



entropy



Article

A Probabilistic Result on Impulsive Noise Reduction in Topological Data Analysis through Group Equivariant Non-Expansive Operators

Patrizio Frosini, Ivan Gridelli and Andrea Pascucci

Special Issue

Topological Data Analysis Meets Information Theory. New Perspectives for the Analysis of Higher-Order Interactions in Complex Systems

Edited by



Dr. Tiziano Squartini, Dr. Fernando Antônio Nóbrega Santos, Dr. Rossana Mastrandrea and Dr. Marco Piangerelli



<https://doi.org/10.3390/e25081150>

Article

A Probabilistic Result on Impulsive Noise Reduction in Topological Data Analysis through Group Equivariant Non-Expansive Operators

Patrizio Frosini ^{*,†} , Ivan Gridelli [†] and Andrea Pascucci [†] 

Department of Mathematics, University of Bologna, 40126 Bologna, Italy; ivangridelli@gmail.com (I.G.); andrea.pascucci@unibo.it (A.P.)

* Correspondence: patrizio.frosini@unibo.it

[†] These authors contributed equally to this work.

Abstract: In recent years, group equivariant non-expansive operators (GENEOs) have started to find applications in the fields of Topological Data Analysis and Machine Learning. In this paper we show how these operators can be of use also for the removal of impulsive noise and to increase the stability of TDA in the presence of noisy data. In particular, we prove that GENEOs can control the expected value of the perturbation of persistence diagrams caused by uniformly distributed impulsive noise, when data are represented by L -Lipschitz functions from \mathbb{R} to \mathbb{R} .

Keywords: GENEO; impulsive noise; persistence diagram; persistent homology; machine learning

MSC: 55N31; 62R40; 60-08; 65D18; 68T09; 68U05



Citation: Frosini, P.; Gridelli, I.; Pascucci, A. A Probabilistic Result on Impulsive Noise Reduction in Topological Data Analysis through Group Equivariant Non-Expansive Operators. *Entropy* **2023**, *25*, 1150. <https://doi.org/10.3390/e25081150>

Academic Editors: Tiziano Squartini, Fernando Antônio Nóbrega Santos, Rossana Mastrandrea and Marco Piangerelli

Received: 5 May 2023

Revised: 22 July 2023

Accepted: 25 July 2023

Published: 31 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the last thirty years, Topological Data Analysis (TDA) developed as a useful mathematical theory for analyzing data [1], benefiting from the dimensionality reduction guaranteed by topology [2–4]. Topology can construct low-dimensional representations of high-dimensional manifolds of data, thereby reducing the dimensionality of the parameter space and ultimately reducing the learning cost of machine learning models. One of the main tools in TDA is the concept of the persistence diagram, which is a collection of points in the real plane that describes the homological changes of the sublevel sets of suitable continuous functions. These changes give important information about the data of interest, focusing on some of their most relevant properties. Persistence diagrams can be used in the presence of noise, since a well-known stability theorem states that these topological descriptors change in a controlled way when we know that the functions expressing the filtrations we are interested in change in a controlled way with respect to the sup-norm. More precisely, the sup-norm distance between two functions is an upper bound for the bottleneck distance between the corresponding persistence diagrams [5]. Furthermore, L^p -stability of persistence diagrams with respect to the sup-norm have been proved in [6]. Unfortunately, in many applications, the sup-norm of the noise is not guaranteed to be small, and hence these results cannot be directly applied. In particular, these results cannot be directly used when data are represented by functions from \mathbb{R} to \mathbb{R} and affected by impulsive noise, i.e., noise that carries a sudden sharp change of short duration, when the variable is seen as a time. It is indeed well known that persistence diagrams can completely change when the data are subject to impulsive noise.

Analogously, in the discrete setting of TDA, the presence of outliers in cloud points can drastically affect the corresponding persistence diagrams. The problem of managing outliers has been studied by several authors with different techniques. In [7], an approach based on confidence sets has been introduced. A method inspired by the k -nearest neighbors

regression and the local median filtering has been used in [8], while the concept of the bagplot has been applied in [9]. In [10], an approach based on reproducing kernels has been proposed.

In our paper, we start exploring a different probabilistic approach, in the topological setting. The main idea is that studying the properties of data should primarily be based on the analysis of the observers that are asked to examine the data, since we cannot ignore that different observers can differently judge the same data. This approach has been initially proposed in [11] and requires both the definition of the space of group equivariant non-expansive operators (GENEOs) and the development of geometrical techniques to move around in this space [12,13]. In other words, in this model we should not wonder how we have to manage the data but rather how we have to manage the observers (i.e., GENEOs) analyzing the data. As a first step in this direction, in this paper we show how GENEOs could be used to obtain stability of persistence diagrams of 1D signals in the presence of impulsive noise. For the use of such operators in the comparison of 1D signals under the action of several transformation groups, we refer the interested reader to the paper [14]. This type of signal is important in many applications, such as those concerning EEG data and time series.

GENEOs have been studied in [14] as a new tool in TDA, since they allow for an extension of the theory that is not invariant under the action of every homeomorphism of the considered domain. This is important in applications where the invariance group is not the group of all homeomorphisms, such as the ones concerning shape comparison. Interestingly, GENEOs are also deeply related to the foliation method used to define the matching distance in two-dimensional persistent homology [15,16] and can be seen as a theoretical bridge between TDA and Machine Learning [11]. Furthermore, these operators make available lower bounds for the natural pseudo-distance $d_G(\varphi_1, \varphi_2) := \inf_{g \in G} \|\varphi_1 - \varphi_2 \circ g\|_\infty$, associated with a group G of self-homeomorphisms of the domain of the signals φ_1, φ_2 [14]. For general information about the interest in the theory of GENEOs and its applications, we refer the reader to [17].

In our paper, we prove that GENEOs can control the expected value of the perturbation of persistence diagrams caused by uniformly distributed impulsive noise when data are represented by L -Lipschitz functions from \mathbb{R} to \mathbb{R} . In order to do that, we choose a “mother” bump function ψ , i.e., a continuous function that is non-negative, upper bounded by 1, and compactly supported, and we assume that our noise is made up of finite bumps each obtained by translating, heightening, and/or widening ψ . The function $\hat{\varphi} = \varphi + R$ represents the corrupted data, where the noise is given by the function $R = \sum_{i=1}^k a_i \psi(b_i(x - c_i))$ for some $a_i, b_i, c_i \in \mathbb{R}$, with $b_i > 0$, and some positive integer k .

In this situation, trying to use a convolution to approximate our starting data is not effective, because even if it does contract the bumps, it does not cut them and hence does not improve the sup-norm distance from the original data. A classical approach for the removal of impulsive noise is the one of using a median filter [18]. Although this approach would be quite efficient in the discrete case, let us remark that in our setting, we are considering continuous functions. The median for the continuous case is defined as the interval of those values m such that

$$\int_{-\infty}^m f(x) dx = \int_m^{+\infty} f(x) dx,$$

where f is a density of probability. However, this operator is not stable: a small alteration in the starting function could lead to a significant change in the median.

The operators we consider in this paper are $F^\delta(\varphi)(x) = \max(\varphi(x - \delta), \varphi(x + \delta))$ and $F_\varepsilon(\varphi)(x) = \min(\varphi(x - \varepsilon), \varphi(x + \varepsilon))$. The main idea is that F^δ removes the noise “directed downwards” and F_ε the noise “directed upwards”, and hence, their composition should be able to eliminate all the bumps. These operators are GENEOs with respect to isometries of the real line. We prove that moving in the space of GENEOs by taking suitable values

for ε and δ , we can get quite close to restoring the original function φ , depending on how the bumps are positioned. The closer the bumps are to being Dirac delta functions and the further they are from each other, the better our approximation can be. On the ground of this result, we finally obtain an estimate of the expected value $E(\|F^\delta \circ F_\varepsilon(\hat{\varphi}) - \varphi\|_\infty)$.

We conclude this introduction by observing that, while the operators we use in this paper are elementary, the theory of GENEOS makes available procedures to learn more complex operators by combining simple GENEOS through techniques belonging to the rising field of Geometric Deep Learning [11–13]. For this reason, we hope that our results can not only illustrate a practical use of some particular GENEOS but also be a first step towards the application of Machine Learning for increasing the stability of TDA in the presence of noise.

The paper is structured as follows. In Section 2, the mathematical background is laid. The case we consider and our notations are explained in Section 3. In Section 4, we prove the results that are needed in order to demonstrate our main results. In Section 5, the main theorems giving us probabilistic upper bounds are formulated. In Section 6, some examples and experiments are presented in order to better illustrate the use of our results. A brief discussion concludes the paper.

2. Mathematical Setting

In this section, we will recall some basic concepts we will use in this paper.

To avoid misunderstandings about pairs and intervals, we will use the symbols (u, v) , $[u, v]$, and $]u, v[$ to denote a pair, a closed interval, and an open interval, respectively.

2.1. Representing Data as Real Functions

Let us consider a set Φ of bounded functions from a set X to \mathbb{R} , which will represent the data we wish to take into account (e.g., functions describing sounds). We shall call Φ the set of *admissible measurements* on X . We endow Φ with the topology induced by the sup-norm $\|\cdot\|_\infty$ and the corresponding distance D_Φ . A pseudo-metric D_X can be defined on X by setting $D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|$ for every $x_1, x_2 \in X$. We recall that a pseudo-metric on a set X is a distance d without the property $d(x, y) = 0 \implies x = y$ (in other words, we do not have the so-called “identity of indiscernibles”). We will consider the topological space (X, τ_{D_X}) where τ_{D_X} is the topology induced by D_X . A base for this topology is given by the set of open balls $\{B(x, r) : x \in X, r \in \mathbb{R}\}$, where $B(x, r) := \{x' \in X : D_X(x, x') < r\}$. The choice of this topology makes every function in Φ a continuous functions. As shown in [11], this fact enables us to use persistence diagrams in the study of Φ .

2.2. GENEOS as Operators Acting on Data

We are interested in considering transformations of data. Let $\text{Homeo}_\Phi(X)$ be the group of Φ -preserving homeomorphisms from X to X with respect to the topology τ_{D_X} , meaning that every g in $\text{Homeo}_\Phi(X)$ is a homeomorphism of X such as both $\varphi \circ g$ and $\varphi \circ g^{-1}$ belong to Φ for every φ in Φ . Let G be a subgroup of $\text{Homeo}_\Phi(X)$. G represents the set of transformations on data for which we will require equivariance to be respected.

Under the previously stated assumptions, we call the ordered pair (Φ, G) a *perception pair*. We can now introduce the concept of GENEOS.

Definition 1. Let (Φ, G) and (Ψ, H) be perception pairs and assume that a homomorphism $T : G \rightarrow H$ is given. A function $F : \Phi \rightarrow \Psi$ is called a *Group Equivariant Non-Expansive Operator (GENEO)* from (Φ, G) to (Ψ, H) with respect to T if the following properties hold:

1. (Group Equivariance) $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for every $\varphi \in \Phi, g \in G$;
2. (Non-Expansivity) $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$.

Let us now consider the set \mathcal{F}_{all} of all GENEOS from (Φ, G) to (Ψ, H) with respect to $T : G \rightarrow H$. The space \mathcal{F}_{all} is endowed with the extended pseudo-metric $D_{\mathcal{F}_{\text{all}}}$, defined by setting $D_{\mathcal{F}_{\text{all}}}(F_1, F_2) = \sup_{\varphi \in \Phi} D_{\Psi}(F_1(\varphi), F_2(\varphi))$ for every $F_1, F_2 \in \mathcal{F}_{\text{all}}$. The word *extended* refers to the possibility that $D_{\mathcal{F}_{\text{all}}}$ takes an infinite value.

The following result can be proven [11]:

Theorem 1. *If $(\Phi, D_{\Phi}), (\Psi, D_{\Psi})$ are compact and convex then the metric space $(\mathcal{F}_{\text{all}}, D_{\mathcal{F}_{\text{all}}})$ is compact and convex.*

If a non-empty set $\mathcal{F} \subseteq \mathcal{F}_{\text{all}}$ is fixed, we can define the following pseudo-distance $D_{\mathcal{F}, \Phi}$ on Φ :

Definition 2. *For any φ_1, φ_2 in Φ we set*

$$D_{\mathcal{F}, \Phi}(\varphi_1, \varphi_2) = \sup_{F \in \mathcal{F}} \|F(\varphi_1) - F(\varphi_2)\|_{\infty}.$$

This pseudo-distance allows us to compare data by taking into account how agents operate on data. Notice how, if G becomes larger, the natural pseudo-distance d_G becomes harder to compute but this new pseudo-distance $D_{\mathcal{F}, \Phi}$ becomes easier to evaluate. In order to find a lower bound for $D_{\mathcal{F}, \Phi}$ it is useful to introduce the notion of persistence diagram.

2.3. Persistence Diagrams

We will now recall some basic definitions and results in persistent homology. The interested reader can find more details in [4].

Let us consider an ordered pair (X, φ) where X is a topological space and $\varphi : X \rightarrow \mathbb{R}$ is a continuous function. For any $t \in \mathbb{R}$ we can set $X_t := \varphi^{-1}([-\infty, t])$. If $u < v$, the inclusion $i_{u,v} : X_u \rightarrow X_v$ induces a homomorphism $i_{u,v}^k : H_k(X_u) \rightarrow H_k(X_v)$ between the k th homology groups of X_u and X_v . We can define the *kth persistent homology group*, with respect to φ and computed at the point (u, v) , as $PH_k(u, v) := i_{u,v}^k(H_k(X_u))$. Moreover, we can define the *kth persistent Betti numbers function* $r_k(u, v)$ as the rank of $PH_k(u, v)$.

The k th persistent Betti numbers function can be represented by the k th persistence diagram. This diagram is defined as the multi-set of all the ordered pairs (u_j, v_j) , where u_j and v_j are the times of birth and death of the j th k -dimensional hole in X , respectively. We call *time of birth* of a hole the first time at which the homology class appears, and *time of death* the first time at which the homology class merges with an older one. When a hole never dies, we set its time of death equal to ∞ . We also add to this set all points of the form (w, w) for $w \in \mathbb{R}$.

In Figure 1 the filtration of the set $X := [0, \frac{3}{4}\pi]$ given by the function $\varphi(x) = 2 \sin x$ is illustrated. In this example, the topology on X is the one defined by the space Φ of admissible functions given by all functions $a \sin(x + b)$ with $a, b \in \mathbb{R}$. The reader can easily check that this topology is the same as the Euclidean topology. The persistence diagram in degree $k = 0$ of the function φ is displayed in Figure 2.

