Hamburger Beiträge
zur Angewandten Mathematik

Interactions between Kernels, Frames, and Persistent Homology

Mijail Guillemard and Armin Iske

Nr. 2016-08
April 2016

Interactions between Kernels, Frames, and Persistent Homology

Mijail Guillemard and Armin Iske

Abstract This contribution discusses interactions between kernel methods, frame analysis, and persistent homology. To this end, we explain recent connections between these research areas, where special emphasis is placed on the discussion of reproducing kernel Hilbert spaces and persistent mechanisms. We show how interactions between these novel methodologies give new opportunities for the construction of numerical algorithms to analyze properties of data that are so far unexplored.

1 Introduction

In the last decades, the concept of kernel methods, along with their related notion of reproducing kernel Hilbert spaces (RKHS), has played an increasingly important role in a broad range of applications in data processing: Interpolation and approximation methods, signal sampling techniques, solution spaces of PDEs, characterization of integral operators, nonlinear dimensionality reduction methods, and machine learning, to mention but a few, are relevant applications, where kernel methods are of fundamental importance. More recent developments are relying on interactions between kernels and frame theory, with providing new opportunities for the construction of high performance numerical algorithms that are combining the advantages of kernels and frames. Quite recently, new tools for the efficient analysis of point cloud datasets arose from computational methods in differential geometry and algebraic topology, where persistent homology is one prominent example for such a new technique. In fact, persistent homology provides a basic algorithmic framework for computing homological information from large point cloud data.

Mijail Guillemard
Technische Universität Hamburg, Germany, e-mail: mijail.guillemard@tuhh.de

Armin Iske
Universität Hamburg, Germany, e-mail: armin.iske@uni-hamburg.de

In this paper, we show how interactions between novel methodologies give new opportunities for the construction of numerical algorithms to analyze properties of data that are so far unexplored. To this end, we first give a short introduction to reproducing kernel Hilbert spaces in Section 2, where we also discuss the fundamental result of Mercer's theorem, along with relevant applications in sampling theory and approximation. Basic concepts of frames and their relation to reproducing kernel Hilbert spaces are explained in Subsection 2.3. We introduce basic features of persistent homology in Section 3, including aspects concerning their stability properties. In Section 4, we finally discuss interactions between frames, kernels and persistent homology. This discussion includes a suitable description concerning the stability of frames and persistent homology and their relations. We finally describe novel concepts for kernels that are tailored to the space of persistent diagrams.

2 Reproducing Kernels and Approximation Theory

In this section, we give a short introduction to the basic concepts of kernels, reproducing kernel Hilbert spaces, and selected of their applications. We primarily focus on interpolation and approximation methods, and new interactions with frame theory. For the main ideas of these concepts we follow along the lines of [16, 17, 27].

2.1 Reproducing Kernel Hilbert Spaces (RKHS)

Despite the multiple and diverse contexts in which the concept of RKHS appears, the main principles can be cast in one unified framework. As a starting point, there are two important, closely related, concepts: a kernel, and an underlying RKHS. Even though there is a close relationship between them, we distinguish the differences of these points of view. One of the main motivations for defining a kernel is to analyze arbitrary unstructured sets by mapping its elements to a set with some useful structure: in the current case, the target will be a Hilbert space.

Definition 1. Given a nonempty set E, a *kernel* is a function $K : E \times E \to \mathbb{R}$, such that a Hilbert space H, and a map $\Phi : E \to \mathsf{H}$, exists with

$$K(x,y) = \langle \Phi(x), \Phi(y) \rangle_{\mathsf{H}}, \quad \forall x,y \in E.$$

The map Φ, and the space H, are denominated *feature map* and *feature space*.

Due to the concept of kernels, K is to measure and analyze the similarity between the elements of E (a set without any predefined structure) using the scalar product of the Hilbert space H. Note that if we use as the scalar field $\mathbb{C}$ instead of $\mathbb{R}$, we have to take care of defining $K(x,y) = \langle \Phi(y), \Phi(x) \rangle_{\mathsf{H}}$ due to the sesquilinearity of the scalar product in $\mathbb{C}$. There are no special constraints on the feature space H, but as we will see in Proposition 5, the interesting candidates are essentially equivalent, and the prototypical examples will be given by reproducing kernel Hilbert spaces.

Definition 2. The Hilbert space H_K of real functions defined on a nonempty set E is a *reproducing kernel Hilbert space* (RKHS), if there exist a map, the *reproducing kernel*, $K : E \times E \to \mathbb{R}$, satisfying:

1. For $K_x : E \to \mathbb{R},\ \ K_x(y) := K(x,y),\ \ y \in E$, we have $K_x \in \mathsf{H}_K$, $\forall x \in E$.
2. Reproduction property:

$$f(x) = \langle f, K_x \rangle, \quad \forall x \in E, \quad \forall f \in \mathsf{H}_K. \tag{1}$$

We have, as in the previous definition, an arbitrary nonempty set E as a starting point, but the focus now is on the particular type of Hilbert space H_K, and the set of functions $\{K_x\}_{x\in E}$ used to generate the reproduction property (the crucial characteristic for the applications of this framework). The work that follows is to analyze the relation of these definitions by constructing adequate feature maps Φ, and presenting specific examples of RKHS with a given kernel K.

Remark 1 (Symmetric and positive semi-definite properties [16]). The reproduction property allows to immediate obtain several basic aspects of reproducing kernels:

1. $K(x,x) \geq 0$, for any $x \in E$.
2. $\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(x_i,x_j) \geq 0$, for any $\{x_i\}_{i=1}^n \subset E$, $\{\lambda_i\}_{i=1}^n \subset \mathbb{C}$.
3. $K(x,y) = \overline{K(y,x)}$, for any $x,y \in E$.

Defining $K_y(x) := K(x,y)$, for $x \in \mathsf{H}_K$, and using (1), we obtain $K_y(x) = \langle K_y, K_x \rangle_{\mathsf{H}}$ for all $x \in E$. With letting $x = y$, we obtain

$$K(y,y) = \langle K_y, K_y \rangle_{\mathsf{H}} = \|K_y\|^2_{\mathsf{H}_K} \geq 0.$$

In a similar spirit, if we select n points $\{x_i\}_{i=1}^n \subset E$, and n complex numbers $\{\lambda_i\}_{i=1}^n \subset \mathbb{C}$, when using the relation $K(x_i,x_j) = K_{x_i}(x_j) = \langle K_{x_i}, K_{x_j} \rangle_{\mathsf{H}}$, we obtain the positive-semidefinite property:

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(x_i,x_j) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \langle K_{x_i}, K_{x_j} \rangle_{\mathsf{H}} = \left\langle \sum_{i=1}^n \lambda_i K_{x_i}, \sum_{j=1}^n \lambda_j K_{x_j} \right\rangle_{\mathsf{H}} \geq 0,$$

where symmetry follows immediately from

$$K(x,y) = K_x(y) = \langle K_x, K_y \rangle_{\mathsf{H}} = \overline{\langle K_y, K_x \rangle_{\mathsf{H}}} = \overline{K_y(x)} = \overline{K(y,x)}.$$

Remark 2 (RKHS Prototype I). Given a symmetric positive-definite kernel K, there is a prototypical example of RKHS that can be constructed by generating a vector space with the functions $K_x : E \to \mathbb{R}$, $K_x(y) = K(x,y)$:

$$\mathsf{H}_K := \overline{\operatorname{span}\{K_x : x \in X\}}.$$

The scalar product is given by $\langle K_x, K_y \rangle_{\mathsf{H}_K} := K(x,y)$, and the feature map is given by

$$\Phi_K : E \to \mathsf{H}_K, \quad \Phi_K(x) := K_x$$

We will see alternative ways of constructing prototypical RKHS with a main result based on the Mercer's theorem, Theorem 3.

