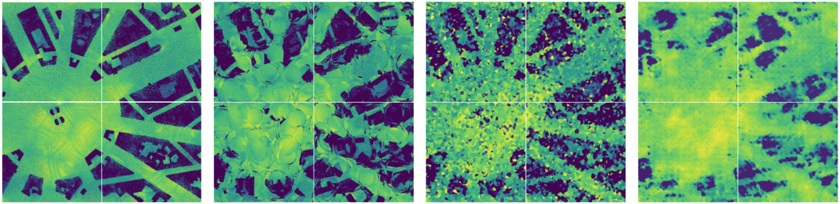
Explainable signal reconstruction



(a) (b) (c) (d)

Comparison of reconstruction methods on the Munich scenario, with only $M = 3\%$ of signals known, whose $Q = 15\%$ of them is featuring errors. (a) ground truth, (b) GENEO, (c) 1-KNN, and (d) U-Net.

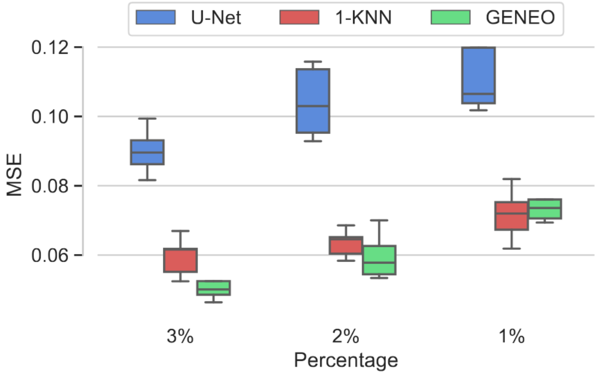
Explainable signal reconstruction



(a) (b) (c) (d)

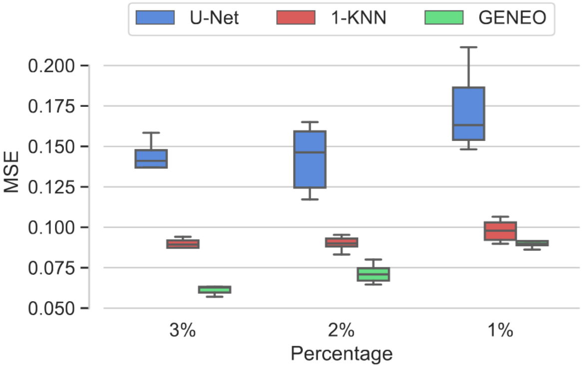
Comparison of reconstruction methods on the Paris scenario, with only $M = 3\%$ of signals known, whose $Q = 15\%$ of them is featuring errors. (a) ground truth, (b) GENEO, (c) 1-KNN, and (d) U-Net.

Explainable signal reconstruction



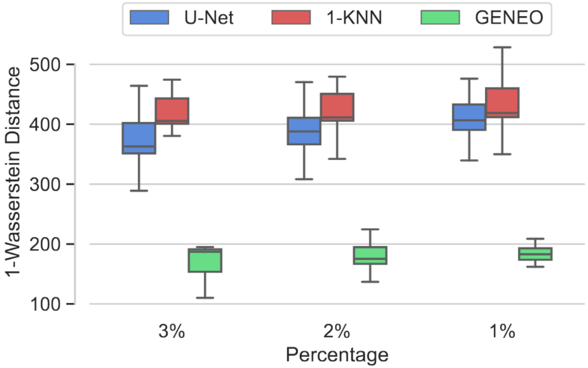
MSE achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Munich scenario, where $M \in \{1,2,3\}$ and $Q = 15$.

Explainable signal reconstruction



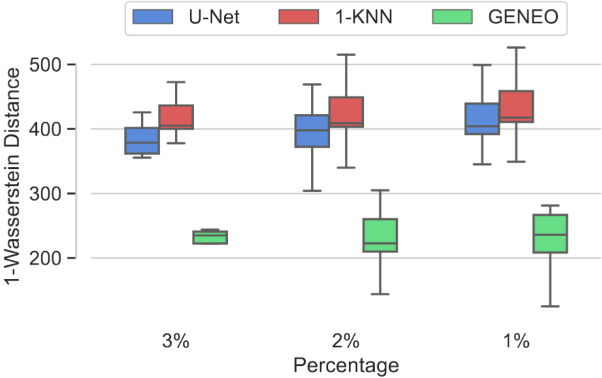
MSE achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Munich scenario, where $M \in \{1, 2, 3\}$ and $Q = 30$.