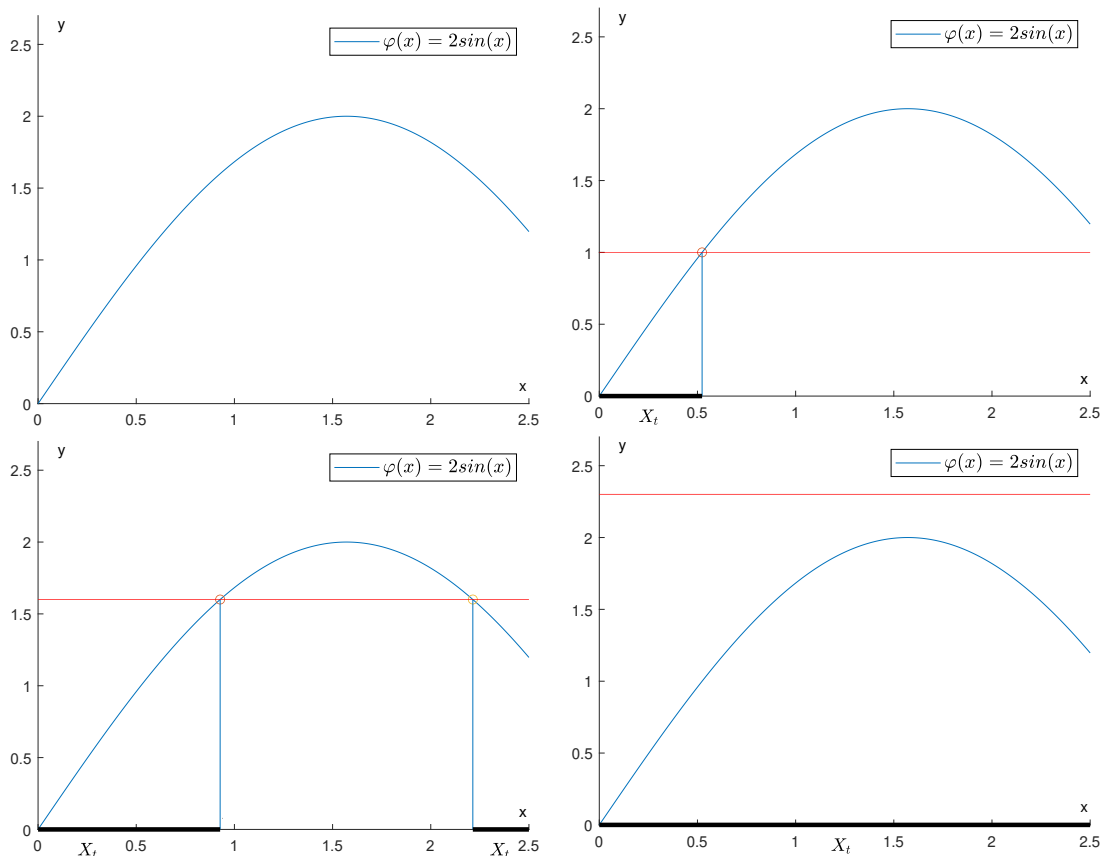


Figure 1. How the sublevel sets change with respect to the filtration induced by φ .

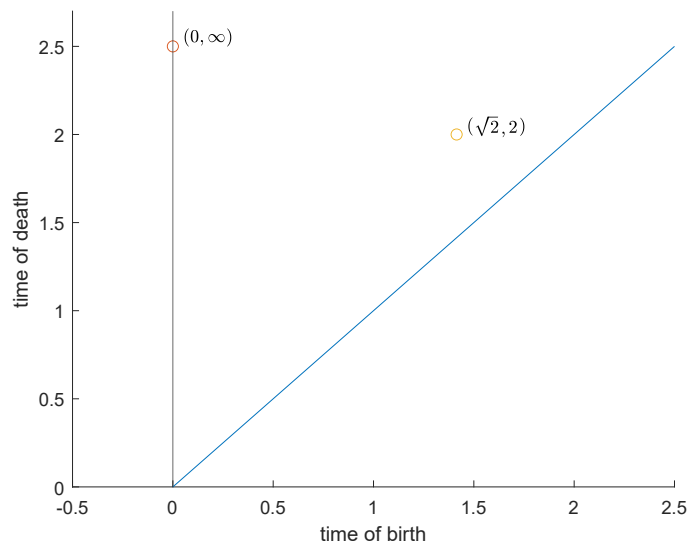


Figure 2. Persistence diagram of the function $2 \sin x$ on $\left[0, \frac{3}{4}\pi\right]$. The point $(0, \infty)$ describes the existence of a connected component that is born at zero and never dies. The point $(\sqrt{2}, 2)$ claims that there is a connected component born at $\sqrt{2}$ that dies (merges with the other one) at 2. The trivial points on the diagonal $u = v$ are not displayed.

2.4. Comparing Persistence Diagrams

Persistence diagrams can be efficiently compared by means of a suitable metric d_{match} . In order to define it, we first define the pseudo-distance

$$\delta((x, y), (x', y')) = \min\{\max\{|x - x'|, |y - y'|\}, \max\{|x - y|/2, |x' - y'|/2\}\}$$

for all $(x, y), (x', y') \in \{(x, y) \in \mathbb{R} \text{ with } x \leq y\} \cup \{(x, \infty) \text{ with } x \in \mathbb{R}\}$ by agreeing that $\infty - y = \infty, y - \infty = -\infty$ for $y \neq \infty, \infty - \infty = 0, \infty/2 = \infty, |\pm\infty| = \infty, \min\{\infty, c\} = c, \max\{\infty, c\} = \infty$.

If two persistence diagrams, D, D' , are given, we can set

$$d_{\text{match}}(D, D') = \inf_{\sigma \in \Sigma} \sup_{P \in D} \delta(P, \sigma(P))$$

where Σ represents the set of all bijections between the multisets D, D' . This function is usually called the *bottleneck distance* between D and D' .

For every degree k we can now define a new pseudo-metric:

$$D_{\mathcal{F}, \Phi}^{\text{match}}(\varphi_1, \varphi_2) = \sup_{F \in \mathcal{F}} d_{\text{match}}(D_{F(\varphi_1)}, D_{F(\varphi_2)})$$

where $D_{F(\varphi_1)}, D_{F(\varphi_2)}$ are the persistence diagrams at degree k of the functions $F(\varphi_1), F(\varphi_2)$, respectively.

In this paper we will limit ourselves to considering data represented as functions from \mathbb{R} to \mathbb{R} , and we recall that for this kind of data, persistence diagrams are non-trivial only in degree $k = 0$ (i.e., when persistent homology is used to count connected components). For this reason, in the following we will always assume $k = 0$.

3. Our Model

In this paper, we will be mainly interested in the set Lip_L of all L -Lipschitz functions from \mathbb{R} to \mathbb{R} , for some fixed constant $L \in \mathbb{R}$, and in the set $C^0(\mathbb{R})$ of all functions from \mathbb{R} to \mathbb{R} that are continuous with respect to the Euclidean topology. We will set $X = \mathbb{R}$ and consider the perception pairs $(\text{Lip}_L, G), (C^0(\mathbb{R}), G)$, where G is the group of the Euclidean isometries of \mathbb{R} .

We will assume our noise originates from a finite number of copies of a “mother” non-negative continuous bump function $\psi : \mathbb{R} \rightarrow \mathbb{R}$, such that $\text{supp}(\psi) \subseteq]-\sigma, \sigma[$ for some $\sigma > 0$ and $\|\psi\|_\infty \leq 1$. We recall that the support of a function is the closure of the set of points where the function is non-zero. After fixing two positive real numbers, η and β , the noise we will be adding is a function R belonging to the space $\mathcal{R}_{\eta, \beta}$ that contains the null function and all functions of the form $\sum_{i=1}^k a_i \psi(b_i(x - c_i))$, where k is a positive integer, and a_i, b_i, c_i are real numbers such that $|c_i - c_j| \geq \eta$ for $i \neq j$ and $b_i \geq \beta$ for every index i . For any $R \in \mathcal{R}_{\eta, \beta}$, we define $S(R) := \bigcup_{i=1}^k [c_i - \frac{\sigma}{\beta}, c_i + \frac{\sigma}{\beta}[$ and remark that if $x \notin S(R)$, then $R(x) = 0$.

Figure 3 shows how drastically persistence diagrams can change when the data are subject to impulsive noise. Our purpose will be to recover $\varphi \in \text{Lip}_L$ as well as possible from the function $\hat{\varphi} = \varphi + R$. An example of such a situation is depicted in Figure 4.

The following result will be of use.

Proposition 1. *Let F_1, F_2 be GENEOS from $(C^0(\mathbb{R}), G)$ to $(C^0(\mathbb{R}), G)$ with respect to the trivial homomorphism $T = \text{id} : G \rightarrow G$. Then $F_1 \circ F_2$ is a GENEOS from $(C^0(\mathbb{R}), G)$ to $(C^0(\mathbb{R}), G)$ with respect to T .*

Proof. For every $\varphi \in C^0(\mathbb{R}), g \in G$ we have that

$$\begin{aligned} F_1 \circ F_2(\varphi \circ g) &= F_1(F_2(\varphi \circ g)) \\ &= F_1(F_2(\varphi) \circ g) \\ &= F_1(F_2(\varphi)) \circ g \\ &= (F_1 \circ F_2)(\varphi) \circ g. \end{aligned}$$

Therefore, $F_1 \circ F_2$ is G -equivariant. Moreover, for any $\varphi_1, \varphi_2 \in C^0(\mathbb{R})$

$$\begin{aligned} D_{C^0(\mathbb{R})}(F_1 \circ F_2(\varphi_1), F_1 \circ F_2(\varphi_2)) &= D_{C^0(\mathbb{R})}(F_1(F_2(\varphi_1)), F_1(F_2(\varphi_2))) \\ &\leq D_{C^0(\mathbb{R})}(F_2(\varphi_1), F_2(\varphi_2)) \\ &\leq D_{C^0(\mathbb{R})}(\varphi_1, \varphi_2). \end{aligned}$$

It follows that $F_1 \circ F_2$ is non-expansive. \square

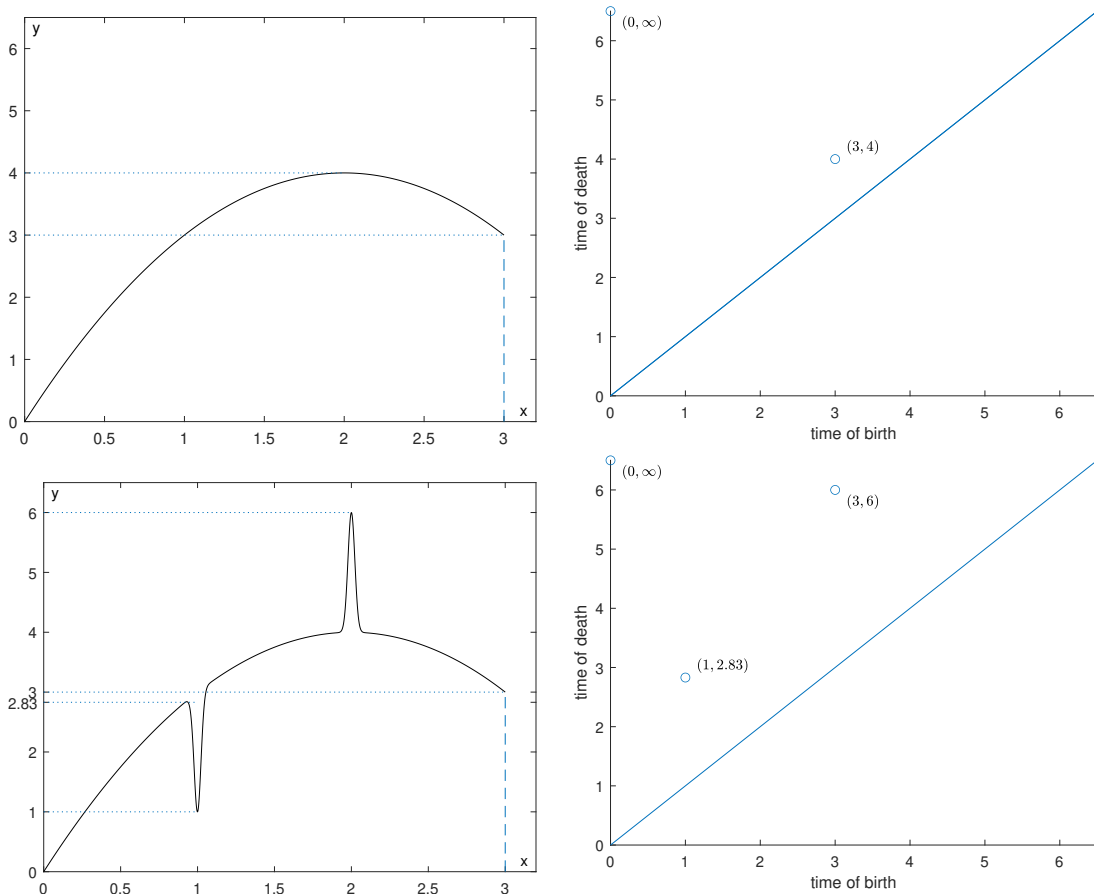


Figure 3. How drastically impulsive noise can influence persistence diagrams.

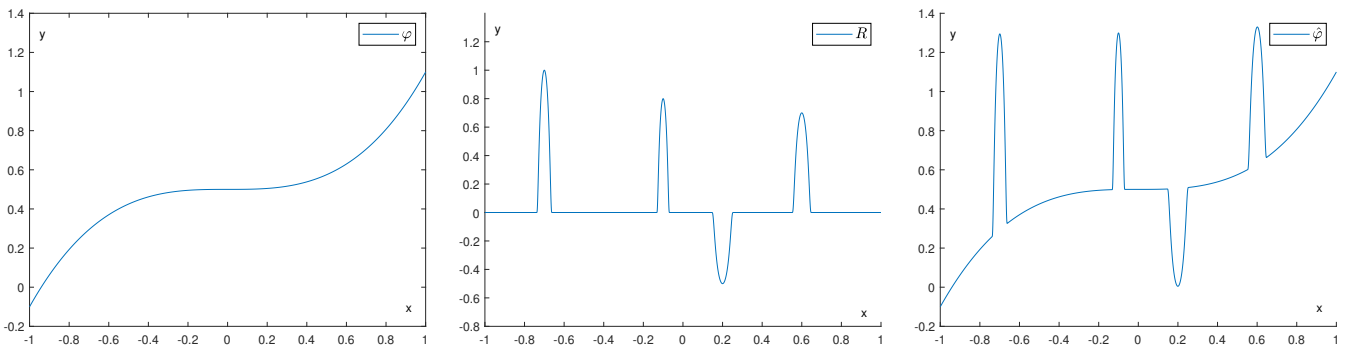


Figure 4. In the figure on the left, we have our original function φ . In the middle we have a noise function R , and in the right figure, we have the corrupted function $\hat{\varphi} := \varphi + R$.

4. Cutting Off the Noise by GENEOS

We start by introducing two families of GENEOS from $(C^0(\mathbb{R}), G)$ to $(C^0(\mathbb{R}), G)$ with respect to the identical homomorphism.

Definition 3. Let $\varphi \in \text{Lip}_L$ and $\varepsilon > 0$. For all $x \in \mathbb{R}$ we define:

1. $F^\varepsilon(\varphi)(x) = \max(\varphi(x - \varepsilon), \varphi(x + \varepsilon))$;
2. $F_\varepsilon(\varphi)(x) = \min(\varphi(x - \varepsilon), \varphi(x + \varepsilon))$.

Proposition 2. The maps F^ε and F_ε are GENEOS from $(C^0(\mathbb{R}), G)$ to $(C^0(\mathbb{R}), G)$ with respect to the identical homomorphism.

Proof. We start by proving that F^ε is G -equivariant.

Let g be the translation $x \mapsto x + k$ with $k \in \mathbb{R}$, then

$$\begin{aligned} F^\varepsilon(\varphi \circ g) &= \max\{\varphi((x + k) - \varepsilon), \varphi((x + k) + \varepsilon)\} \\ &= \max\{\varphi((x - \varepsilon) + k), \varphi((x + \varepsilon) + k)\} \\ &= F^\varepsilon(\varphi) \circ g. \end{aligned}$$

Let g be the symmetry $x \mapsto -x$, then

$$\begin{aligned} F^\varepsilon(\varphi \circ g) &= \max\{\varphi((-x) - \varepsilon), \varphi((-x) + \varepsilon)\} \\ &= \max\{\varphi(-(x + \varepsilon)), \varphi(-(x - \varepsilon))\} \\ &= \max\{\varphi(x + \varepsilon), \varphi(x - \varepsilon)\} \circ g \\ &= F^\varepsilon(\varphi) \circ g. \end{aligned}$$

Since every isometry in G can be written as the composition of a symmetry and a translation, our statement follows.