One important characteristic of a Hilbert space, equivalent to the reproduction property is the continuity of the *point evaluation functionals* (or *Dirac functionals*), namely, that given $x \in E$, the map $f \to f(x)$ is continuous for all $f \in \mathsf{H}$. This fact is a straightforward consequence of the Riesz representation theorem. Recall that for any Hilbert space H, the map $L_g : \mathsf{H} \to \mathbb{R}$, $L_g(f) = \langle f, g \rangle, f \in \mathsf{H}$ is a linear and bounded (continuous) functional for any $g \in \mathsf{H}$. Conversely, the Riesz representation theorem, a fundamental property of Hilbert spaces [8], specifies that for any linear and bounded functional $L : \mathsf{H} \to \mathbb{R}$, there exist a unique vector $g \in \mathsf{H}$ satisfying $L(f) = \langle f, g \rangle$, for any $f \in \mathsf{H}$.

Theorem 1. *Let H be a Hilbert space of real functions defined in a nonempty set E, and let $L_x : \mathsf{H} \to \mathbb{R}$, $L_x(f) := f(x)$, be the point evaluation functional at $x \in E$. The linear map L_x is continuous for any $x \in E$, if and only if H has a reproduction property, $f(x) = \langle f, K_x \rangle_{\mathsf{H}}$, for a set $\{K_x\}_{x \in E} \subset \mathsf{H}$, and any $f \in \mathsf{H}$, $x \in E$.*

Proof. If the functional L_x is continuous, with the Riesz representation theorem, we have a vector $K_x \in \mathsf{H}$ with the reproduction property $L_x(f) = f(x) = \langle f, K_x \rangle_{\mathsf{H}}$, for any $x \in E$, and $f \in \mathsf{H}$. Conversely, with the reproduction property, we can construct a bounded linear functional $L_x(f) = \langle f, K_x \rangle_{\mathsf{H}}$ due to the continuity of the scalar product: that is, the point evaluation functionals are continuous. □

Remark 3 (Pointwise, uniform, strong and weak convergence). Another specific property of reproducing kernel Hilbert spaces is the fact that strong convergence implies pointwise convergence. Recall that for any nonempty set A, and any metric space (M, d), a sequence of mappings $f_n : A \to M$ converges *pointwise (or simply)* to $f : A \to M$, if $f_n(x)$ converges to $f(x)$, for any $x \in A$, i.e., $\lim_{n\to\infty} d(f_n(x), f(x)) = 0$. The convergence is called *uniform* if $\lim_{n\to\infty} \sup_{x\in E}(d(f_n(x), f(x))) = 0$. Uniform convergence obviously implies pointwise convergence, but the opposite does not hold in general.

We recall two other important notions of convergence: strong (or norm) convergence and weak convergence. In order to set these definitions we require a normed, (or Banach) space $\mathscr{B}$. The sequence x_n *converges weakly* to x, if $f(x_n)$ converges to $f(x)$ (i.e., $\lim_{n\to\infty} f(x_n) = f(x)$ in $\mathbb{C}$), for every bounded linear operator f in $\mathscr{B}$. If $\mathscr{B}$ is a Hilbert space, with the Riesz representation theorem we can write $\langle x_n, y \rangle \to \langle x, y \rangle$ (in $\mathbb{C}$), for any $y \in \mathscr{B}$. The sequence x_n *converges normwise (or strong)*, if $\|x_n - x\|_{\mathscr{B}}$ converges to zero. Due to the Cauchy Schwarz inequality, strong convergence implies weak convergence (see [16, p. 18]). But the converse is, in general, not true: take for instance any complete orthogonal system $\{\phi_n\}_{n\in\mathbb{N}}$ of a Hilbert space H. The sequence ϕ_n converges to zero weakly, since

$$\sum_{n\in\mathbb{N}} |\langle \phi_n, f \rangle_{\mathsf{H}}| < \infty, \qquad \forall f \in \mathsf{H},$$

but $\|\phi_n\|_{\mathsf{H}}$ does not necessarily converge to zero.

Moreover, note that strong convergence does not necessarily imply pointwise convergence. A standard example is the function space L^p, for $1 \leq p < \infty$, where we may have pointwise divergence on any set of measure zero, without affecting norm convergence. This is in contrast to the situation in a RKHS, where we can establish the following result, due to the reproduction property.

Proposition 1. *In a RKHS, strong convergence implies pointwise convergence.*

Proof. In a reproducing kernel Hilbert space H, each point evaluation functional is continuous, and so we have $|f(x)| \leq \|f\|_{\mathsf{H}} M_x$. Therefore,

$$|f_n(x) - f(x)| = |(f_n - f)(x)| \leq \|f_n - f\|_{\mathsf{H}} M_x.$$

As the metric in the Hilbert space H is given by $d(f,g) := \|f - g\|_{\mathsf{H}}$, we see that strong convergence implies pointwise convergence. $\square$

Let us now recall a few elementary properties concerning the uniqueness of reproducing kernels, characterization of linear subspaces of RKHS, and orthogonal projections on RKHS. All of the following basic results are straightforward applications of the reproduction property.

Proposition 2. *The reproducing kernel K in a RKHS H_K is unique.*

Proof. Let K' another reproducing kernel of H_K. We set $K_x(y) := K(x,y)$, and $K'_x(y) := K'(x,y)$. By the reproduction property in H_K we obtain the identity

$$\begin{aligned}
\|K_x - K'_x\|^2 &= \langle K_x - K'_x, K_x - K'_x \rangle \\
&= \langle K_x - K'_x, K_x \rangle - \langle K_x - K'_x, K_x \rangle \\
&= K_x - K'_x - K_x + K'_x = 0
\end{aligned}$$

which completes our proof. $\square$

Proposition 3. *Any linear subspace H of a RKHS H_K is also a RKHS.*

Proof. Let the point evaluation functionals be $L_x(f) = f(x)$, $f \in \mathsf{H}_K$, $x \in E$, and let $L_x|_{\mathsf{H}}$ be their restrictions to H. By Proposition 1, and the continuity of the $L_x|_{\mathsf{H}}$, the subspace H is a RKHS. $\square$

Proposition 4. *If a RKHS H_K is a linear subspace of a Hilbert space H, the orthogonal projection in H_k is given by*

$$P_{\mathsf{H}_K}(f)(x) = \langle f, K_x \rangle, \quad f \in \mathsf{H}.$$

Proof. For any $f \in \mathsf{H}$, we can write $f = f' + g$, with $f' \in \mathsf{H}_K$, and $g \in \mathsf{H}_K^{\perp}$. We have then $P_{\mathsf{H}_K}(f)(x) = \langle f', K_x \rangle + \langle g, K_x \rangle$. As $K_x \in \mathsf{H}_K$, we obtain $P_{\mathsf{H}_K}(f)(x) = \langle f, K_x \rangle$. $\square$

We now explore relations between the reproducing kernel K and a basis of the space H_K (see Theorem 2). This relation can conveniently be used for checking whether a given Hilbert space has a reproducing kernel (see Remark 7). We will see that in the case of a RKHS, the kernel can be expanded as a product of the basis elements, provided that the Hilbert space is topologically separable. Although many commonly used spaces are separable, it is important to recognize counterexamples.

Remark 4 (Separable spaces). A topological space (X,τ) is *separable* if there is a countable dense subset D, namely, $\overline{D}=X,\ D\subset X$. Recall that a Hilbert space H is separable if and only if it has a countable orthonormal basis. Indeed, let $\{\psi_i\}_{i\in I}\subset \mathsf{H}$ be a orthonormal set, i.e., $\langle\psi_i,\psi_j\rangle=\delta_{ij}$ for $i,j\in I$, with the usual Kronecker symbol δ_{ij}. For any $i,j\in I, i\neq j$, we have

$$\|\psi_i-\psi_j\|^2=\langle\psi_i-\psi_j,\psi_i-\psi_j\rangle=\langle\psi_i,\psi_i\rangle-\langle\psi_i,\psi_j\rangle-\langle\psi_j,\psi_i\rangle+\langle\psi_j,\psi_j\rangle=2.$$

If H is separable with a countable dense set D, and I is not countable, we have a contradiction using the density condition of D. The argument is to consider an injective map from I to D by selecting for every $\phi_i, i\in I$ an element in D. We have then a countable identification which contradicts the hypothesis of a non-countable I. Conversely, we can use the countable property of the base field ($\mathbb{C}$ or $\mathbb{R}$) of H in order to construct a countable dense set given a countable orthonormal basis.