Explainable signal reconstruction



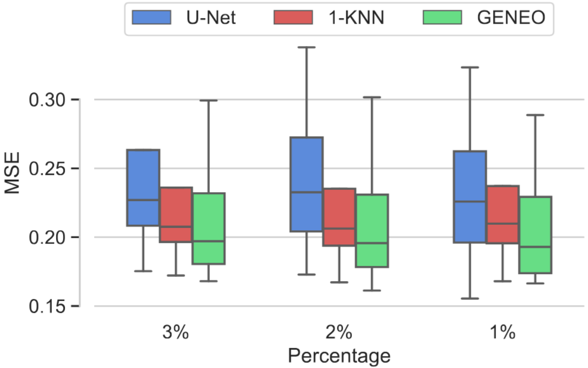
1-Wasserstein achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Munich scenario, where $M \in \{1,2,3\}$ and $Q = 15$.

Explainable signal reconstruction



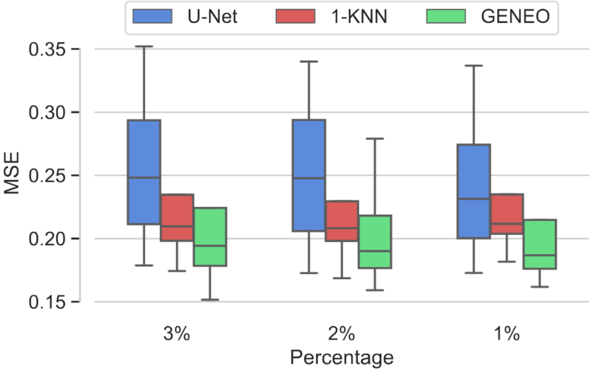
1-Wasserstein achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Munich scenario, where $M \in \{1, 2, 3\}$ and $Q = 30$.

Explainable signal reconstruction



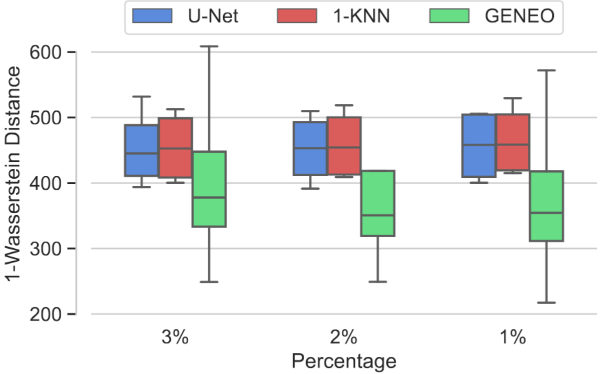
MSE achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Paris scenario, where $M \in \{1, 2, 3\}$ and $Q = 15$.

Explainable signal reconstruction



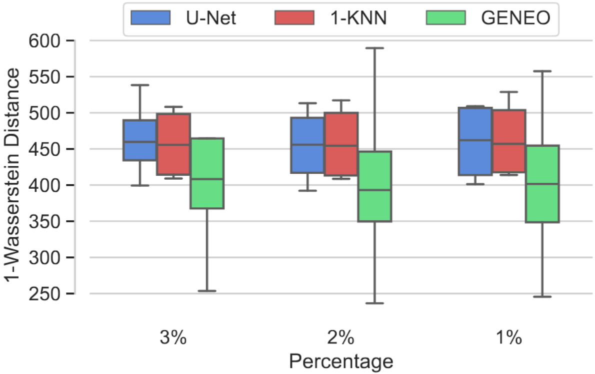
MSE achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Paris scenario, where $M \in \{1, 2, 3\}$ and $Q = 30$.

Explainable signal reconstruction



1-Wasserstein achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Paris scenario, where $M \in \{1, 2, 3\}$ and $Q = 15$.

Explainable signal reconstruction



1-Wasserstein achieved by GENEO, U-Net, and 1-KNN for normalized signal reconstruction in Paris scenario, where $M \in \{1, 2, 3\}$ and $Q = 30$.

Explainable signal reconstruction

MSE Performance: For $Q = 15$, GENEO (green) achieves the lowest reconstruction error in terms of MSE across all sampling ratios M. Under heavier corruption (Q=30), GENEO's advantage persists: it consistently outperforms 1-KNN and U-Net, with the largest margin at M=1.

1-Wasserstein Performance: For $Q = 15$, GENEO again yields substantially lower topological error, improving over 1-KNN and U-Net across all M. For Q=30, GENEO maintains its lead in topological fidelity, while both baselines exhibit worse performance in terms of signal reconstruction.

TAKE-AWAY MESSAGE

To sum up, GENEOs are novel mathematical tools initially developed in TDA and useful for approximating equivariant neural networks using a compositional approach. GENEOs are generally interpretable, making them potentially beneficial for explainable artificial intelligence (XAI).

THANKS FOR YOUR ATTENTION