Furthermore, for any $x \in \mathbb{R}$

$$\begin{aligned} |F^\varepsilon(\varphi_1)(x) - F^\varepsilon(\varphi_2)(x)| &= \\ &= |\max\{\varphi_1(x - \varepsilon), \varphi_1(x + \varepsilon)\} - \max\{\varphi_2(x - \varepsilon), \varphi_2(x + \varepsilon)\}| \\ &\leq \max\{|\varphi_1(x - \varepsilon) - \varphi_2(x - \varepsilon)|, |\varphi_1(x + \varepsilon) - \varphi_2(x + \varepsilon)|\} \\ &\leq \max\{\|\varphi_1 - \varphi_2\|_\infty, \|\varphi_1 - \varphi_2\|_\infty\} \\ &= \|\varphi_1 - \varphi_2\|_\infty. \end{aligned}$$

It follows that $\|F^\varepsilon(\varphi_1) - F^\varepsilon(\varphi_2)\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty$, and hence F^ε is non-expansive.

We observe that $\min\{a, b\} = -\max\{-a, -b\}$, and hence $F_\varepsilon(\varphi) = -F^\varepsilon(-\varphi)$ for every $\varphi \in C^0(\mathbb{R})$. Moreover, the map that takes φ to $-\varphi$ is a GENEOS from $(C^0(\mathbb{R}), G)$ to $(C^0(\mathbb{R}), G)$ with respect to the identical homomorphism. Therefore, Proposition 1 ensures that G -equivariance and non-expansivity also hold for F_ε . \square

Since Proposition 1 shows that the composition of GENEOS is still a GENEOS, the operator $F^\delta \circ F_\varepsilon$ is a GENEOS from $(C^0(\mathbb{R}), G)$ to $(C^0(\mathbb{R}), G)$ with respect to the identical homomorphism.

We now want to prove that if a function $\hat{\varphi}$ is obtained by adding impulsive noise to a function φ , then the value of

$$\|F^\delta \circ F_\varepsilon(\hat{\varphi}) - \varphi\|_\infty$$

is bounded, and possibly small, provided that δ and ε are chosen appropriately. The main idea is that the operator F_ε cuts the noise “directed upwards” and F^δ cuts the noise “directed downwards”.

In order to proceed, we need two lemmas.

Lemma 1. Let $R \in C^0(\mathbb{R})$, $\varphi \in \text{Lip}_L$ for some $L \in \mathbb{R}$, and set $\hat{\varphi} := \varphi + R$. Then for any $\varepsilon > 0$ and $\delta > 0$

- (i) $-L\varepsilon + F_\varepsilon(R) \leq F_\varepsilon(\hat{\varphi}) - \varphi \leq L\varepsilon + F_\varepsilon(R)$;
- (ii) $-L\delta + F^\delta(R) \leq F^\delta(\hat{\varphi}) - \varphi \leq L\delta + F^\delta(R)$;

$$(iii) \quad -L(\delta + \varepsilon) + F^\delta \circ F_\varepsilon(R) \leq F^\delta \circ F_\varepsilon(\hat{\varphi}) - \varphi \leq L(\delta + \varepsilon) + F^\delta \circ F_\varepsilon(R).$$

Proof. Since φ is Lipschitz of constant L , we have that for any value $x \in \mathbb{R}$ $|\varphi(x - \varepsilon) - \varphi(x)| \leq L\varepsilon$ and $|\varphi(x + \varepsilon) - \varphi(x)| \leq L\varepsilon$. Therefore,

$$\begin{aligned} F_\varepsilon(\hat{\varphi})(x) &= F_\varepsilon(\varphi + R)(x) \\ &= \min\{\varphi(x - \varepsilon) + R(x - \varepsilon), \varphi(x + \varepsilon) + R(x + \varepsilon)\} \\ &\leq \min\{\varphi(x) + L\varepsilon + R(x - \varepsilon), \varphi(x) + L\varepsilon + R(x + \varepsilon)\} \\ &= \varphi(x) + L\varepsilon + \min\{R(x - \varepsilon), R(x + \varepsilon)\} \\ &= \varphi(x) + L\varepsilon + F_\varepsilon(R)(x). \end{aligned}$$

Analogously, $F_\varepsilon(\hat{\varphi})(x) \geq \varphi(x) - L\varepsilon + F_\varepsilon(R)(x)$. The same steps applied to F^δ yield the second statement of the lemma. As for the last claim, we can see that:

$$\begin{aligned} F^\delta \circ F_\varepsilon(\hat{\varphi})(x) &= \max\{F_\varepsilon(\hat{\varphi})(x - \delta), F_\varepsilon(\hat{\varphi})(x + \delta)\} \\ &\leq \max\{\varphi(x - \delta) + L\varepsilon + F_\varepsilon(R)(x - \delta), \varphi(x + \delta) + L\varepsilon + F_\varepsilon(R)(x + \delta)\} \\ &\leq \max\{\varphi(x) + L\delta + L\varepsilon + F_\varepsilon(R)(x - \delta), \varphi(x) + L\delta + L\varepsilon + F_\varepsilon(R)(x + \delta)\} \\ &= \varphi(x) + L\delta + L\varepsilon + \max\{F_\varepsilon(R)(x - \delta), F_\varepsilon(R)(x + \delta)\} \\ &= \varphi(x) + L\delta + L\varepsilon + F^\delta \circ F_\varepsilon(R)(x). \end{aligned}$$

Analogously, we can prove the lower bound. \square

Henceforth, we will assume that any summation on an empty set of indexes is the null function.

Lemma 2. Let $R \in \mathcal{R}_{\eta,\beta}$ and $\lambda \geq \frac{\sigma}{\beta}$. If $\lambda \leq \rho \leq \frac{\eta}{2} - \lambda$ then

$$(a) \quad F_\rho(R)(x) = \sum_{a_i < 0} [a_i \psi(b_i(x - \rho - c_i)) + a_i \psi(b_i(x + \rho - c_i))] \leq 0$$

$$(b) \quad F^\rho(R)(x) = \sum_{a_i > 0} [a_i \psi(b_i(x - \rho - c_i)) + a_i \psi(b_i(x + \rho - c_i))] \geq 0$$

for all $x \in \mathbb{R}$. Moreover $F_\rho(R), F^\rho(R)(x) \in \mathcal{R}_{2\lambda,\beta}$.

Proof. We will suppose without loss of generality that $c_i < c_{i+1}$ for all $i = 1, \dots, k - 1$ and $a_i \neq 0$ for all indexes i .

We want to show that at least one of $x - \rho$ and $x + \rho$ must always be outside of $\bigcup_{i=1}^k]c_i - \lambda, c_i + \lambda[$. If by contradiction both $x - \rho$ and $x + \rho$ were in the same $]c_i - \lambda, c_i + \lambda[$, then we would have $\rho < \lambda$, against our hypotheses. Moreover, $x - \rho$ and $x + \rho$ cannot belong to different intervals $]c_i - \lambda, c_i + \lambda[$ and $]c_j - \lambda, c_j + \lambda[$ for $i < j$: if by contradiction $x - \rho \in]c_i - \lambda, c_i + \lambda[$ and $x + \rho \in]c_j - \lambda, c_j + \lambda[$ for some $i < j$, then $2\rho = x + \rho - (x - \rho) > (c_j - \lambda) - (c_i + \lambda) \geq c_j - c_i - 2\lambda \geq \eta - 2\lambda$, and hence we would have $\rho > \frac{\eta}{2} - \lambda$, against our hypotheses.

Since $]c_i - \frac{\sigma}{\beta}, c_i + \frac{\sigma}{\beta}[\subseteq]c_i - \lambda, c_i + \lambda[$, it follows that at least one among the values $R(x - \rho)$ and $R(x + \rho)$ must always be zero. Let us now set $I_i^- :=](c_i - \rho) - \lambda, (c_i - \rho) + \lambda[$ and $I_i^+ :=](c_i + \rho) - \lambda, (c_i + \rho) + \lambda[$. These two intervals must be disjoint, since $(c_i + \rho - \lambda) - (c_i - \rho + \lambda) = 2\rho - 2\lambda \geq 0$.

Let us now consider $\{I_1^-, I_1^+, \dots, I_k^-, I_k^+\}$. We will now prove that any two distinct elements from this set must be a disjoint. Since we have just proven that $I_i^- \cap I_i^+ = \emptyset$ for $i = 1, \dots, k$, the following holds in the case $i < j$:

$$1. \quad I_i^+ \cap I_j^+ = \emptyset, \text{ since } c_j + \rho - \lambda - (c_i + \rho + \lambda) \geq c_j - c_i - 2\lambda \geq \eta - 2\lambda \geq 0;$$

2. $I_i^- \cap I_j^- = \emptyset$, since $c_j - \rho - \lambda - (c_i - \rho + \lambda) \geq c_j - c_i - 2\lambda \geq \eta - 2\lambda \geq 0$;
3. $I_i^+ \cap I_j^- = \emptyset$, since $c_j - \rho - \lambda - (c_i + \rho + \lambda) \geq c_j - c_i - 2\rho - 2\lambda \geq \eta - 2(\frac{\eta}{2} - \lambda) - 2\lambda = 0$;
4. $I_j^+ \cap I_i^-$, since $c_j + \rho - \lambda - (c_i - \rho + \lambda) \geq c_j - c_i + 2\rho - 2\lambda \geq \eta + 2\lambda - 2\lambda \geq 0$.

Let us now fix $k \in \{1, \dots, k\}$.

Since $\text{supp}(a_k \psi(b_k(x - c_k))) \subseteq]c_k - \lambda, c_k + \lambda[$, then $\text{supp}(a_k \psi(b_k(x - \rho - c_k))) \subseteq I_k^-$ and $\text{supp}(a_k \psi(b_k(x + \rho - c_k))) \subseteq I_k^+$.

Hence, for any $x \in \mathbb{R}$ we have that

$$F_\rho(R)(x) = \min\{R(x - \rho), R(x + \rho)\} = \begin{cases} 0 & \text{if } x \notin \bigcup_{i=1}^k (I_i^- \cup I_i^+) \\ a_j \psi(b_j(x + \rho - c_j)) & \text{if } x \in I_j^- \text{ and } a_j < 0 \\ a_j \psi(b_j(x - \rho - c_j)) & \text{if } x \in I_j^+ \text{ and } a_j < 0. \end{cases}$$

This means that

$$F_\rho(R)(x) = \sum_{a_i < 0} [a_i \psi(b_i(x - \rho - c_i)) + a_i \psi(b_i(x + \rho - c_i))] \leq 0.$$

Now, given $c_i + \rho$ and $c_j - \rho$ centers of bumps of $F_\rho(R)$, we have that

$$|(c_i + \rho) - (c_j - \rho)| \geq \min\{\eta - 2\rho, 2\rho\} \geq 2\lambda.$$

It follows that $F_\rho(R) \in \mathcal{R}_{2\lambda, \beta}$.

By noting that $F^\rho(R) = -F_\rho(-R)$, we obtain the second part of the thesis. \square

Let us remark that, in particular, if $\lambda = 2\frac{\sigma}{\beta}$, then $F_\rho(R), F^\rho(R) \in \mathcal{R}_{4\frac{\sigma}{\beta}, \beta}$. We observe that for there to exist a ρ that satisfies the hypotheses of Lemma 2, the inequality $\eta \geq 4\lambda$ must hold.

We are now actually ready to prove a result that will be useful in the remainder of our paper.

Theorem 2. Given $\theta, \beta, L \in \mathbb{R}$, let $\bar{R} \in \mathcal{R}_{\theta, \beta}$, $\varphi \in \text{Lip}_L$ and set $\hat{\varphi} := \varphi + \bar{R}$. If $2\frac{\sigma}{\beta} \leq \varepsilon \leq \frac{\theta}{2} - 2\frac{\sigma}{\beta}$ then for any $\frac{\sigma}{\beta} \leq \delta \leq \frac{1}{2} \min\{\theta - 2\varepsilon, 2\varepsilon\} - \frac{\sigma}{\beta}$ the following inequality holds:

$$\|F^\delta \circ F_\varepsilon(\hat{\varphi}) - \varphi\|_\infty \leq L(\varepsilon + \delta).$$

Proof. Let $x \in \mathbb{R}$. Firstly, let us notice that the condition $2\frac{\sigma}{\beta} \leq \varepsilon \leq \frac{\theta}{2} - 2\frac{\sigma}{\beta}$ implies that $\frac{\sigma}{\beta} \leq \frac{1}{2} \min\{\theta - 2\varepsilon, 2\varepsilon\} - \frac{\sigma}{\beta}$. From Lemma 1, we know that

$$-L(\delta + \varepsilon) + F^\delta \circ F_\varepsilon(\bar{R})(x) \leq F^\delta \circ F_\varepsilon(\hat{\varphi})(x) - \varphi(x) \leq L(\delta + \varepsilon) + F^\delta \circ F_\varepsilon(\bar{R})(x).$$

Let us now prove that $F^\delta \circ F_\varepsilon(\bar{R})(x) = 0$.

By applying Lemma 2 with $\eta := \theta, \lambda := 2\frac{\sigma}{\beta}, \rho := \varepsilon$ and $R := \bar{R}$, we obtain

$$F_\varepsilon(\bar{R})(x) = \sum_{a_i < 0} [a_i \psi(b_i(x - \varepsilon - c_i)) + a_i \psi(b_i(x + \varepsilon - c_i))] \leq 0.$$

Let us remark that $F_\varepsilon(\bar{R}) \in \mathcal{R}_{4\frac{\sigma}{\beta}, \beta}$. Moreover,

1. $F^\delta \circ F_\varepsilon(\bar{R})(x) = \max\{F_\varepsilon(\bar{R})(x - \delta), F_\varepsilon(\bar{R})(x + \delta)\} \leq 0$ since both terms are negative;
2. $F^\delta \circ F_\varepsilon(\bar{R})(x) = F^\delta(F_\varepsilon(\bar{R}))(x) \geq 0$.

These inequalities follow from Lemma 2, by setting $\eta := 4\frac{\sigma}{\beta}, \lambda := \frac{\sigma}{\beta}, \rho := \delta$ and $R := F_\varepsilon(\bar{R})$.

Therefore, we have proved that $|F^\delta \circ F_\varepsilon(\hat{\varphi})(x) - \varphi(x)| \leq L(\delta + \varepsilon)$ for any $x \in \mathbb{R}$. It follows that $\|F^\delta \circ F_\varepsilon(\hat{\varphi}) - \varphi\|_\infty \leq L(\varepsilon + \delta)$. \square

Let us remark that Theorem 2 works under the (only) implicit assumption that $\theta \geq 8\frac{\sigma}{\beta}$. In our setting, this should not be restrictive since it means that the noise added is made up of scattered, thin bumps, without any reference to the height of the bumps. This is what we expect when considering additive impulsive noise.

Corollary 1. *Given $\theta, \beta, L \in \mathbb{R}$, let $R \in \mathcal{R}_{\theta, \beta}$, $\varphi \in \text{Lip}_L$, and set $\hat{\varphi} := \varphi + R$. If $\theta \geq 8\frac{\sigma}{\beta}$, then*

$$\left\| F^{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi \right\|_\infty \leq 3L\frac{\sigma}{\beta}.$$

Proof. The claim follows from Theorem 2 by taking $\varepsilon := 2\frac{\sigma}{\beta}$ and $\delta := \frac{\sigma}{\beta}$. \square

Corollary 1 and the well-known stability of persistence diagrams with respect to the max-norm [5] immediately imply the following result, which is of interest in TDA (the symbol d_{match} denotes the usual bottleneck distance between persistence diagrams).