Remark 5 (Counterexamples for separable spaces). As a topological concept, the separability of a Banach space depends on the underlying norm. For instance, a classical example of a nonseparable Banach space is the set of bounded operators in a Hilbert space, $B(\mathsf{H})$, with the norm topology, namely, the topology induced by the operator norm $\|T\|_{\mathrm{op}}:=\sup_{\|x\|_{\mathsf{H}}\le 1}\|T(x)\|_{\mathsf{H}}$, $T\in B(\mathsf{H})$. Although this topology is standard for the vector space $B(\mathsf{H})$, it turns out to be too fine to allow the construction of countable dense sets. Another standard example of a non-separable topological set can be described when considering sequences of complex numbers $(c_n)_{n=1}^{\infty}$, with $c_n\in\mathbb{C}$ for all $n\in\mathbb{N}$. To this end, we recall the linear sequence spaces

$$\ell^p(\mathbb{C}):=\left\{c=(c_n)_{n=1}^{\infty}\ \middle|\ \|c\|_p:=\left(\sum_{n\in\mathbb{Z}}|c_n|^p\right)^{1/p}<\infty\right\}\qquad\text{for } 1\le p<\infty,$$

and the linear space of bounded sequences,

$$\ell^\infty(\mathbb{C}):=\left\{c=(c_n)_{n=1}^{\infty},(c_n)\ \middle|\ \|c\|_\infty:=\sup_{n\in\mathbb{N}}c_n<\infty\right\}.$$

Any bounded sequence $c_0:=(c_n)_{n=1}^{\infty}$ converging to zero, $\lim_{n\to\infty}c_n=0$ has finite norm $\|c\|_\infty<\infty$, and so c_0 is an element of ℓ^∞, Among these examples, the space of bounded sequences, ℓ^∞, is the only case of a non-separable space.

For a separable RKHS H_K, we can give an important characterization of H_K.

Theorem 2. *Let H_K be a reproducing kernel Hilbert space. Then, we have*

$$K(x,y) = \sum_{i=1}^{\infty} \psi_i(x)\overline{\psi_i(y)}, \qquad \forall x, y \in E, \tag{2}$$

with a countable orthonormal system $\{\psi_i\}_{i\in\mathbb{N}}$ of H_K, if and only if H_K is separable.

Proof. If the decomposition (2) holds, we define $K_x : E \to \mathbb{R}$ by $K_x := \sum_{i=1}^{\infty} \psi_i(x)\psi_j$, for $x \in E$. Then, we have $K_x(y) = K(x,y)$, and with the reproduction property (1) we obtain:

$$f(x) = \langle f, K_x\rangle = \sum_{i=1}^{\infty} \langle f, \psi_i(x)\psi_i\rangle = \sum_{i=1}^{\infty} \langle f, \psi_i(x)\rangle \psi_i,$$

for any $f \in \mathsf{H}_K$ and $x \in E$. Therefore, the orthonormal system $\{\psi_i\}_{i\in\mathbb{N}}$ is a countable orthonormal basis, i.e., H_K is separable. Conversely, let H_K be a separable RKHS, with K its reproducing kernel. We then define $K_y(x) := K(x,y)$, with $K_y \in \mathsf{H}_K$, for $y \in E$. Since H_K is separable, there exists a countable orthonormal basis $\{\psi_i\}_{i\in\mathbb{N}}$ with $K_y = \sum_{i=1}^{\infty} c_i(y)\psi_i$. The coefficients $c_i(y)$ can be computed as scalar products $c_i(y) = \langle K_y, \psi_i\rangle = \overline{\psi_i(y)}$, and so we obtain the decomposition (2). □

Remark 6 (Kernel decomposition is basis independent). Note that the only requirement for the set $\{\psi_i\}_{i\in\mathbb{N}}$ in the proof of Theorem 2 is to be a countable orthonormal system. For any other countable orthonormal system $\{\phi_i\}_{i\in\mathbb{N}}$, we obtain, under the same hypothesis, the decomposition $K(x,y) = \sum_{i=1}^{\infty} \phi_i(x)\overline{\phi_i(y)}$.

Remark 7 (Examples and counterexamples for RKHS [28]). We are now in a position, where we can present elementary examples and counterexamples for RKHS.

The sequence space ℓ^2 is a RKHS. In fact, using $|c_n| \le 1$, for all $c \in \ell^2$, $\|c\|_{\ell^2} = 1$, we see that the point evaluation functional $L_n(c) := c_n$, for $c := (c_n)_{n=1}^{\infty} \in \ell^2$, and $n \in \mathbb{N}$, is continuous with operator norm $\|L_n\|_{op} = 1$. In this case, the kernel is $K(n,m) = \delta_{nm}$.

The linear space of square integrable functions $L^2([-\pi,\pi])$ is not a RKHS. In fact, for $L^2([-\pi,\pi])$ there is countable orthonormal system $\{\psi_l\}_{l\in\mathbb{N}}$, e.g., $\psi_l(t) = \sin(tl)/\sqrt{\pi}$. Therefore, $L^2([-\pi,\pi])$ is separable. Now we see that the expression

$$K(t,s) := \frac{1}{\pi}\sum_{\ell=1}^{\infty} \sin(t\ell)\sin(s\ell),$$

is not necessarily convergent (take for instance $t = s = \pi/2$), in which case the reproduction property does not hold. Alternatively, we can analyze the point evaluation functionals, and check that $L_x(f) := f(x)$, for $f \in L^2([-\pi,\pi])$ and $x \in [-\pi,\pi]$, is not continuous (for a fixed x, the expression $|f(x)|$ is for $\|f\|_2 = 1$ unbounded).

In the following subsection we discuss more elaborate examples of RKHS.

Mercer's theorem, feature maps and feature spaces

We now address an important result in RKHS theory, Mercer's theorem, which allows us to construct concrete examples of feature maps and feature spaces, thereby providing an important link between the concepts of kernels and RKHS (see Definitions 1 and 2). Up to now we have worked in a very general setting (a nonempty set E without any particular structure), which is of importance in relevant applications of kernel methods. For the sake of convenience, when setting the framework of the Mercer's theorem, we assume a measurable space (E,μ), with $E \subset \mathbb{R}^n$ and μ a Borel measure. The following step is to construct our main tool, which is an integral operator L_K defined on the space of square integrable functions

$$L^2_\mu(E) := \{f : E \to \mathbb{C}, \int_E |f(x)|^2 \, d\mu(x) < \infty\}.$$

Now we transfer the information from the rarefied environment given by (E,μ) and K, into the richer structural setting of the linear space $L^2_\mu(E)$ and the linear operator L_K. Once we ensure that the spectral theorem machinery can be applied to the operator L_K, we obtain a useful decomposition for K, which allows to construct a prototypical example of RKHS H_K.

Theorem 3 (Mercer's theorem [17]). *Let K be a continuous, symmetric, positive-semidefinite kernel defined in a measurable space (E,μ), with $E \subset \mathbb{R}^n$ closed, and let μ be a Borel measure. Further assume*

$$\iint_{E^2} K(x,y)^2 \, d\mu(x) \, d\mu(y) < \infty.$$

and, moreover, let the integral operator $L_K : L^2(E) \to L^2(E)$ be defined as

$$L_K f(x) := \int_E K(x,y) f(y) \, d\mu(y).$$

Then, we have the decomposition

$$K(x,y) = \sum_{i=1}^{\infty} \lambda_i \psi_i(x) \psi_i(y), \tag{3}$$

where $\{\lambda_i\}_{i=1}^\infty$ and $\{\psi_i\}_{i=1}^\infty$ are the eigenvalues and eigenvectors of the operator L_K.