Corollary 2. *Given $\theta, \beta, L \in \mathbb{R}$, let $R \in \mathcal{R}_{\theta, \beta}$, $\varphi \in \text{Lip}_L$, and set $\hat{\varphi} := \varphi + R$. If $\theta \geq 8\frac{\sigma}{\beta}$, and D and D' are the persistence diagrams in degree 0 of the filtering functions φ and $F^{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi})$, respectively, then*

$$d_{\text{match}}(D, D') \leq 3L\frac{\sigma}{\beta}.$$

5. Our Main Results

We are now ready to prove our main results. We start by recalling a simple statement concerning the probability p that any two distinct points in a randomly chosen set of cardinality k in an interval of length ℓ have a distance greater than η (cf., e.g., [19]). For the reader’s convenience, we report the proof here.

Lemma 3. *Let X_1, \dots, X_k , with $k \geq 2$, be independent random variables, uniformly distributed on the interval $[0, \ell]$, for some $\ell > 0$. Let*

$$M := \min_{\substack{1 \leq i, j \leq k \\ i \neq j}} |X_i - X_j|$$

be the minimal distance between two distinct random variables. Then, we have

$$P(M > \eta) = \begin{cases} 1 & \text{if } \eta \leq 0, \\ \left(1 - \frac{(k-1)\eta}{\ell}\right)^k & \text{if } 0 < \eta < \frac{\ell}{k-1}, \\ 0 & \text{if } \eta \geq \frac{\ell}{k-1}. \end{cases}$$

Proof. It suffices to consider the case $0 < \eta < \frac{\ell}{k-1}$. By symmetry, we have

$$P(M > \eta) = k! P((M > \eta) \cap (X_1 < X_2 < \dots < X_k))$$

(since X_1, \dots, X_k are uniformly distributed)

$$= k! \frac{\text{Leb}(S)}{\ell^k} \tag{1}$$

where $S = \{x \in [0, \ell]^k \mid x_1 < x_2 - \eta < x_3 - 2\eta < \dots < x_k - (k - 1)\eta\}$, and Leb denotes the Lebesgue measure. Setting $y_i = x_i - (i - 1)\eta$ for $i = 1, \dots, k$, we have that $\text{Leb}(S) = \text{Leb}(S')$ where

$$S' = \{y \in [0, \ell - (k - 1)\eta]^k \mid y_1 < y_2 < \dots < y_k\}.$$

On the other hand, again by symmetry, we have

$$\text{Leb}(S') = \frac{\text{Leb}([0, \ell - (k - 1)\eta]^k)}{k!} = \frac{(\ell - (k - 1)\eta)^k}{k!}$$

and plugging this last identity into (1) we obtain the thesis. \square

We can now prove the following result, concerning the expected value of the error $\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty}$. We remark explicitly that the error is random because it is a function of c_1, \dots, c_k , which are independent random variables uniformly distributed on the interval $[0, \ell]$.

Theorem 3. *Let us choose a function $\varphi \in \text{Lip}_L$, a non-negative continuous function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ with $\|\psi\|_{\infty} \leq 1$ and $\text{supp}(\psi) \subseteq]-\sigma, \sigma[$ for some $\sigma > 0$, two positive numbers β and ℓ , and an integer $k \geq 2$. For $i = 1, \dots, k$, let us fix $a_i \in \mathbb{R}$ and $b_i \geq \beta$, and set $\bar{\alpha} := \max|a_i|$. Moreover, let c_1, \dots, c_k be independent random variables, uniformly distributed on the interval $[0, \ell]$. Let us consider the random variable $\hat{\varphi} := \varphi + R$, where $R(x) := \sum_{i=1}^k a_i \psi(b_i(x - c_i))$ for any $x \in \mathbb{R}$. If $\frac{\sigma}{\beta} < \frac{\ell}{8(k-1)}$, then*

$$E\left(\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty}\right) \leq 3L\frac{\sigma}{\beta} + k\bar{\alpha} \left(1 - \left(1 - 8\frac{(k-1)\sigma}{\ell\beta}\right)^k\right).$$

Proof. By setting $\delta = \frac{\sigma}{\beta}$ and $\varepsilon = 2\frac{\sigma}{\beta}$ in statement (iii) of Lemma 1, we have that

$$\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty} \leq 3L\frac{\sigma}{\beta} + \left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(R)\right\|_{\infty}.$$

Since the operator $F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}$ is non-expansive and $F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\mathbf{0}) = \mathbf{0}$, it follows that

$$\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(R)\right\|_{\infty} \leq \|R\|_{\infty} \leq k\bar{\alpha}\|\psi\|_{\infty} \leq k\bar{\alpha}.$$

Therefore, $\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty} \leq 3L\frac{\sigma}{\beta} + k\bar{\alpha}$. If we apply Lemma 3 with $\eta = 8\frac{\sigma}{\beta}$, we obtain that $R \in \mathcal{R}_{8\frac{\sigma}{\beta}, \beta}$ with probability $p := \left(1 - 8\frac{(k-1)\sigma}{\ell\beta}\right)^k$. If $R \in \mathcal{R}_{8\frac{\sigma}{\beta}, \beta}$, we can apply Theorem 2 by setting $\delta = \frac{\sigma}{\beta}$, $\varepsilon = 2\frac{\sigma}{\beta}$ and $\theta = 8\frac{\sigma}{\beta}$, and hence we obtain that $\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty} \leq 3L\frac{\sigma}{\beta}$ with probability at least p . Since $\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty} \leq 3L\frac{\sigma}{\beta} + k\bar{\alpha}$ in any case, it follows that

$$\begin{aligned} E\left(\left\|F_{\frac{\sigma}{\beta}}^{\sigma} \circ F_{2\frac{\sigma}{\beta}}^{\sigma}(\hat{\varphi}) - \varphi\right\|_{\infty}\right) &\leq 3L\frac{\sigma}{\beta}p + \left(3L\frac{\sigma}{\beta} + k\bar{\alpha}\right)(1 - p) \\ &= 3L\frac{\sigma}{\beta} + k\bar{\alpha}(1 - p). \end{aligned}$$

\square

Theorem 3 and the well-known stability of persistence diagrams with respect to the max-norm [5] immediately imply the following result, which is of interest in TDA (the symbol d_{match} denotes the usual bottleneck distance between persistence diagrams).

Corollary 3. *Let us make the same assumptions of Theorem 3. Let D and D' be the persistence diagrams in degree 0 of the filtering functions φ and $F^{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi})$, respectively. If $\frac{\sigma}{\beta} < \frac{\ell}{8(k-1)}$, then*

$$E(d_{\text{match}}(D, D')) \leq 3L\frac{\sigma}{\beta} + k\bar{\alpha} \left(1 - \left(1 - 8\frac{(k-1)\sigma}{\ell\beta} \right)^k \right).$$

This result shows that the use of suitable GENEOS can make TDA (relatively) stable also in the presence of impulsive noise, under the assumptions we have considered in this paper.

Theorem 3 and Corollary 3 can be easily extended to the case that the number k of bumps is a random variable:

Theorem 4. *Let k be a random variable that takes values in the subset $S = \{2, 3, \dots, K\}$ of the positive integers, where we assume that the probability of the integer $j \in S$ is p_j . Let us choose a function $\varphi \in \text{Lip}_L$, a non-negative continuous function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ with $\|\psi\|_{\infty} \leq 1$ and $\text{supp}(\psi) \subseteq]-\sigma, \sigma[$ for some $\sigma > 0$, two positive numbers β and ℓ . Let c_1^k, \dots, c_k^k be random variables that, conditioned to k , are independent and uniformly distributed on the interval $[0, \ell]$. Moreover, for $i = 1, \dots, k$, let us fix $a_i^k \in \mathbb{R}$ and $b_i^k \geq \beta$, and set $\bar{\alpha}_k := \max_i |a_i^k|$. Let us also consider the random variable $\hat{\varphi} := \varphi + R_k$, where $R_k(x) := \sum_{i=1}^k a_i^k \psi(b_i^k(x - c_i^k))$ for any $x \in \mathbb{R}$. If $\frac{\sigma}{\beta} < \frac{\ell}{8(K-1)}$, then*

$$E\left(\left\|F^{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\right\|_{\infty}\right) \leq 3L\frac{\sigma}{\beta} + \sum_{j=2}^K p_j j \bar{\alpha}_j \left(1 - \left(1 - 8\frac{(j-1)\sigma}{\ell\beta} \right)^j \right).$$

Proof. It follows immediately from Theorem 3. \square

Corollary 4. *Let us make the same assumptions of Theorem 4. Let D and D' be the persistence diagrams in degree 0 of the filtering functions φ and $F^{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi})$, respectively. If $\frac{\sigma}{\beta} < \frac{\ell}{8(K-1)}$, then*

$$E(d_{\text{match}}(D, D')) \leq 3L\frac{\sigma}{\beta} + \sum_{j=2}^K p_j j \bar{\alpha}_j \left(1 - \left(1 - 8\frac{(j-1)\sigma}{\ell\beta} \right)^j \right).$$

Proof. It follows immediately from Corollary 3. \square

6. Examples and Experiments

We will now validate our approach based on GENEOS by giving two examples and illustrating some experimental results.

6.1. Examples

In order to verify how our approach works, we will set $\tau_n := \left(1 - \frac{1}{n}\right)2\frac{\sigma}{\beta} + \frac{1}{n}\left(\frac{\theta}{2} - 2\frac{\sigma}{\beta}\right)$ and consider the upper bound $\left\|F^{\frac{\tau_n}{2}} \circ F_{\tau_n}(\hat{\varphi}) - \varphi\right\|_{\infty} \leq \frac{3}{2}L\tau_n$, obtained by applying Corollary 1. We observe that $\tau_n \geq 2\frac{\sigma}{\beta}$ for every index n , and $\lim_{n \rightarrow +\infty} \tau_n = 2\frac{\sigma}{\beta}$. We will examine two examples that use the GENEOS $F^{\frac{\tau_n}{2}} \circ F_{\tau_n}$ and show how our method based on such operators and the method based on convolutions differ, as for their capability in removing additive impulsive noise. Moreover, we will compare the actual error $\left\|F^{\frac{\tau_n}{2}} \circ F_{\tau_n}(\hat{\varphi}) - \varphi\right\|_{\infty}$ to its

upper bound $\frac{3}{2}L\tau_n$, by running several simulations. The convolutions that will be applied in our examples use the functions $T_h : \mathbb{R} \rightarrow \mathbb{R}$ defined by setting

$$T_h(x) := \begin{cases} \frac{h}{2} & \text{if } -\frac{1}{h} \leq x \leq \frac{1}{h} \\ 0 & \text{otherwise} \end{cases}$$

for $h > 0$. We will see that, although the convolution with such functions is also a GENEO, it will not be able to efficiently remove the noise.

Our noise function R will be constructed starting from the mother function ψ defined by setting $\psi(x) := e^{1-\frac{1}{1-x^2}}$ for $x \in]-1, 1[$ and $\psi(x) := 0$ for $x \notin]-1, 1[$. Using the notation introduced in the previous sections, we will set $\sigma = 1.1$, thus satisfying the condition $\text{supp}(\psi) \subseteq]-\sigma, \sigma[$. The impulsive noise will be added in an interval $[-\ell, \ell]$. The following parameters are considered, with these respective uniform distributions:

- $k \sim \text{Unif}_{\{1, \dots, 10\}}$
- $a_i \sim \text{Unif}_{]-100, 100[}$ for $i = 1, \dots, k$
- $b_i \sim \text{Unif}_{]0, 100[}$ for $i = 1, \dots, k$
- $c_i \sim \text{Unif}_{]-\ell + \frac{\sigma}{\beta}, \ell - \frac{\sigma}{\beta}[}$ for $i = 1, \dots, k$.

We set $\beta := \min_{i=1, \dots, N} b_i$, $\bar{\alpha} := \max_{i=1, \dots, N} |a_i|$, and $\eta := \min_{i \neq j} |c_i - c_j|$. After producing random values for the parameters N, a_i, b_i, c_i , our algorithm checks whether $\eta > 8\frac{\sigma}{\beta}$; otherwise it generates another set of parameters.

6.1.1. First Example

Let us consider the function

$$\varphi(x) := \begin{cases} \sin x & \text{if } -4\pi \leq x \leq 4\pi \\ 0 & \text{otherwise} \end{cases}$$

for $x \in \mathbb{R}$. We observe that $\varphi \in \text{Lip}_L$, for $L = 1$. We will add noise in the interval $[-4\pi, 4\pi]$ (meaning $\ell = 4\pi$) and visualize the results in such an interval. Figure 5 illustrates how the function $\hat{\varphi}$ looks compared to φ .

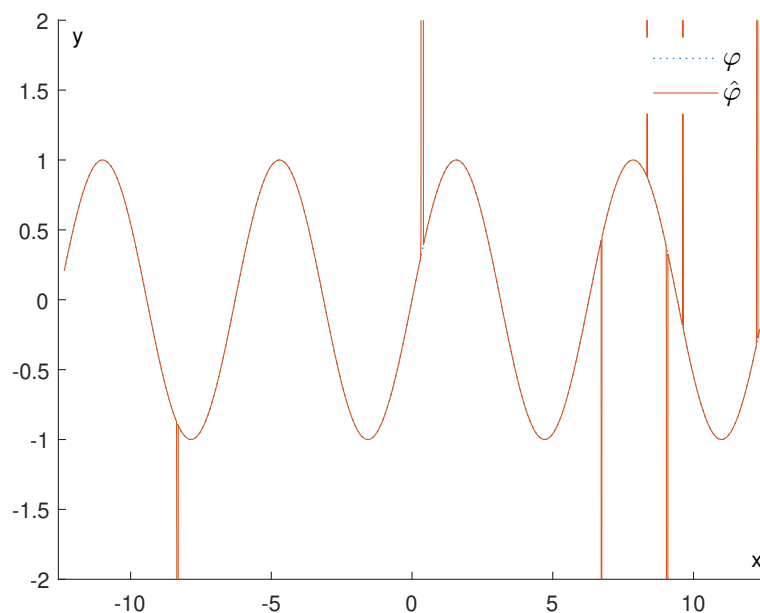


Figure 5. Example 1: Comparison between φ and $\hat{\varphi}$.

We will start by considering how well the convolution $\hat{\varphi} * T_n$ can approximate the original function φ , when n goes from 3 to 100. From Figure 6, it is immediately apparent that the max-norm distance between $\hat{\varphi} * T_n$ and φ remains quite large.

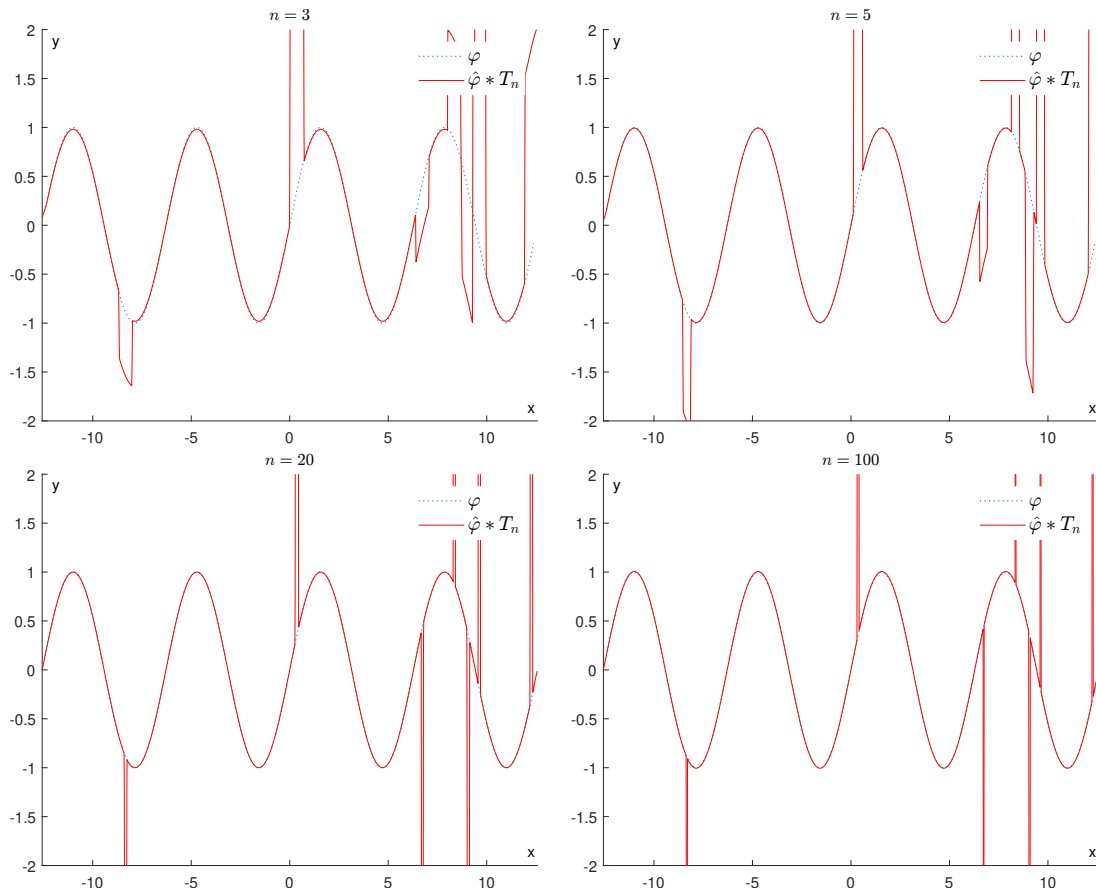


Figure 6. Example 1: Denoising via convolution with T_h ($h = n$).

If we apply a convolution with $T_{\frac{1}{n}}$, for $3 \leq n \leq 100$, we obtain the results displayed in Figure 7, showing that all information represented by the function φ is progressively destroyed.

In contrast, if we apply the operator $F_{\frac{\sigma}{2}} \circ F_{\tau_n}(\hat{\varphi})$, for $n = 3, 5, 20, 100$, we obtain the results displayed in Figure 8. As we can see, this operator is much more efficient in removing the bumps and restoring the function φ . We recall that as n increases, τ_n tends to $2\frac{\sigma}{\beta}$, and our operator becomes more effective at removing the noise.

As a matter of fact, when we apply a convolution with the function T_h and check the corresponding errors via the sup-norm, we obtain the results displayed in Figure 9. Since $\lim_{h \rightarrow +\infty} \hat{\varphi} * T_h = \hat{\varphi}$ and $\lim_{h \rightarrow +\infty} \hat{\varphi} * T_{\frac{1}{h}} = 0$, we obtain that the errors tend to $\|\hat{\varphi} - \varphi\|_{\infty} = \max_{i=1, \dots, N} |a_i| = \bar{a}$ and $\|\varphi\|_{\infty}$, respectively.

In contrast, if we apply the operator $F_{\frac{\sigma}{2}} \circ F_{\tau_n}$, we obtain the results displayed in Figure 10. This shows that the upper bound for the error stated in Corollary 1 is quite tight. As we can expect, the best denoising is achieved when we replace τ_n with $2\frac{\sigma}{\beta}$ (see Figure 11).

Finally, we executed 1000 simulations. In each simulation, we added random impulsive noise to the function φ to produce a function $\hat{\varphi}$. We then applied $F_{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}$ to $\hat{\varphi}$ to see how close the upper bound in Corollary 1 is to the actual error $\|F_{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\|_{\infty}$. The same parameters as in the beginning of this example were used. As we can see in Figure 12,

the overestimation committed by our upper bound is often quite close to zero relative to the Lipschitz constant L of the function φ .

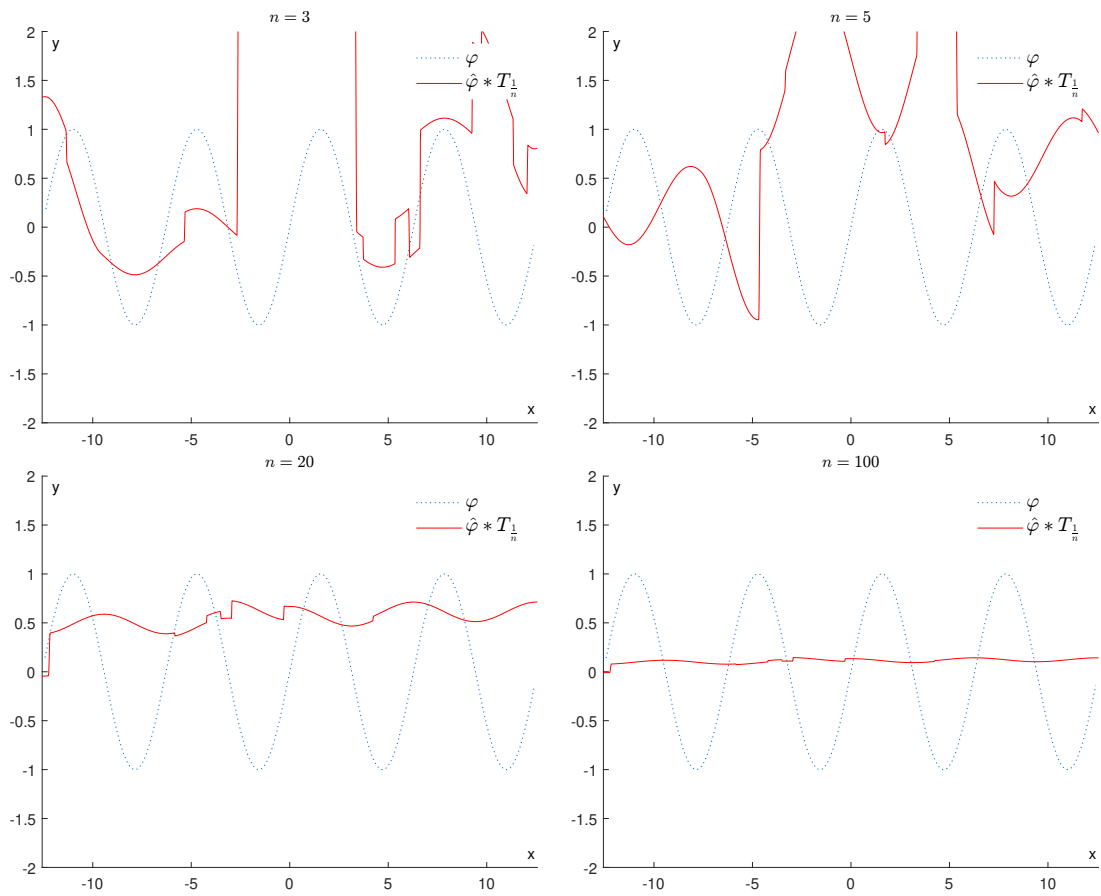


Figure 7. Example 1: Denoising via convolution with T_h ($h = \frac{1}{n}$).

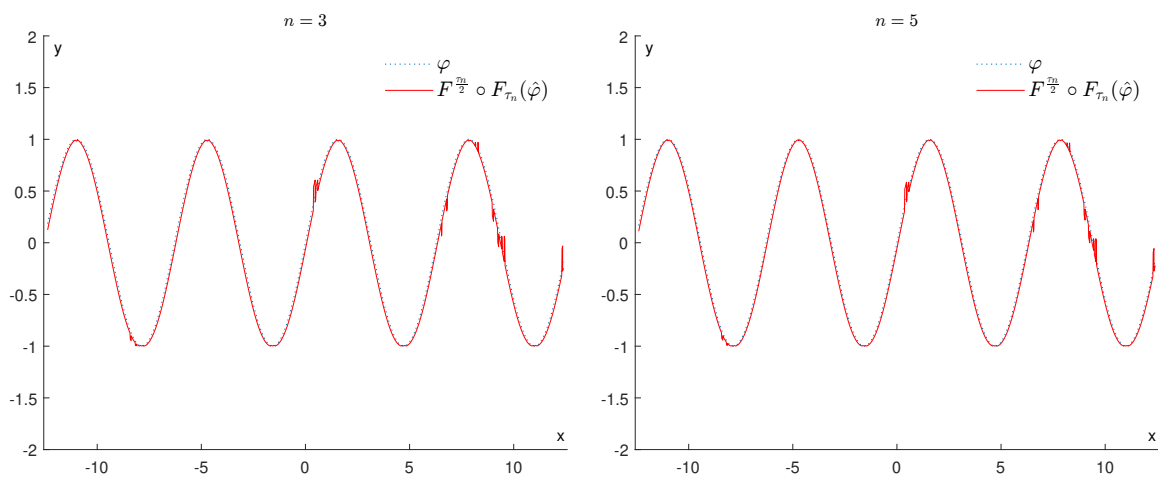


Figure 8. Cont.

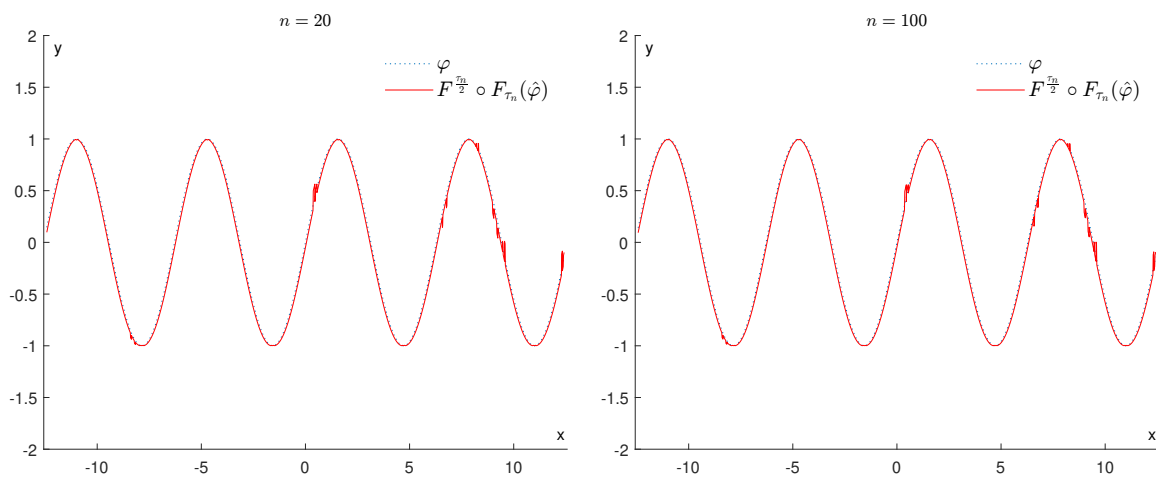


Figure 8. Example 1: Denoising via the proposed GENE $F^{\frac{\tau_n}{2}} \circ F_{\tau_n}$.

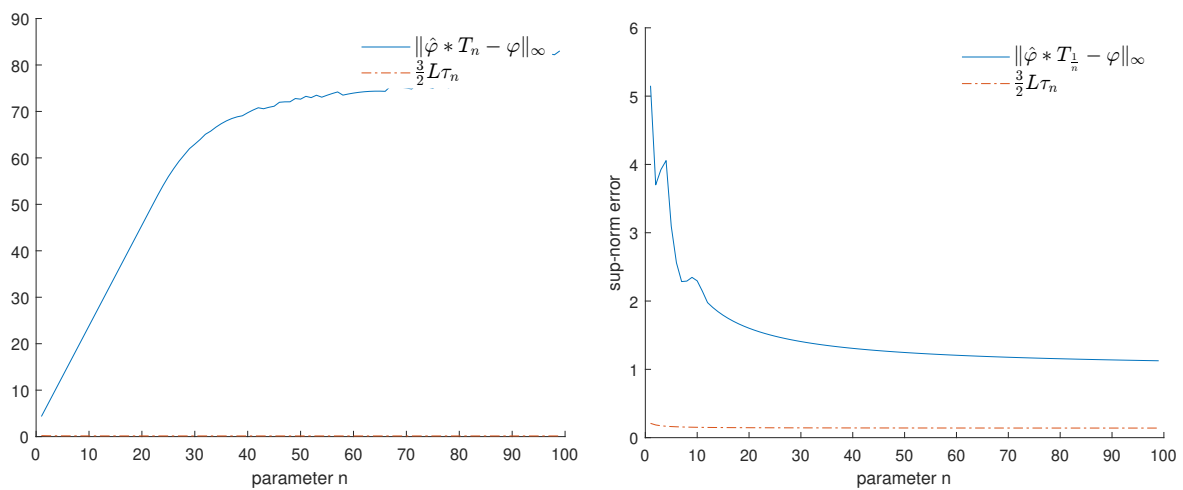


Figure 9. Example 1: Error made by applying a convolution with T_n (left) or $T_{\frac{1}{n}}$ (right).

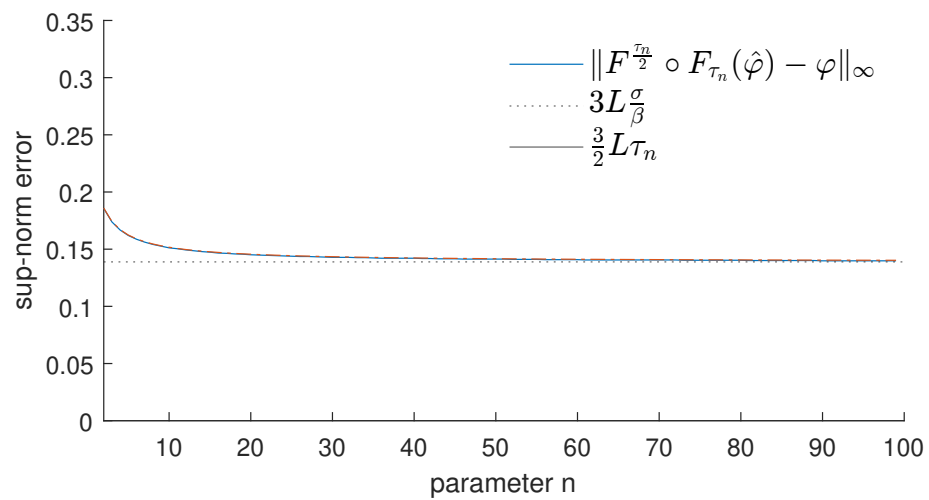


Figure 10. Example 1: Error made by using the GENE $F^{\frac{\tau_n}{2}} \circ F_{\tau_n}$.