Proof (sketch). First verify that L_K is a positive, self-adjoint compact operator. Then we can construct an orthonormal basis $\{\psi_i\}_{i=1}^\infty$ of $L^2(E)$, consisting of eigenvectors, with corresponding positive eigenvalues λ_i. This basis is then used for building the kernel expansion (3). To compute the adjoint of L_K, we use $\langle L_H f, g\rangle = \langle f, L_H^* g\rangle$ with

$$\langle L_H f, g\rangle = \int_E \int_E K(x,y) f(y) \overline{g(x)} \, d\mu(x) \, d\mu(y) = \int_E f(z) \overline{L_H^* g(z)} d\mu(z),$$

where we let

$$L_H^* g(z) := \int_E \overline{K(x,z)} g(x)\, d\mu(x) = \int_E K^*(z,x) g(x)\, d\mu(x) \quad \text{with } K^*(x,z) := \overline{K(z,x)}.$$

Using the symmetric property of K, we obtain $L_K = L_K^*$ (see [15, p. 91]). With the spectral theorem we obtain the orthonormal basis $\{\psi_i\}_{i=1}^{\infty}$ and corresponding eigenvalues λ_i, used to construct the decomposition (3). □

Remark 8 (RKHS Prototype II [17]). Given a symmetric positive semi-definite kernel K, Mercer's theorem allows us to construct another prototypical example of a RKHS. By using decomposition (3), we can construct a feature map $\Phi_\mu : E \to l^2(\mathbb{C})$, with $\Phi_\mu(x) = (\sqrt{\lambda_i}\psi_i(x))_{i=1}^{\infty}$. If the number of nonzero eigenvalues is $N < \infty$, we use the vector space $\mathbb{R}^N$ instead of ℓ^2. As explained in [17], the eigenvectors $\{\psi_k\}$ and eigenvalues $\{\lambda_i\}$ depend on the measure μ, i.e., by selecting a different measure, we obtain a different feature map Φ.

We now show that the different prototypes of RKHS we used so far, and any other that can be constructed, are essentially equivalent. The main idea is to construct an isometry between the first RKHS prototype we described in Remark 2, and any arbitrary feature space H with feature map $\Phi : E \to \mathsf{H}$, and $\langle \Phi(x), \Phi(y) \rangle = K(x,y)$.

Proposition 5. *Let E be a nonempty set, $K : E \times E \to \mathbb{R}$ be a positive semi-definite kernel, and Φ be an arbitrary feature map with feature space H, that is, $\Phi : E \to \mathsf{H}$, with $\langle \Phi(x), \Phi(y) \rangle_{\mathsf{H}} = K(x,y)$. We define the vector spaces*

$$\mathsf{H}_\Phi := \overline{span\{\Phi_x, x \in E\}}, \quad \mathsf{H}_K := \overline{span\{K_x, x \in E\}},$$

with $\Phi_x := \Phi(x)$, and $K_x(y) := K(x,y)$, for any $x, y \in E$. Denoting by $\mathbb{R}^E$ the vector space of real-valued functions on E, we define the linear operator

$$L_\Phi : \mathsf{H} \to \mathbb{R}^E, \quad L_\Phi(v)(x) := \langle v, \Phi_x \rangle_{\mathsf{H}}.$$

Then, the restriction $L_\Phi|_{\mathsf{H}_\Phi}$ is an isometry (isometric isomorphism) from H_Φ to H_K.

Proof. For $L_\Phi(v) = 0$ we have $\langle v, \Phi_x \rangle = 0$, for all $x \in E$. Then, $\ker L_\Phi = \mathsf{H}_\Phi^\perp$, and therefore $L_\Phi|_{\mathsf{H}_\Phi}$ is bijective. Now note that $L_\Phi(\Phi_y)(x) = \langle \Phi_y, \Phi_x \rangle = K(x,y) = K_y(x)$ for any $x \in E$, therefore L_Φ maps the function Φ_y to K_y, which implies that $\mathrm{span}\{\Phi_x, x \in E\}$ is isomorphic to $\mathrm{span}\{K_x, x \in E\}$. The isometry property follows from $\langle \Phi_x, \Phi_y \rangle_{\mathsf{H}} = K(x,y) = \langle K_x, K_y \rangle_{\mathsf{H}_K}$. □

Remark 9 (Applications to sampling theory [28]). We apply the general framework of RKHS to sampling as follows. Given a RKHS H_K with kernel $K : E \times E \to \mathbb{R}$, the main component for constructing a sampling procedure is an adequate selection of points $\{t_k\}_{k \in \mathbb{N}} \subset E$, such that $\{K_{t_k}\}_{k \in \mathbb{N}}$ is a complete orthogonal system of H_K. In this context we have

$$\|K_{t_k}\|^2 = \langle K_{t_k}, K_{t_k} \rangle = K(t_k, t_k), \quad f(t_k) = \langle f, K_{t_k} \rangle.$$

Therefore, the sampling reconstruction formula is given as

$$f(x) = \sum_{k\in\mathbb{N}} \langle f, K_{t_k}\rangle \frac{K_{t_k}(x)}{\sqrt{K(t_k,t_k)}} = \sum_{k\in\mathbb{N}} f(t_k) \frac{K_{t_k}(x)}{\sqrt{K(t_k,t_k)}}, \quad \forall f \in \mathsf{H}_K,\ x \in E.$$

We can use this scheme to obtain the well known case of Nyquist-Shannon sampling framework when using B_ω, the space of square integrable functions whose Fourier transform is supported in the interval $[-\omega, \omega]$ (the space of bandlimited functions, or the Paley-Wiener space). This space turns out to be a RKHS with kernel

$$K(x,y) = \frac{\sin(\omega(x-y))}{\omega(x-y)}.$$

Using the function $\operatorname{sinc}(x) := \sin(x)/x$, and the sequence $\{t_k := k\Delta, k \in \mathbb{Z}\}$, for a sampling step Δ, we obtain the well-known sampling formula

$$f(x) = \frac{1}{2\omega} \sum_{k\in\mathbb{Z}} f(t_k) \operatorname{sinc}\left(\frac{\omega}{\pi}(x-t_k)\right), \qquad \forall f \in B_\omega,\ x \in \mathbb{R}.$$

2.2 Further Aspects in Approximation Theory

Approximation is concerned with the design and analysis of computational methods for function reconstruction. A standard problem consist in recovering a function $f : \Omega \to \mathbb{R}$, from a finite set of values $f(x_1), \ldots, f(x_n)$, with $X := \{x_1, \ldots x_n\}$, and $\Omega \subset \mathbb{R}^d$. To set up a suitable framework for measuring the error behavior, execution speed, and quality of the approximation procedures, the first step is to identify the structure of the spaces in which the target function f lies. To this end, it is common standard to first select a normed linear space, or a Banach space, and then to fix a suitable linear subspace for building efficient approximation methods. Basic definitions and features can conveniently be presented in the context of a metric space.

Definition 3. Let (M,d) be a metric space, and $U \subset M$. A *best approximation* of $f \in M$ in U is an element $u^\star \in U$ satisfying

$$d(f,u^\star) = d(f,U) \quad \text{for} \quad d(f,U) := \inf\{d(f,u), u \in M\}. \tag{4}$$

Basic questions that need to be addressed are the existence, uniqueness, and construction of best approximations. An elementary property for a subset U to fulfill in order to guarantee existence of a best approximation, is compactness. Further details on this can be found in standard texts on approximation theory (e.g. [5, 6, 20, 26]).

Proposition 6. *If $U \subset M$ is a compact set in the metric space (M,d), then for every $f \in M$, there exist a best approximation $u^\star \in U$ of f.*

Proof. For $d := \inf\{d(f,u), u \in M\}$, we take a minimal sequence $\{u_k\}_{k\in\mathbb{N}} \subset U$ satisfying $d(f,u_k) \to d$ for $n \to \infty$. Since U is compact, the minimal sequence

$\{u_k\}_{k\in\mathbb{N}}$ has a limit point $u^\star \in U$, giving a best approximation of f: To see that $u^\star$ satisfies (4), we use the triangle inequality $d(f,u^\star) \leq d(f,u_k)+d(u_k,u^\star)$. Since $d(u_k,u^\star) \to 0$ for $n \to \infty$, we have $d(f,u^\star) = d$. □

We now reagard normed linear spaces V (instead of metric spaces) to develop criteria for best approximations.