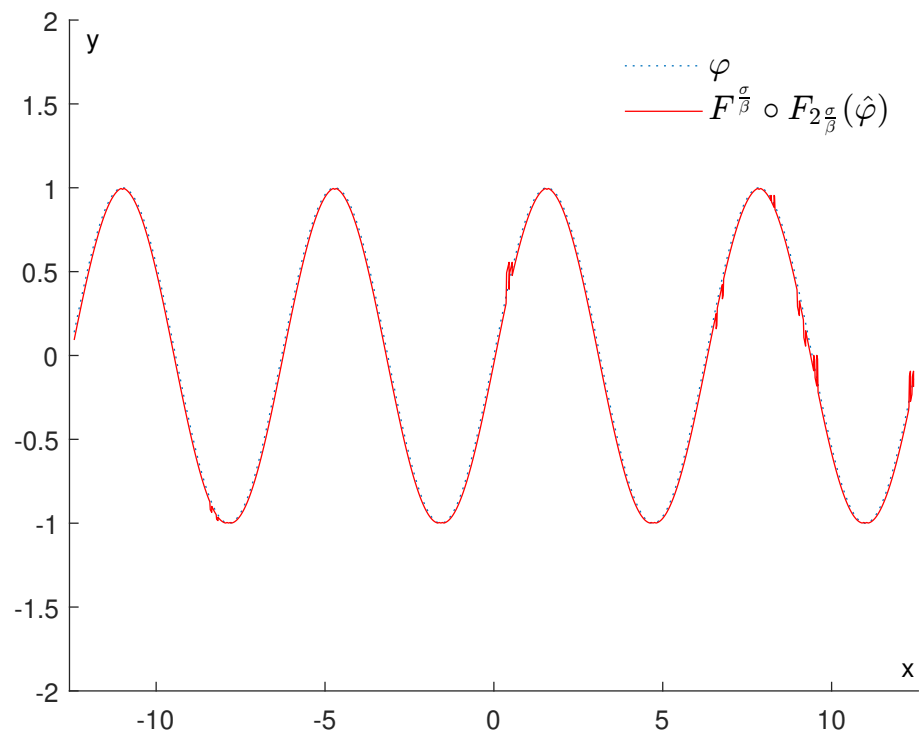


Figure 11. Example 1: Denoising via $F^{\frac{\sigma}{\beta}} \circ F_{2^{\frac{\sigma}{\beta}}}(\hat{\varphi})$.

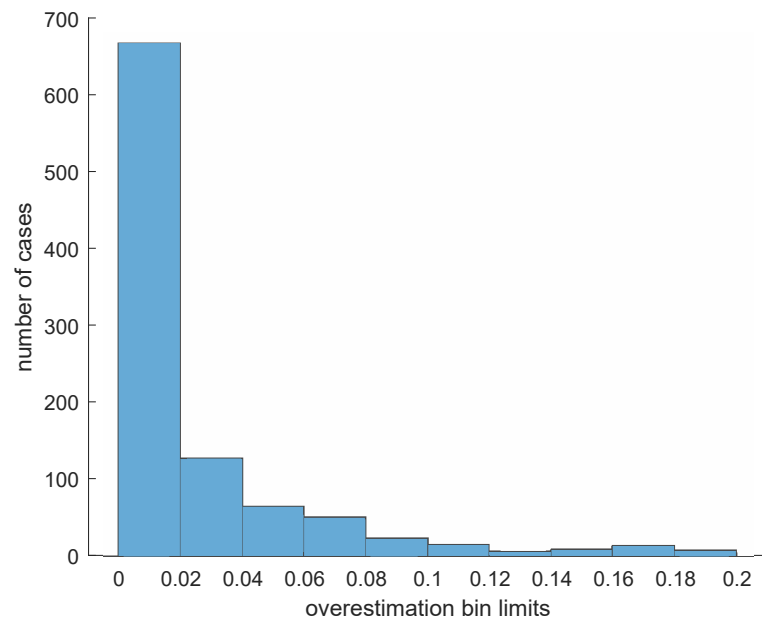


Figure 12. Example 1: Histogram counting the number of cases in each of ten bins concerning the overestimation value $3L^{\frac{\sigma}{\beta}} - \|F^{\frac{\sigma}{\beta}} \circ F_{2^{\frac{\sigma}{\beta}}}(\hat{\varphi}) - \varphi\|_{\infty}$ obtained in our simulations. Most of the cases belong to the first bin, specifically 665 cases, containing the functions φ for which $0 \leq 3L^{\frac{\sigma}{\beta}} - \|F^{\frac{\sigma}{\beta}} \circ F_{2^{\frac{\sigma}{\beta}}}(\hat{\varphi}) - \varphi\|_{\infty} \leq 0.02$.

This suggests that this upper bound is quite accurate.

6.1.2. Second Example

Let us consider the function

$$\varphi(x) := \begin{cases} \frac{1}{1000}(x-5)(x-3)(x+1)(x+4)(x+5) & \text{if } -5 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

for $x \in \mathbb{R}$. The coefficient $\frac{1}{1000}$ was chosen so that the Lipschitz constant L would be comparable to the one in the previous example. In this example, $L = \frac{27}{25}$. Figure 13 illustrates how the function $\hat{\varphi}$ looks compared to φ .

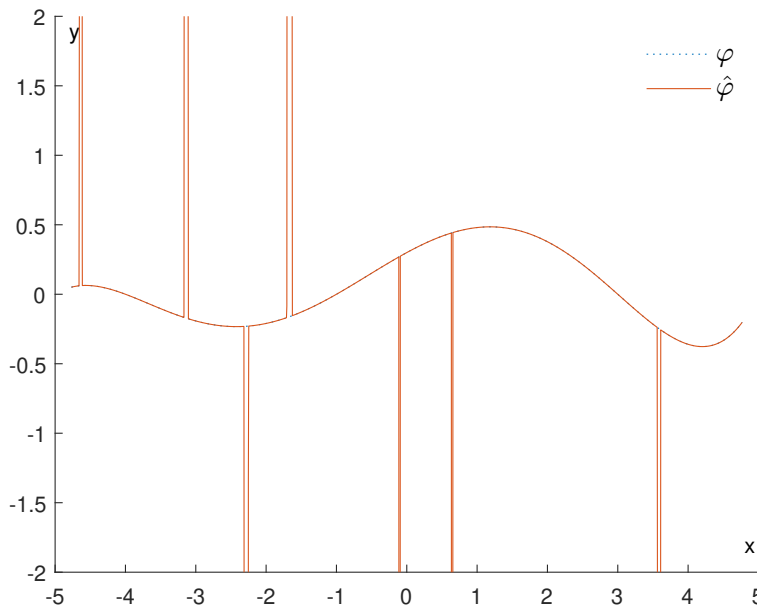


Figure 13. Example 2: Comparison between φ and $\hat{\varphi}$.

We will start by considering how well the convolution $\hat{\varphi} * T_n$ can approximate the original function φ , when n goes from 3 to 100. From Figure 14, it is immediately apparent that the max-norm distance between $\hat{\varphi} * T_n$ and φ remains quite large.

If we apply a convolution with $T_{\frac{1}{n}}$, for $3 \leq n \leq 100$, we obtain the results displayed in Figure 15, showing that all information represented by the function φ is progressively destroyed.

In contrast, if we apply the operator $F^{\frac{\sigma}{2}} \circ F_{\tau_n}(\hat{\varphi})$, for $n = 3, 5, 20, 100$, we obtain the results displayed in Figure 16. As we can see, this operator is much more efficient in removing the bumps and restoring the function φ .

When we apply a convolution with the function T_h and check the corresponding errors via the sup-norm, we obtain the results displayed in Figure 17. As we have already seen, since $\lim_{h \rightarrow +\infty} \hat{\varphi} * T_h = \hat{\varphi}$ and $\lim_{h \rightarrow +\infty} \hat{\varphi} * T_{\frac{1}{h}} = 0$, we obtain that the errors tend to $\|\hat{\varphi} - \varphi\|_{\infty} = \max_{i=1, \dots, k} |a_i| = \bar{a}$ and $\|\varphi\|_{\infty}$, respectively.

In contrast, if we apply the operator $F^{\frac{\sigma}{2}} \circ F_{\tau_n}$, we obtain the results displayed in Figure 18. This shows that the upper bound for the error stated in Corollary 1 is quite tight. As we can expect, the best denoising is achieved when we replace τ_n with $2\frac{\sigma}{\beta}$ (see Figure 19).

Finally, we again executed 1000 simulations. We used the same methodology as in the previous case, but this time we considered the polynomial presented at the beginning of this example. As we can see in Figure 20, the overestimation committed by our upper bound is often quite close to zero relative to the Lipschitz constant L of the polynomial φ .

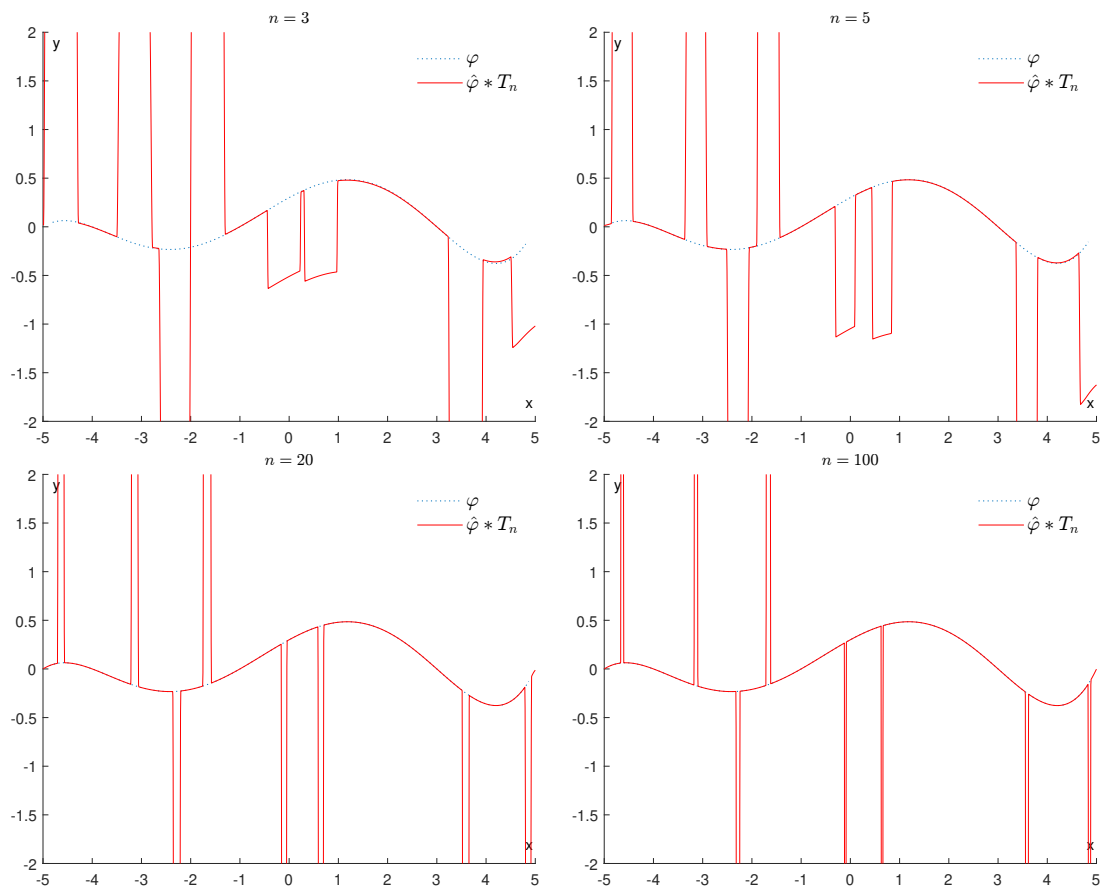


Figure 14. Example 2: Denoising via convolution with T_h ($h = n$).

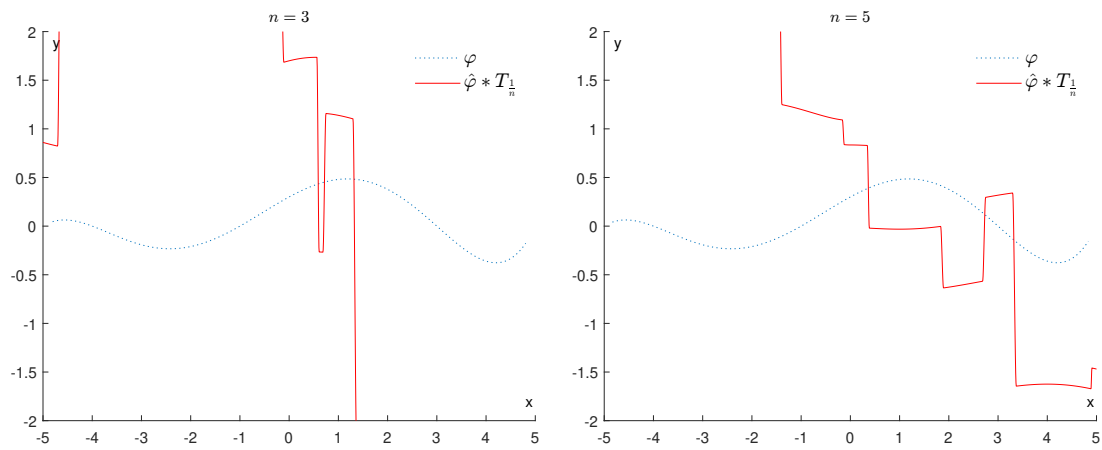


Figure 15. Cont.

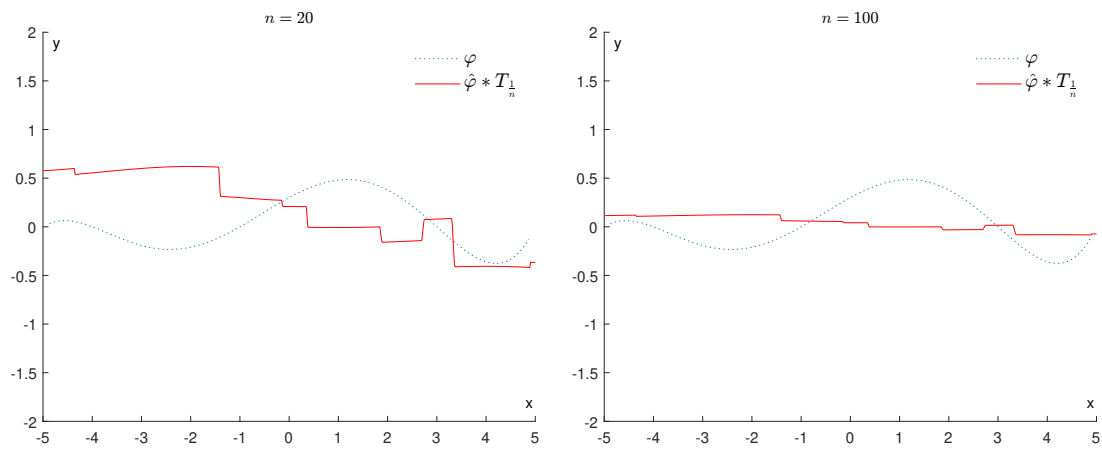


Figure 15. Example 2: Denoising via convolution with T_h ($h = \frac{1}{n}$).