Proposition 7. *Let U be a finite dimensional subspace of a normed linear space V. For every $f \in V$ there exist a best approximation in U.*

Proof. For any $u_0 \in U$, the subset $U_0 := \{u \in U \,:\, \|f-u\| \leq \|f-u_0\|\} \subset V$ is compact. Due to Proposition 6 there exists a best approximation $u^\star \in U$ of f. □

Up to know we have only addressed the existence of best approximations. To guarantee uniqueness, we work with strictly convex norms.

Remark 10 (Strictly convex norm). A norm $\|\,\|$ in a vector space V is said to be *strictly convex*, iff $\|\alpha v+\beta w\| < \alpha\|v\|+\beta\|w\|$ for all $v,w \in V$, and all $\alpha+\beta = 1$, $\alpha,\beta \in (0,1)$. Due to the sublinearity, every norm is convex, but in the case of L^p spaces, the $\|\|_\infty$ and $\|\|_1$ are the only cases that are not strictly convex.

Proposition 8. *Let U be a linear subspace of a normed linear space V with strictly convex norm. Then every element $f \in V$ has at most one best approximation.*

Proof. Let $f \in V$, and assume that we have two different best approximations $u_1,u_2 \in U$ with $d := \|f-u_1\| = \|f-u_2\|$. We can use the strict convexity of the norm to compute

$$\left\|f-\frac{1}{2}(u_1+u_2)\right\| < \frac{1}{2}\|f-u_1\|+\frac{1}{2}\|f-u_2\| = d,$$

which contradicts our assumption on the optimality of u_1 and u_2, and so there can only exist at most one best approximation of f. □

We now consider best approximations in pre-Hilbert spaces H.

Proposition 9. *Let H be a pre-Hilbert space, and $U \subset \mathsf{H}$ a linear subspace. Then, an element $u^\star \in U$ is a best approximation of an element $f \in \mathsf{H}$, if and only if*

$$\langle f-u^\star, v\rangle = 0, \quad \forall v \in U. \tag{5}$$

Proof. Assuming that an element $u^\star \in U$ satisfies the orthogonality (5), we can apply the Pythagoras theorem to obtain

$$\|f-u\|^2 = \|(f-u^\star)+(u^\star-u)\| = \|f-u^\star\|+\|u^\star-u\| > \|f-u^\star\|$$

for any $u \in U$. In this case, $u^\star$ is a best approximation of f.

Assuming that the orthogonality (5) does not hold for some $v \in U$, we can select one $\lambda := -\langle f-u^\star, v\rangle/\|v\|^2$ to compute

$$\|f-u^\star+\lambda v\|^2 = \|f-u^\star\| + 2\lambda\langle f-u^\star, v\rangle + \lambda^2\|v\|^2 < \|f-u^\star\|^2,$$

in which case $u^\star$ cannot be a best approximation of f. □

RKHS and Approximation Theory

We now demonstrate how to construct approximation algorithms in a RKHS. To this end, we follow along the lines of [25]. For further details, see [2, 14, 23, 24, 29].

Remark 11 (RKHS approximation scheme). The basic setup for approximation in a RKHS is as follows. The first ingredient is an (unknown) function $f : \Omega \to \mathbb{R}$ to be reconstructed from a discrete set of (given) values $f(x_1), \ldots, f(x_n)$, based on the sampling elements $X = \{x_1, \ldots, x_n\} \subset \Omega \subset \mathbb{R}^d$. The second ingredient is a RKHS H_K with kernel $K : \Omega \times \Omega \to \mathbb{R}$, where we assume $f \in \mathsf{H}_K$. Finally, the third ingredient a finite dimensional approximation space, a linear subspace of H_K of the form

$$S_X := \mathrm{span}\{K_x,\ x \in X\}, \text{ with } K_x : \Omega \to \mathbb{R},\ K_x(y) := K(x,y).$$

Given these ingredients, we can directly apply the basic framework of kernel-based approximation to efficiently reconstruct f. From an application point of view, one important decision is the selection of a suitable kernel K, with corresponding RKHS H_K. We show two important properties of the resulting approximation scheme: uniqueness and optimality of the best approximation.

Theorem 4. *Let H_K be a RKHS with kernel $K : \Omega \times \Omega \to \mathbb{R}$, for $\Omega \subset \mathbb{R}^d$. Let $f \in \mathsf{H}_K$ be a function, whose sampling values $f(x_1), \ldots, f(x_n)$ on $X := \{x_1, \ldots, x_n\} \subset \Omega$ are given. Then, there exists a best approximation $f_X^* \in S_X := span\{K_x : x \in X\}$, whose coefficients are the unique solution of the linear system*

$$f_X^*(x_k) = \sum_{i=1}^{n} a_i^* K(x_i, x_k) = f(x_k), \quad 1 \le k \le n. \tag{6}$$

Proof. Since $S_X \subset \mathsf{H}_K$ is a finite-dimensional linear subspace, there exists a best approximation $f_X^* \in S_X$ of $f \in \mathsf{H}_K$ satisfying

$$\langle f - f_X^*, s\rangle = 0, \quad \forall s \in S_X,$$

due to Propositions 7 and 9. From the reproduction property (1), for the function $K_{x_k}(y) := K(x_k, y)$, we get

$$0 = \langle f - f_X^*, K_{x_k}\rangle = f(x_k) - f_X^*(x_k), \quad 1 \le k \le n.$$

Since f_X^* is an element of S_X, we obtain (6). □

Theorem 5. *Under the conditions und with the notations of Theorem 4, we find the optimality property*

$$\min_{g \in O_X(f)} \|g\|_K = \|f^*\|_K, \qquad O_X(f) := \{g \in S_X,\ g|_X = f|_X\}.$$

Proof. For any $g \in O_X(f)$, we have

$$\|g\|_K^2 = \|f_X^* + (g - f_X^*)\|_K^2 = \|f_X^*\|^2 + 2\langle f_X^*, g - f_X^*\rangle + \|g - f_X^*\|^2. \tag{7}$$

The inner product (7) can be analyzed by using $K_{x_i}(y) := K(x_i, y)$, $y \in X$, to obtain

$$< f_X^*, g - f_X^* >= \left\langle \sum_{i=1}^{n} a_i^* K_{x_i}, g - f_X^* \right\rangle = \sum_{i=1}^{n} a_i^* \langle K_{x_i}, g - f_X^*\rangle = \sum_{i=1}^{n} a_i^* (g(x_i) - f_X^*(x_i)).$$

Since $f_X^*(x_i) = f(x_i)$, and $g(x_i) = f(x_i)$, for $i = 1, \ldots, n$, we have $\langle f_X^*, g - f_X^*\rangle = 0$. Therefore, we obtain $\|g\| \geq \|f_X^*\|$, for all $g \in O_X(f)$ from (7). □

2.3 Interactions between Kernels and Frames

In this section, we address interactions between frame theory and reproducing kernel Hilbert spaces, whose connections are investigated in [19, 21]. The main goal is to gain additional flexibility (when using a frame instead of a basis in a Hilbert space) to enlarge the set of admissible kernels and RKHS. In frame theory one considers a family of vectors $\{\psi_x\}_{x \in \mathsf{X}}$ in a Hilbert space H, where X is a locally compact Hausdorff space with a positive Radon measure μ (see [11]). When X is finite or discrete (e.g., $\mathsf{X} = \mathbb{N}$), we consider using a counting measure μ, and the resulting concept will be a generalization of an orthogonal basis. This provides quite flexible tools for the analysis and synthesis of a signal $f \in \mathsf{H}$.

A frame $\{\psi_x\}_{x \in \mathsf{X}} \subset \mathsf{H}$ relies on the stabilization of the analysis operator.

Definition 4. A set of vectors $\{\psi_x\}_{x \in \mathsf{X}} \subset \mathsf{H}$ in a Hilbert space H is a *frame*, if

$$A||f||^2 \leq ||Vf||^2 \leq B||f||^2, \quad \forall f \in \mathsf{H}$$

for $0 < A \leq B < \infty$, the *lower and upper frame bounds*, where $V : \mathsf{H} \to L^2(\mathsf{X})$, $(Vf)(x) = \langle f, \psi_x\rangle$, for $x \in \mathsf{X}$, is the *analysis operator*.

Reducing the difference between A and B improves the stability of V, and for the case of $A = B$, or $A = B = 1$, the resulting frame is denominated *tight frame* and *Parseval frame*, respectively. The synthesis operator $V^* : L^2(\mathsf{X}) \to V$,

$$V^*((a_x)_{x \in \mathsf{X}}) = \int_{\mathsf{X}} a_x \psi_x \, d\mu(x),$$

is defined through a positive Radon measure μ, when X is a locally compact Hausdorff space (see [11]). The maps V^* and V are combined in the frame operator

$$S = V^*V : \mathsf{H} \to \mathsf{H}, Sf = \int_{\mathsf{X}} \langle f, \psi_x\rangle \psi_x d\mu(x),$$

which plays an important role, since the operator norm of S can be bounded by A and B,

$$A \leq ||S||_{op} \leq B. \tag{8}$$

Now we turn to the interaction with reproducing kernel Hilbert spaces.

Theorem 6. *Let* H *be a Hilbert space of functions over* $\Omega \subset \mathbb{R}^d$*, with a frame* $\{\phi_i\}_{i\in I} \subset \mathsf{H}$*. Let* $K_x(y) := \sum_{i\in I} \overline{\phi_i}(y)\phi_i(x)$*, where* $\{\overline{\phi_i}\}_{i\in I}$ *is the dual frame of* $\{\phi_i\}_{i\in I}$*. Then,* H *is a reproducing kernel Hilbert space, if*

$$||K_x||_{\mathsf{H}} < \infty, \quad \forall x \in \Omega.$$

Proof. We use the frame property to proof the continuity of point evaluation functionals (see Theorem 1). The frame property gives

$$f = \sum_{i\in I} \langle f, \overline{\phi_i} \rangle \phi_i, \quad \forall f \in \mathsf{H}, \tag{9}$$

where the inner product in (9) is defined on H.

Now we use the semi-norm $\|f\|_x = |f(x)|$, for $x \in \Omega$, on H. This allows us to rewrite (9), by pointwise convergence, as

$$f(x) = \sum_{i\in I} \langle f, \overline{\phi_i} \rangle \phi_i(x), \quad \forall x \in \Omega,$$

which in turn can be restated as

$$f(x) = \left\langle f, \sum_{i\in I} \overline{\phi_i}\phi_i(x) \right\rangle, \quad \forall x \in \Omega.$$

Using the Cauchy Schwarz inequality, we obtain

$$|f(x)| = \left| \left\langle f, \sum_{i\in I} \overline{\phi_i}\phi_i(x) \right\rangle \right| \leq \|f\| \left\| \sum_{i\in I} \overline{\phi_i}\phi_i(x) \right\|.$$

Since

$$K_x = \sum_{i\in I} \overline{\phi_i}\phi_i(x)$$

is assumed to be bounded, the linear point evaluation functional $L_x(f) := f(x)$ is bounded, which implies that H is a RKHS, due to Theorem 1. □

Theorem 7. *If* H_K *is a reproducing kernel Hilbert space of functions over* $\Omega \subset \mathbb{R}^d$*, which contains a frame* $\{\phi_i\}_{i\in I}$*, then the reproducing kernel can be expressed as*

$$K(x,y) = \sum_{i\in I} \overline{\phi_i}(x)\phi_i(y). \tag{10}$$

Proof. On the one hand, in the spirit of in Theorem 6, any function in $f \in \mathsf{H}_K$ can (due to the frame property) be written as

$$f(x) = \left\langle f, \sum_{i\in I} \overline{\phi_i}\phi_i(x) \right\rangle, \quad \forall x \in \Omega, \quad \forall f \in \mathsf{H}_K.$$

On the other hand, since H_K is a RKHS, there is a kernel K, such that

$$f(x) = \langle f, K_x \rangle, \quad \forall x \in \Omega, \quad \forall f \in \mathsf{H}_K.$$

Now, due to the unicity of the reproducing kernel (see Proposition 2), we obtain the kernel decomposition based on a frame $\{\phi_i\}_{i\in I}$ in (10). □

3 Persistent Homology

In this section, we present a short introduction to the basic ideas of persistent homology, which is an important algorithmic and theoretical tool developed over the last decade as a topic of computational topology. First, we present basic concepts on persistent homology as an important new development in computational topology for extracting qualitative information from a point cloud data $X = \{x_i\}$. Here, our interest lies mostly on datasets arising from time frequency analysis and signal processing problems.

3.1 Simplicial and Persistent Homology

We first recall elementary concepts on simplicial homology as a basic homology theory used for constructing algebraic data from topological spaces (see [12] for similar material).

Remark 12 (Simplicial complexes). A basic component in this context is a (finite) *abstract simplicial complex* which is a nonempty family of subsets K of a vertex set $V = \{v_i\}_{i=1}^m$ such that $V \subseteq K$ (here we simplify the notation and we identify the vertex v with the set $\{v\}$) and if $\alpha \in K, \beta \subseteq \alpha$, then $\beta \in K$. The elements of K are denominated *faces*, and their *dimension* is defined as their cardinality minus one. Faces of dimension zero and one are called vertices and edges, respectively. A *simplicial map* between simplicial complexes is a function respecting their structural content by mapping faces in one structure to faces in the other. These concepts represent combinatorial structures capturing the topological properties of many geometrical constructions. Given an abstract simplicial complex K, an explicit topological space is defined by considering the associated *geometric realization* or *polyhedron*, denoted by $|K|$. These are constructed by thinking of faces as higher dimensional versions of triangles or tetrahedrons in a large dimensional Euclidean spaces and gluing them according to the combinatorial information in K.

Remark 13 (Homology groups). A basic analysis tool of a simplicial complex K, is the construction of algebraic structures for computing topological invariants, which are properties of $|K|$ that do not change under homeomorphisms and even continuous deformations. From an algorithmic point of view, we compute topological invariants of K by translating its combinatorial structure in the language of linear algebra. For this task, a basic scenario is to consider the following three steps. First, we construct the *groups of k-chains* C_k, defined as the formal linear combinations of k-dimensional faces of K with coefficients in a commutative ring R (with e.g. $R = \mathbb{Z}$, or $R = \mathbb{Z}_p$). We then consider linear maps between the group of k-chains by constructing the *boundary operators* ∂_k, defined as the linear transformation which maps a face $\sigma = [p_0, \cdots p_n] \in C_n$ into C_{n-1} by $\partial_n \sigma = \sum_{k=0}^{n} (-1)^k [p_0, \cdots, p_{k-1}, p_{k+1}, \cdots p_n]$. As a third step, we construct the *homology groups* defined as the quotient $H_k := \ker(\partial_k)/\mathrm{im}(\partial_{k+1})$. Finally, the concept of *number of k-dimensional holes* are defined using the rank of the homology groups, $\beta_k = \mathrm{rank}(H_k)$ (Betti numbers). For instance, in a sphere we have zero 1-dimensional holes, and one 2-dimensional hole. In the case of a torus, there are two 1-dimensional holes, and one 2-dimensional hole.

3.2 Basics on Persistent Homology

In many application problems a main objective is to analyze experimental datasets $X = \{x_i\}_{i=1}^m \subset \mathbb{R}^n$ and understand their content by computing qualitative information. Topological invariants are important characteristics of geometrical objects, and their properties would be fundamental tools for understanding experimental datasets. The major problem when computing topological invariants of datasets are their finite characteristics and the corresponding inherent instability when computing homological information. Indeed, minor variations (e.g. noise and error in measurements) on how topological structures are constructed from X, could produce major changes on the resulting homological information. Persistent homology [3, 10, 9] is an important computational and theoretical strategy developed over the last decade for computing topological invariants of finite structures. We now describe its motivations, main principles, and theoretical background.