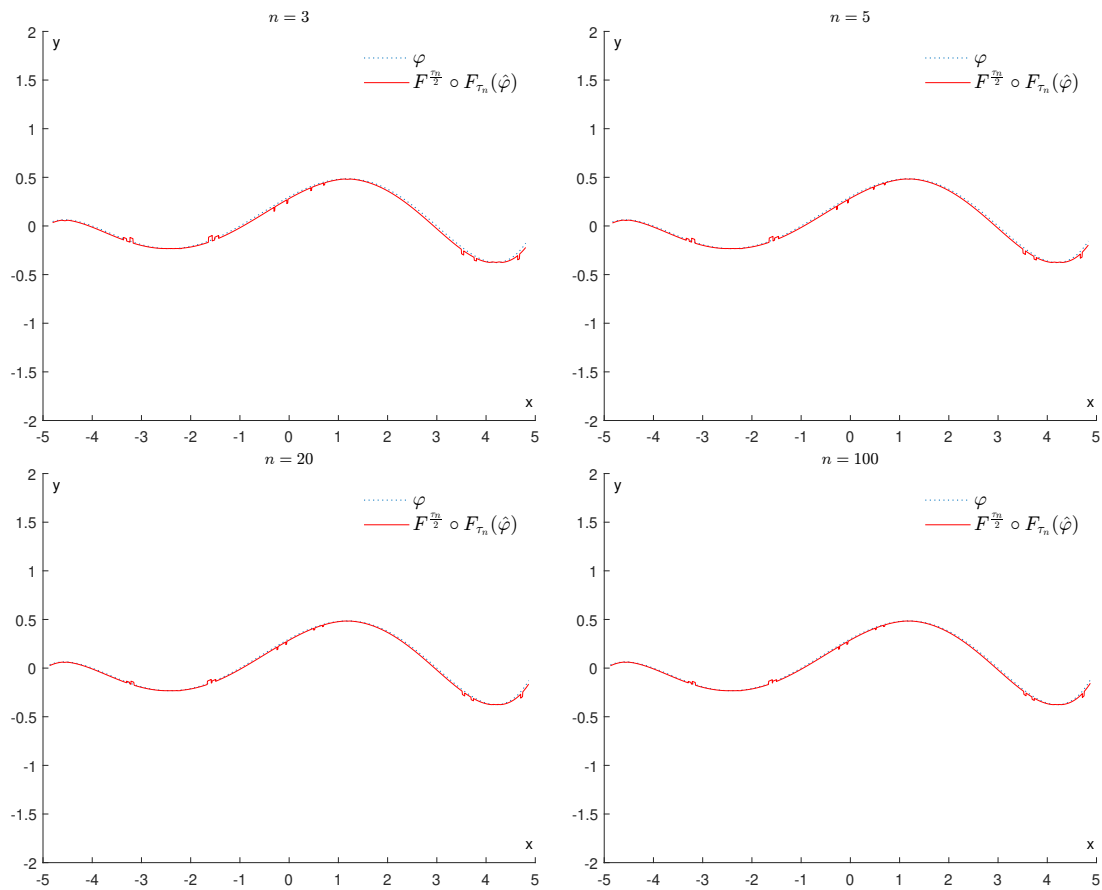


Figure 16. Example 2: Denoising via the proposed GENE $F^{\frac{m}{2}} \circ F_{\tau_n}$.

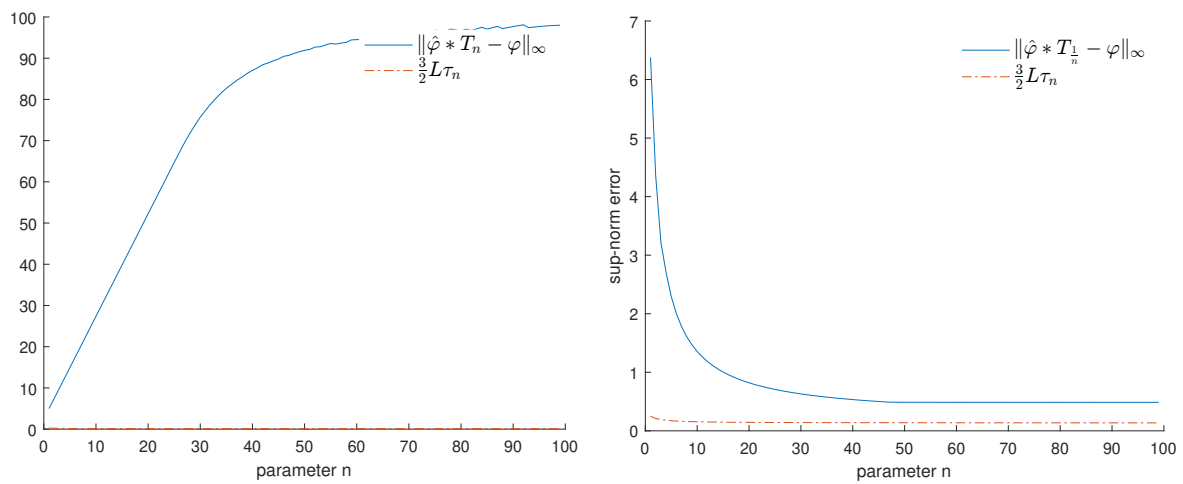


Figure 17. Example 2: Error made by applying a convolution with T_n (left) or $T_{\frac{1}{n}}$ (right).

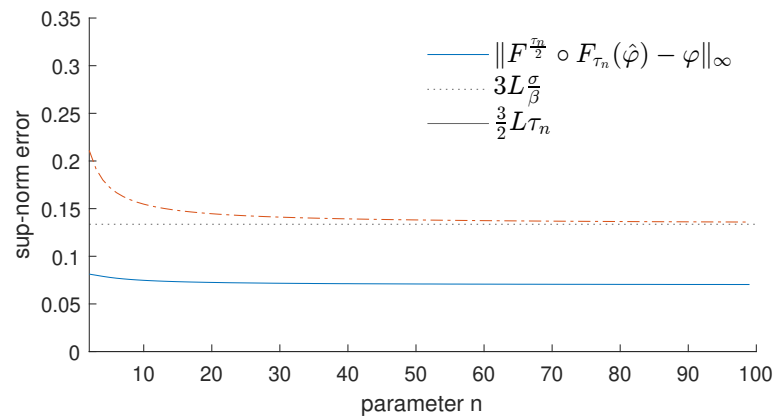


Figure 18. Example 2: Error made by using the GENE $F^{\frac{\alpha}{2}} \circ F_{\tau_n}$.

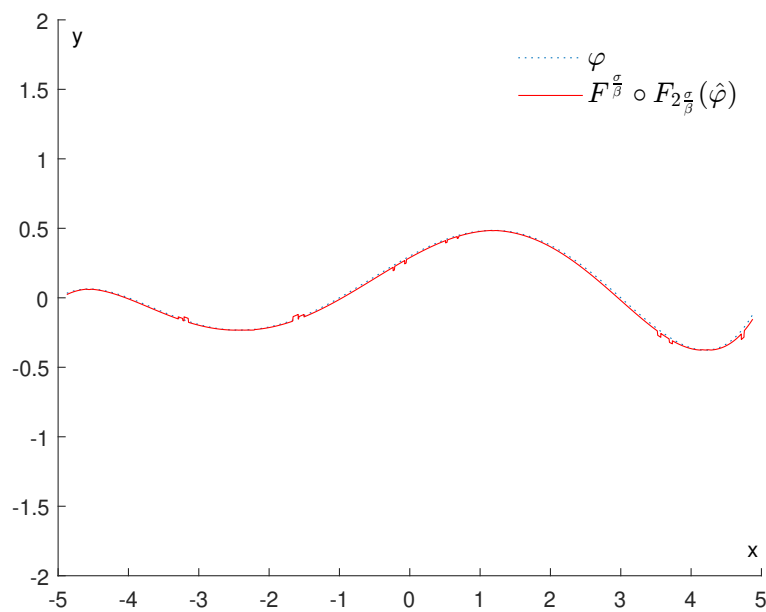


Figure 19. Example 2: Denoising via $F^{\frac{\alpha}{2}} \circ F_{2\frac{\alpha}{2}}$.

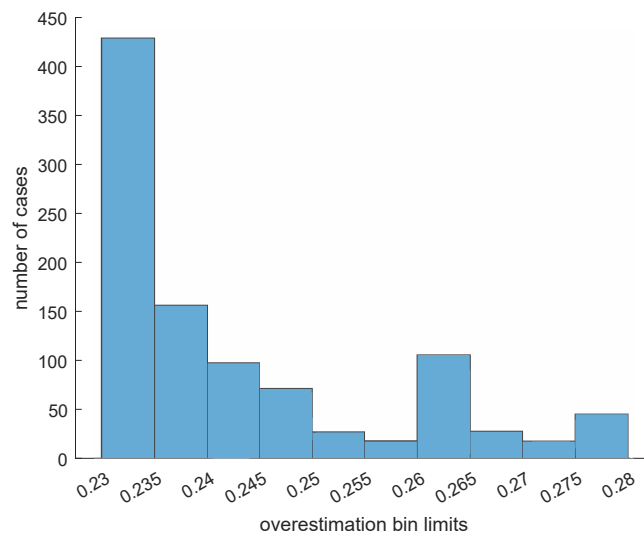


Figure 20. Example 2: Histogram counting the number of cases in each of ten bins concerning the overestimation value $3L\frac{\sigma}{\beta} - \|F_{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\|_{\infty}$ obtained in our simulations: 431 cases belong to the first bin and 154 belong to the second one; hence, most of the cases belong to the first two bins containing the functions φ for which $0 \leq 3L\frac{\sigma}{\beta} - \|F_{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\|_{\infty} \leq 0.24$.

6.2. Experiments

In order to check how good the upper bound stated in Theorem 3 is, we have made the following experiment. In the first step, we fixed $\ell = 20$ and $\sigma = 1.1$, and assumed the following parameters to be given: $L > 0, N \in \mathbb{N}, \alpha > 0, \beta > 0, k \in \mathbb{N}, a_1, \dots, a_k \in]0, \alpha[, b_1, \dots, b_k > \beta, c_1, \dots, c_k \in]0, \ell[$.

Firstly, we used the parameters $L > 0, N \in \mathbb{N}$ to generate a random L -Lipschitz function φ in the following way. We randomly chose and sorted in ascending order N points x_1, \dots, x_N in the open interval $]0, \ell[$, with uniform distribution. Hence, we obtained the following decomposition: $]0, \ell[=]0, x_1] \cup]x_1, x_2] \cup \dots \cup]x_N, \ell[$. We defined our Lipschitz function φ to be 0 outside $]0, \ell[$. After setting $(x_0, y_0) = (0, 0)$ and $(x_{N+1}, y_{N+1}) = (\ell, 0)$ for $i \in \{1, \dots, N + 1\}$, the value y_i of the function at x_i was randomly chosen, with uniform distribution in an interval that allows for an L -Lipschitz extension to $[0, \ell]$ of the function, i.e.,

$$[\max\{y_{i-1} - L(x_i - x_{i-1}), -L(\ell - x_i)\}, \min\{y_{i-1} + L(x_i - x_{i-1}), L(\ell - x_i)\}].$$

Finally, the graph of the Lipschitz function φ on $[0, \ell]$ was obtained by connecting each point (x_{i-1}, y_{i-1}) to (x_i, y_i) with a segment for $i \in \{1, \dots, N + 1\}$. We observe that φ constructed this way is an L -Lipschitz function.

Secondly, we used the parameters $\alpha > 0, \beta > 0, k \in \mathbb{N}, a_1, \dots, a_k \in]0, \alpha[, b_1, \dots, b_k > \beta, c_1, \dots, c_k \in]0, \ell[$ to generate a noise function as follows. We considered the mother function ψ defined by setting $\psi(x) := e^{1 - \frac{1}{1-x^2}}$ for $x \in]-1, 1[$ and $\psi(x) := 0$ for $x \notin]-1, 1[$. For each L -Lipschitz function φ produced in the previously described way, we considered the function $\hat{\varphi} = \varphi + \sum_{i=1}^k a_i \psi(b_i(x - c_i))$.

In Figures 21 and 22, some examples of the functions we have produced are displayed for $N = 3$ and $N = 7$, respectively. In each figure, the functions are displayed without noise (left) and with added noise (right).

In the second step, we fixed $\ell = 20$ and $\sigma = 1.1$ again, and considered a probabilistic model where α, β, L are given and the values N, k, a_i, b_i, c_i are random variables. In this setting, we compared the noise $\|\hat{\varphi} - \varphi\|_{\infty}$ with the probabilistic upper bound stated in Theorem 3 and the value $\left\| F_{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi \right\|_{\infty}$, which represents the reduced noise that we can obtain by applying our method. In order to average our results, for each triplet (α, β, L)

with $\alpha \in \{50, 55, 60, \dots, 100\}$, $\beta \in \{3, 4, 5, \dots, 13\}$, and $L \in \{1, 2, \dots, 10\}$, we randomly generated 100 examples of an L -Lipschitz function φ and its noisy version $\hat{\varphi}$ by randomly choosing the parameters N, k, a_i, b_i, c_i according to the following distributions:

- $N \sim \text{Unif}_{\{1, \dots, 10\}}$
- $k \sim \text{Unif}_{\{1, \dots, 10\}}$
- $a_i \sim \text{Unif}_{]0, \alpha[}$ for $i = 1, \dots, k$
- $b_i \sim \text{Unif}_{] \beta, 20[}$ for $i = 1, \dots, k$
- $c_i \sim \text{Unif}_{] 3 \frac{\sigma}{\beta}, \ell - 3 \frac{\sigma}{\beta} [}$ for $i = 1, \dots, k$.

Then, for the chosen values of each one of the three variables α, β, L , we computed the mean, with respect to the other two variables, of the average of the noise $\|\hat{\varphi} - \varphi\|_\infty$ and the average of the reduced noise $\|F_{\frac{\sigma}{\beta}} \circ F_{2 \frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\|_\infty$ obtained by our method, both evaluated for $(\varphi, \hat{\varphi})$ varying in the set of cardinality 100 that we produced. We also computed the mean of the probabilistic upper bound stated in Theorem 3 with respect to the same two variables. The results are displayed in Figures 23–25. We remind the reader that α and β , respectively, express the maximum height and the thinness of the noise bumps, while L is a Lipschitz constant for each function φ that we are interested in. These results illustrate the effectiveness of using GENEOS in the reduction of impulsive noise.

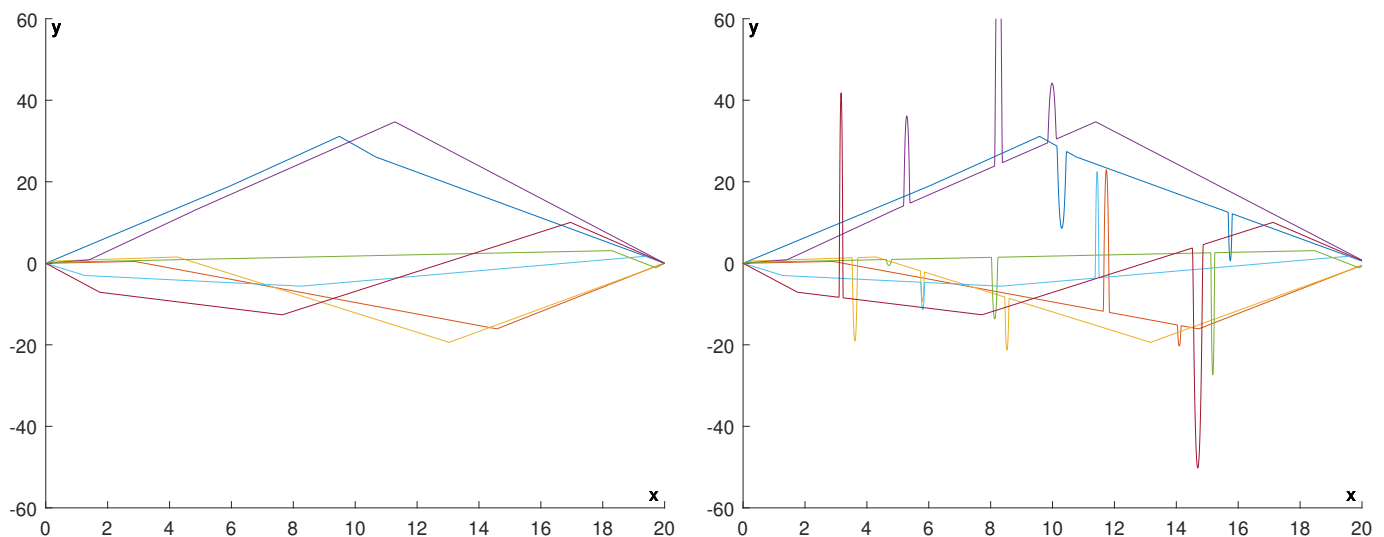


Figure 21. Five examples of the functions we produced for $N = 3$. In each figure, the functions are displayed without noise (left) and with added noise (right).