Motivations

A major problem when using tools from simplicial homology for studying a dataset $X = \{x_i\}_{i=1}^m \subset \mathbb{R}^n$ is the fact that we do not have a simplicial complex structure at hand. If we assume that X is sampled from a manifold (e.g. $X \subset \mathscr{M}$, with $\mathscr{M}$ being a submanifold of $\mathbb{R}^n$), a main objective would be to compute homological information of $\mathscr{M}$ using only the dataset X. We remark that more generalized settings, where $\mathscr{M}$ is not necessarily a manifold, are fundamental cases for many applications and experimental scenarios. But we can discuss, for illustration purposes, the simplified situation of $\mathscr{M}$ being a manifold. We also notice that the crucial problem of finding density conditions for X to be a meaningful sampling set of a manifold $\mathscr{M}$ has been recently addressed in [18], and we discuss these issues later in this report.

Attempting to construct a simplicial complex structure from X can be a very difficult problem. A simple strategy would be to consider the homology of the spaces

$$\mathbb{X}_\varepsilon = \cup_{i=1}^m B(x_i, \varepsilon)$$

where a ball $B(x_i, \varepsilon)$ of radius ε is centered around each point of X. A naive approach would be to try to find an optimal ε_o such that the homology of $\mathbb{X}_{\varepsilon_o}$ corresponds to the homology of $\mathscr{M}$. But this approach is highly unstable, as different homological values might be obtained when considering small perturbations of ε_o.

The proposal in *persistent homology* is to consider topological information for all $\varepsilon > 0$ simultaneously, and not just a single value ε_o. The key concept is that a general homological overview for all values $\varepsilon > 0$ is a useful tool when studying the topology of finite datasets. From a computational point of view, estimating homological data for all continuous values $\varepsilon > 0$ might sound unreasonable, but there are two crucial remarks for implementing these ideas in an efficient computational framework. On the one hand, despite the fact that we are considering a continuous parameter $\varepsilon > 0$, it can be verified that for a given dataset X, there is actually only a finite number of non-homeomorphic simplicial complexes

$$K_1 \subset K_2 \subset \cdots \subset K_r$$

(which is the concept of a *filtration* to be explicitly defined later on) that can be constructed from $\{\mathbb{X}_\varepsilon, \varepsilon > 0\}$. On the other hand, another crucial property is that the persistent homology framework includes efficient computational procedures for calculating homological information of the whole family $K_1 \subset K_2 \subset \cdots \subset K_r$, [30].

We also remark that, given a parameter ε with corresponding set $\mathbb{X}_\varepsilon$, there are various topological structures useful for studying homological information of X. In particular, an efficient computational construction is given by the *Vietoris-Rips complexes* $R_\varepsilon(X)$, defined with X as the vertex set, and setting the vertices $\sigma = \{x_0, \ldots, x_k\}$ to span a k-simplex of $R_\varepsilon(X)$ if $d(x_i, x_j) \leq \varepsilon$ for all $x_i, x_j \in \sigma$. For a given ε_k the Vietoris-Rips complex $R_{\varepsilon_k}(X)$ provides an element of the filtration $K_1 \subset K_2 \subset \cdots \subset K_r$, with $K_k = R_{\varepsilon_k}(X)$. In conclusion, there is only a finite set of positive values $\{\varepsilon_i\}_{i=1}^r$, that describe homological characteristics of X, each of which generate a Vietoris Rips complex $\{K_i\}_{i=1}^m$ representing the topological features of the family $\{\mathbb{X}_\varepsilon, \varepsilon > 0\}$. Therefore, the topological analysis of a point cloud data X boils down to the analysis of a filtration $K_1 \subset K_2 \subset \cdots \subset K_r$, which is the main object of study in persistent homology. We now describe the main conceptual ingredients in this framework.

3.3 Conceptual Setting

The input in the persistent homology framework is a *filtration* of a simplicial complex K, defined as a nested sequence of subcomplexes $\emptyset = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_r = K$. Given a simplicial complex K, we recall that the boundary operators ∂_k connect the chain groups C_k, and define a *chain complex*, denoted by C_*, and depicted with the diagram:

$$\ldots C_{k+1} \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \rightarrow \ldots.$$

Recall that given a chain complex C_* one defines the k-cycle groups and the k-boundary groups as $Z_k = \ker \partial_k$, and $B_k = \operatorname{im} \partial_{k+1}$, respectively. As we have nested subgroups $B_k \subseteq Z_k \subseteq C_k$, the k-homology group $H_k = Z_k / B_k$ is well defined.

There are several basic definitions required for the setting of persistent homology. A *persistent complex* is defined as a family of chain complexes $\{C_*^i\}_{i \geq 0}$ over a commutative ring R, together with maps

$$f^i : C_*^i \rightarrow C_*^{i+1} \quad \text{related as} \quad C_*^0 \xrightarrow{f_0} C_*^1 \xrightarrow{f_1} C_*^2 \xrightarrow{f_2} \ldots,$$

or more explicitly, described with the following diagram

$$\begin{tikzcd}
\vdots \arrow[d, "\partial_3"] & \vdots \arrow[d, "\partial_3"] & \vdots \arrow[d, "\partial_3"] & \\
C_2^0 \arrow[r, "f^0"] \arrow[d, "\partial_2"] & C_2^1 \arrow[r, "f^1"] \arrow[d, "\partial_2"] & C_2^2 \arrow[r, "f^2"] \arrow[d, "\partial_2"] & \ldots \\
C_1^0 \arrow[r, "f^0"] \arrow[d, "\partial_1"] & C_1^1 \arrow[r, "f^1"] \arrow[d, "\partial_1"] & C_1^2 \arrow[r, "f^2"] \arrow[d, "\partial_1"] & \ldots \\
C_0^0 \arrow[r, "f^0"] \arrow[d, "\partial_0"] & C_0^1 \arrow[r, "f^1"] \arrow[d, "\partial_0"] & C_0^2 \arrow[r, "f^2"] \arrow[d, "\partial_0"] & \ldots \\
0 \arrow[r, "f^0"] & 0 \arrow[r, "f^1"] & 0 \arrow[r, "f^2"] & \ldots
\end{tikzcd}$$

We remark that, due to the applications we have in mind, we will assume that chain complexes are trivial in negative dimensions. Given a filtration of a simplicial complex K, a basic example of a persistent complex is given by considering the functions f^i as the inclusion maps between each simplicial complex in the nested sequence $\emptyset = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_r = K$. Another fundamental concept is a *persistent module*, defined as a family of R-modules M^i and homomorphisms $\phi^i : M^i \rightarrow M^{i+1}$. We say that the persistent module is of *finite type* if each M^i is finitely generated, and the maps ϕ^i are isomorphisms for $i \geq k$ and some integer k. The basic example of a persistent module is given by the homology of the simplicial complexes of a filtration. We now define the *p-persistent homology group* of K_i as the group

$$H_k^{i,p} = Z_k^i / (B_k^{i+p} \cap Z_k^i),$$

where Z_k^i and B_k^i stand respectively for the k-cycles and k-boundaries groups in C^i. This group can also be described in terms of the inclusions $K^i \subset K^{i+p}$, their induced homomorphisms $f_k^{i,p} : H_k^i \to H_k^{i+p}$, and the corresponding relation

$$\mathrm{im}(f_k^{i,p}) \cong H_k^{i,p}.$$

These persistent homology groups contain homology classes that are stable in the interval i to $i+p$: they are born before the "time" index i and are still alive at $i+p$. Persistent homology classes alive for large values of p, are stable topological features of X, while classes alive only for small values of p are unstable or noise-like topological components. We will see, in the following paragraphs, alternative views for explaining generalized versions of persistent objects as functors between special categories.

The output of the persistent homology algorithm are representations of the evolution, with respect to the parameter $\varepsilon > 0$, of the topological features of X. These representations are depicted with *persistent diagrams* indicating, for each homology level k, the amount and stability of the different k-dimensional holes of the point cloud X. We now present a more precise explanation of the concepts related to persistent diagrams and some of its properties.