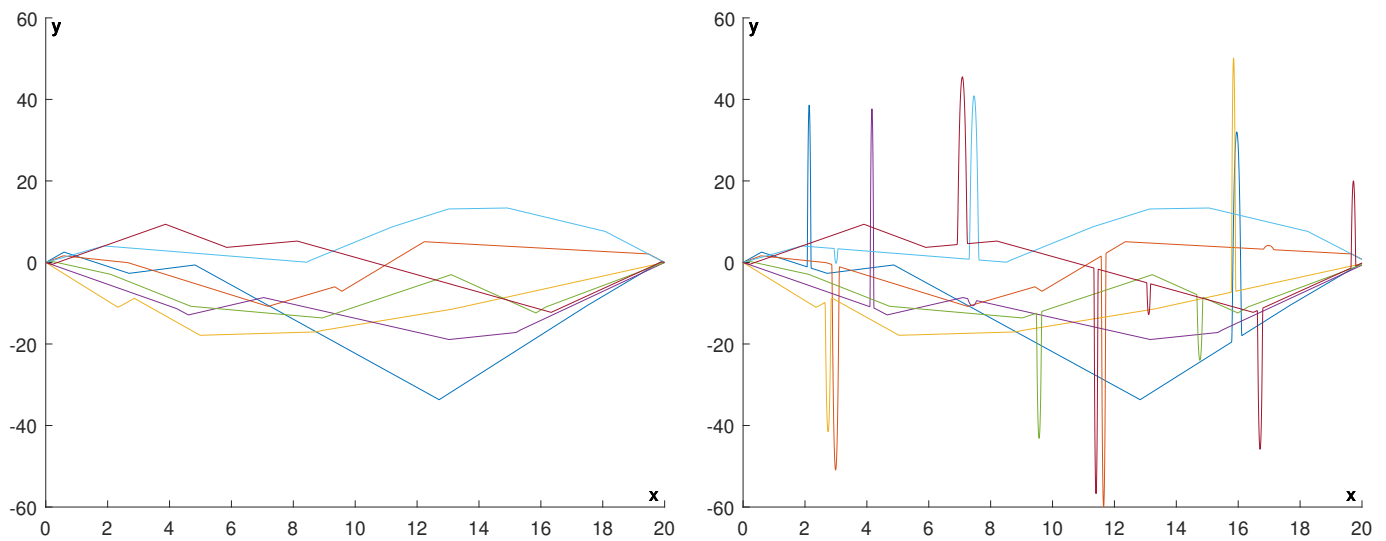


Figure 22. Five examples of the functions we produced for $N = 7$. In each figure, the functions are displayed without noise (left) and with added noise (right).

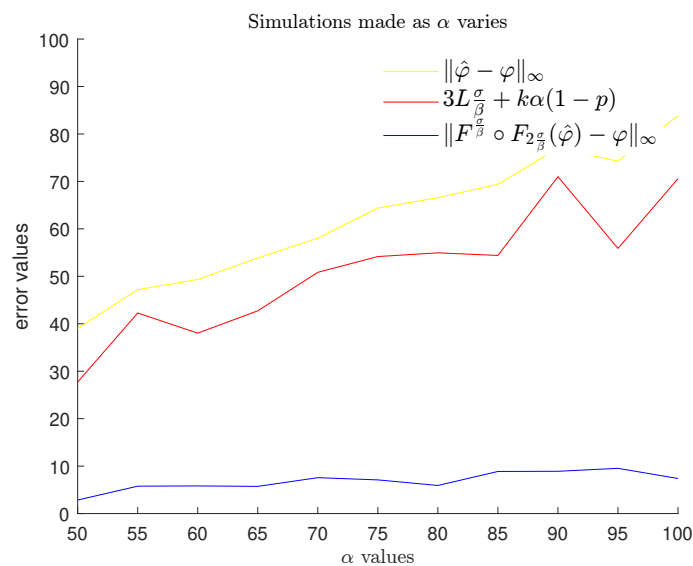


Figure 23. Plots of the means of the averaged values of $\|\hat{\varphi} - \varphi\|_\infty$ (yellow), the means of $3L\frac{\sigma}{\beta} + k\alpha\left(1 - \left(1 - 8\frac{(k-1)\sigma}{\ell\beta}\right)^k\right)$ (brown), and the means of the averaged values of $\left\|F_{\frac{\sigma}{\beta}} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\right\|_\infty$ (blue) for $\beta \in \{3, 4, 5, \dots, 13\}$ and $L \in \{1, 2, \dots, 10\}$, when α varies in the set $\{50, 55, 60, \dots, 100\}$.

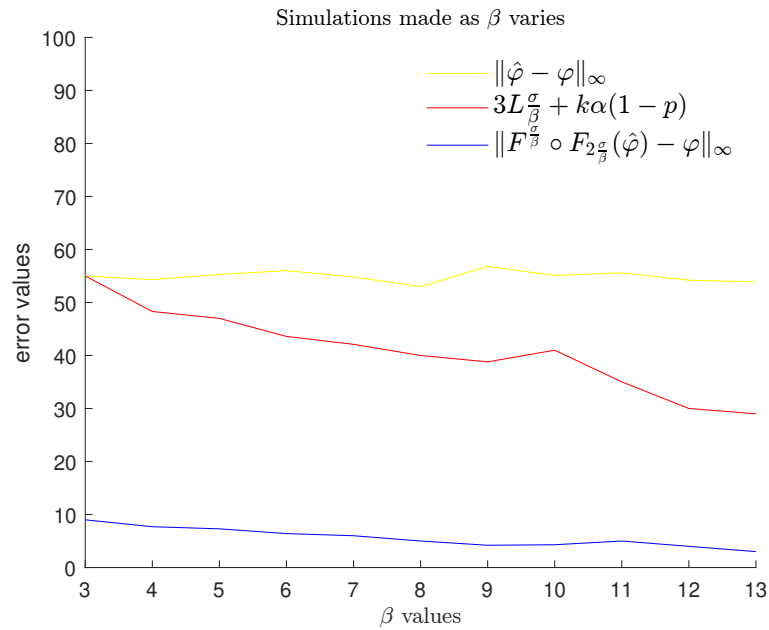


Figure 24. Plots of the means of the averaged values of $\|\hat{\varphi} - \varphi\|_\infty$ (yellow), the means of $3L\frac{\sigma}{\beta} + k\alpha(1 - p)$ (brown), and the means of the averaged values of $\|F\frac{\sigma}{\beta} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\|_\infty$ (blue) for $\alpha \in \{50, 55, 60, \dots, 100\}$ and $L \in \{1, 2, \dots, 10\}$, when β varies in the set $\{3, 4, 5, \dots, 13\}$.

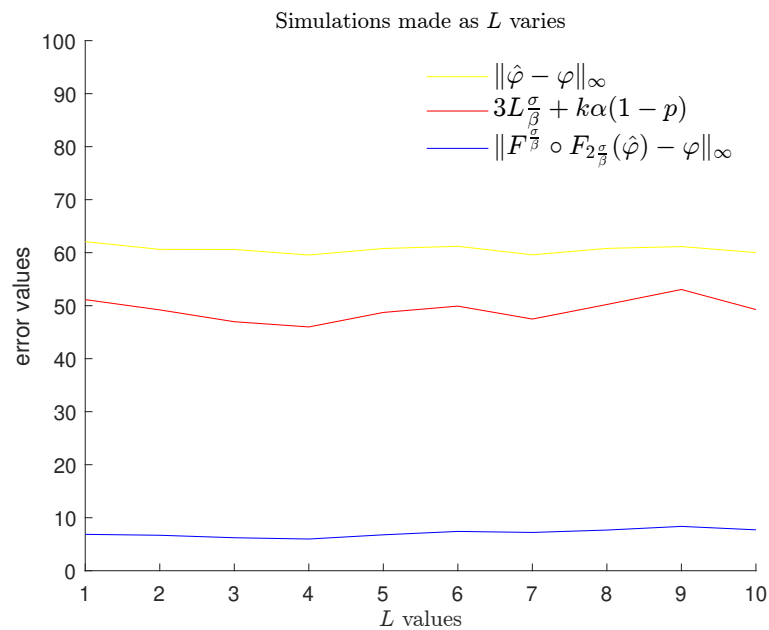


Figure 25. Plots of the means of the averaged values of $\|\hat{\varphi} - \varphi\|_\infty$ (yellow), the means of $3L\frac{\sigma}{\beta} + k\alpha(1 - p)$ (brown), and the means of the averaged values of $\|F\frac{\sigma}{\beta} \circ F_{2\frac{\sigma}{\beta}}(\hat{\varphi}) - \varphi\|_\infty$ (blue) for $\alpha \in \{50, 55, 60, \dots, 100\}$ and $\beta \in \{3, 4, 5, \dots, 13\}$, when L varies in the set $\{1, 2, \dots, 10\}$.

7. Conclusions

In our paper, we have proved a stability property for persistence diagrams of functions from \mathbb{R} to \mathbb{R} in the presence of impulsive noise. This property shows that TDA can also be useful when noise drastically changes the topology of the sublevel sets of the filtering functions we are considering, and highlights some new potential interactions between TDA and the theory of GENEOS. The experimental section shows that our approach is indeed able to remove impulsive noise in most cases. It would be interesting

to investigate the possibility of extending our method to real-valued functions defined on n -dimensional domains and allowing the function ψ to be a random variable, by selecting suitable GENEOS. We observe that the operators F^ε and F_ε introduced in Definition 3 can be easily adapted to the case of functions from \mathbb{R}^n to \mathbb{R} , by setting $F^\varepsilon(\varphi)(x) = \max\{\varphi(x+v) | v \in \mathbb{R}^n, \|v\| = \varepsilon\}$ and $F_\varepsilon(\varphi)(x) = \min\{\varphi(x+v) | v \in \mathbb{R}^n, \|v\| = \varepsilon\}$ for all $x \in \mathbb{R}^n$. Applying a suitable operator $F^\delta \circ F_\varepsilon$ to real-valued filtering functions defined on \mathbb{R}^n should allow us to control the expected value of the perturbation of persistence diagrams caused by uniformly distributed impulsive noise, even for $n > 1$ and in degrees greater than 0. Our long-term goal is to realize Machine Learning methods in which optimal operators for impulse noise removal can be learned and selected in the convex hull of a given finite set of GENEOS, using a theorem which guarantees that the convex combination of GENEOS is still a GENEOS (cf. [11]). We plan to devote our research to this topic in the future.

Author Contributions: P.F. devised the project. All authors contributed equally to the manuscript. The authors of this paper have been listed in alphabetical order. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the European Union in the context of the H2020 EU Marie Curie Initial Training Network project named WAKEUPCALL, and by INdAM-GNSAGA.

Data Availability Statement: Data available on request from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Edelsbrunner, H.; Morozov, D. Persistent homology: Theory and practice. In *European Congress of Mathematics*; European Mathematical Society: Zürich, Germany, 2013; pp. 31–50.
- Biasotti, S.; Floriani, L.D.; Falcidieno, B.; Frosini, P.; Giorgi, D.; Landi, C.; Papaleo, L.; Spagnuolo, M. Describing shapes by geometrical-topological properties of real functions. *ACM Comput. Surv.* **2008**, *40*, 12:1–12:87. [\[CrossRef\]](#)
- Carlsson, G. Topology and data. *Bull. Amer. Math. Soc.* **2009**, *46*, 255–308. [\[CrossRef\]](#)
- Edelsbrunner, H.; Harer, J. Persistent homology—A survey. In *Surveys on Discrete and Computational Geometry*; American Mathematical Society: Providence, RI, USA, 2008; Volume 453, pp. 257–282.
- Cohen-Steiner, D.; Edelsbrunner, H.; Harer, J. Stability of persistence diagrams. *Discrete Comput. Geom.* **2007**, *37*, 103–120. [\[CrossRef\]](#)
- Cohen-Steiner, D.; Edelsbrunner, H.; Harer, J.; Mileyko, Y. Lipschitz functions have L_p -stable persistence. *Found. Comput. Math.* **2010**, *10*, 127–139. [\[CrossRef\]](#)
- Fasy, B.T.; Lecci, F.; Rinaldo, A.; Wasserman, L.; Balakrishnan, S.; Singh, A. Confidence sets for persistence diagrams. *Ann. Stat.* **2014**, *42*, 2301–2339. [\[CrossRef\]](#)
- Buchet, M.; Chazal, F.; Dey, T.K.; Fan, F.; Oudot, S.Y.; Wang, Y. Topological analysis of scalar fields with outliers. In Proceedings of the 31st International Symposium on Computational Geometry (SoCG 2015), Eindhoven, The Netherlands, 22–25 June 2015; Leibniz International Proceedings in Informatics (LIPIcs); Arge, L., Pach, J., Eds.; Schloss Dagstuhl–Leibniz-Zentrum für Informatik: Dagstuhl, Germany, 2015; Volume 34; pp. 827–841.
- Adler, R.J.; Agami, S. Modelling persistence diagrams with planar point processes, and revealing topology with bagplots. *J. Appl. Comput. Topol.* **2019**, *3*, 139–183. [\[CrossRef\]](#)
- Vishwanath, S.; Fukumizu, K.; Kuriki, S.; Sriperumbudur, B.K. Robust persistence diagrams using reproducing kernels. In *Advances in Neural Information Processing Systems*; Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F., Lin, H., Eds.; Curran Associates, Inc.: New York, NY, USA, 2020; Volume 33, pp. 21900–21911.
- Bergomi, M.G.; Frosini, P.; Giorgi, D.; Quercioli, N. Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning. *Nat. Mach. Intell.* **2019**, *1*, 423–433. [\[CrossRef\]](#)
- Conti, F.; Frosini, P.; Quercioli, N. On the construction of Group Equivariant Non-Expansive Operators via permutants and symmetric functions. *Front. Artif. Intell.* **2022**, *5*, 786091. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bocchi, G.; Botteghi, S.; Brasini, M.; Frosini, P.; Quercioli, N. On the finite representation of linear group equivariant operators via permutant measures. *Ann. Math. Artif. Intell.* **2023**. [\[CrossRef\]](#)
- Frosini, P.; Jabłoński, G. Combining persistent homology and invariance groups for shape comparison. *Discrete Comput. Geom.* **2016**, *55*, 373–409. [\[CrossRef\]](#)
- Cerri, A.; Fabio, B.D.; Ferri, M.; Frosini, P.; Landi, C. Betti numbers in multidimensional persistent homology are stable functions. *Math. Methods Appl. Sci.* **2013**, *36*, 1543–1557. [\[CrossRef\]](#)
- Cerri, A.; Ethier, M.; Frosini, P. On the geometrical properties of the coherent matching distance in 2D persistent homology. *J. Appl. Comput. Topol.* **2019**, *3*, 381–422. [\[CrossRef\]](#)

17. Micheletti, A. A new paradigm for artificial intelligence based on group equivariant non-expansive operators. *Eur. Math. Soc. Mag.* **2023**, *128*, 4–12. [[CrossRef](#)]
18. Vaseghi, S.V. *Impulsive Noise: Modelling, Detection and Removal*; John Wiley and Sons, Ltd.: Chichester, UK, 2008; Chapter 13, pp. 341–358.
19. Earnest, M. Average Minimum Distance between n Points Generate i.i.d. with Uniform Dist. Mathematics Stack Exchange. Available online: <https://math.stackexchange.com/q/2001026> (accessed on 13 July 2021).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.