The main task we now describe is the analysis of persistent homology groups by capturing their properties in a single algebraic entity represented by a finitely generated module. Recall that a main objective of persistent homology is to construct a summary of the evolution (with respect to ε) of the topological features of X using the sets $\{\mathbb{X}_\varepsilon, \varepsilon > 0\}$. This property is analyzed when constructing, with the homology groups of the complexes K_i, a module over the polynomial ring $R = \mathbb{F}[t]$ with a field $\mathbb{F}$. The general setting for this procedure is to consider a persistent module $M = \{M^i, \phi_i\}_{i\geq 0}$ and construct the graded module $\alpha(M) = \bigoplus_{i\geq 0} M^i$ over the graded polynomial ring $\mathbb{F}[t]$, defined with the action of t given by the shift $t \cdot (m^0, m^1, \dots) = (0, \phi^0(m^0), \phi^1(m^1), \dots)$. The crucial property of this construction is that α is a functor that defines an equivalence of categories between the category of persistent modules of finite type over $\mathbb{F}$, and the category of finitely generated non-negatively graded modules over $\mathbb{F}[t]$. In the case of a filtration of complexes K_0 to K_r, this characterization of persistent modules provides the finitely generated $\mathbb{F}[t]$ module:

$$\alpha(M) = H_p(K_0) \oplus H_p(K_1) \oplus \cdots \oplus H_p(K_r).$$

These modules are now used in a crucial step that defines and characterizes the output of persistent homology. The main tool is the well-know structure theorem characterizing finitely generated modules over principle ideal domains (this is why we need $\mathbb{F}$ to be a field). This property considers a finitely generated non-negatively graded module $\mathfrak{M}$, and ensures that there are integers $\{i_1, \dots, i_m\}$, $\{j_1, \dots, j_n\}$, $\{l_1, \dots, l_n\}$, and an isomorphism:

$$\mathfrak{M} \cong \bigoplus_{s=1}^{m} \mathbb{F}[t](i_s) \oplus \bigoplus_{r=1}^{n} (\mathbb{F}[t]/(t^{l_r}))(j_r).$$

This decomposition is unique up to permutation of factors, and the notation $\mathbb{F}[t](i_s)$ denotes an i_s shift upward in grading. The relation with persistent homology is given by the fact that when a persistent homology class τ is born at K_i and dies at K_j it generates a torsion module of the form $\mathbb{F}[t]\tau/t^{j-i}(\tau)$. When a class τ is born at K_i but does not die, it generates a free module of the form $\mathbb{F}[t]\tau$.

We can now explain the concept of persistent diagrams using an additional characterization of $\mathbb{F}[t]$-modules. We first define a P-interval as an ordered pair (i,j) where $0 \leq i < j$ for $i,j \in \mathbb{Z} \cup \{\infty\}$. We now construct the function Q mapping a P-interval as $Q(i,j) = (\mathbb{F}[t]/t^{j-i})(i)$, $Q(i,\infty) = \mathbb{F}[t](i)$, and for a set of P-intervals $S = \{(i_1, j_1), \ldots, (i_n, j_n)\}$, we have the $\mathbb{F}[t]$-module

$$Q(S) = \bigoplus_{\ell=1}^{n} Q(i_\ell, j_\ell).$$

This map Q turns out to be a bijection between the sets of finite families of P-intervals and the set of finitely generated graded modules over $\mathbb{F}[t]$.

Now, we can recap all these results by noticing that the concept of persistent diagrams can be described as the corresponding set of P-intervals associated to the finitely generated graded module over $\mathbb{F}[t]$, constructed with the functor α from a given filtration $\emptyset = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_r = K$. There are several graphical representations for persistent diagrams, and two well known examples are the so called barcodes, and triangular regions of index-persistent planes.

3.4 Generalizations with Functorial Properties

In order to design useful generalizations of persistent homology, it is important to understand its setting in a deeper conceptual level. A recent formulation, providing the core features of persistent homology, has been presented in [3], and describes this concept as a functor between well chosen categories. Indeed, a crucial aspect of persistent homology is the association from an index set to a sequence of homology groups constructed from a filtration $\emptyset = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_r = K$. An important generalization of this construction considers a general partially ordered set P as an index set which we associate to a family of objects in a given category $\mathbf{C}$. Notice that we can consider the partially ordered set P as a category $\mathbf{P}$, whose objects are P, and a morphism from x to y is defined whenever $x \leq y$. With this setting, a P-persistent object in $\mathbf{C}$ is defined as a functor $\Phi : \mathbf{P} \to \mathbf{C}$, described also as a family of objects $\{c_x\}_{x \in P}$ in $\mathbf{C}$, and morphisms $\phi_{xy} : c_x \to c_y$, when $x \leq y$.

These concepts are of fundamental importance for extending the main ideas of persistent homology in more general situations. Notice that in standard persistent homology we use the partial ordered sets $P = \mathbb{N}$ or $P = \mathbb{R}$, but important extensions have been recently developed in the context of multidimensional persistence. Here, we consider multidimensional situations where the partial ordered sets are, for instance, $P = \mathbb{N}^k$ or $P = \mathbb{R}^k$, $k > 1$. These developments are motivated by multiple practical considerations, such as the analysis of point cloud using both density estimations and the Vietoris Rips Complex construction.

3.5 Stability Properties

A crucial property in persistent homology is the concept of *stability of persistent diagrams*. We recall that for a topological space X, and a map $h: X \to \mathbb{R}$, we say that h is *tame* if the homology properties of $\{X_\varepsilon, \varepsilon > 0\}$, for $X_\varepsilon = h^{-1}(]-\infty, \varepsilon[)$, can be completely described with a finite family of sets $X_{a_0} \subset X_{a_1}, \subset \cdots \subset X_{a_r}$, where the positive values $\{a_i\}_{i=0}^r$ are *homology critical points*. If we denote the *persistent diagram* for X and $h: X \to \mathbb{R}$, as $\mathrm{dgm}_n(h)$, we have a summary of the *stable and unstable* holes generated by the filtration

$$X_{a_0} \subset X_{a_1}, \subset \cdots \subset X_{a_r}$$

(see [9]). With these concepts, the *stability of persistent diagrams* is a property indicating that small changes in the persistent diagram $\mathrm{dgm}_n(h)$ can be controlled with small changes in the tame function $h: X \to \mathbb{R}$ (see [7] for details on the stability properties of persistent diagrams).

An important theoretical and engineering problem to investigate is the sensibility of the persistent homology features of X_f when applying signal transformations to f. This is in relation to the question of finding useful signal invariants using the persistent diagram of X_f. For instance, in the case of audio analysis, a crucial task is to understand the effects in the persistent diagram of X_f when applying audio transformations to f as, for instance, delay filters or convolution transforms (e.g room simulations). This task requires both theoretical analysis and numerical experiments. For a conceptual analysis, a possible strategy is to consider these recent theorems explaining the stability of persistent diagrams.

In order to introduce the idea of *stability of persistent diagrams* we now introduce with more detail the basic concepts.

Definition 5. Let X be a topological space, and $\alpha: \mathsf{X} \to \mathbb{R}$ a continuous function. A *homological critical value* (or HCV) is a number $a \in \mathbb{R}$ for which the map induced by α

$$H_n(\alpha^{-1}(]-\infty, a-\varepsilon[)) \to H_n(\alpha^{-1}(]-\infty, a+\varepsilon[))$$

is not an isomorphism for all $\varepsilon > 0$. Remember that each $\alpha^{-1}(]-\infty, a[)$ is a *level sets* of α, and it plays a crucial role in Morse theory, as well as in our current setting. A *tame function* is now defined to be a function $\alpha: \mathsf{X} \to \mathbb{R}$ that has only a finite number of HCV.

Typical examples of tame functions are *Morse functions* on compact manifolds, and piecewise linear functions on finite simplicial complexes [7].

Definition 6. For a tame function $\alpha : \mathsf{X} \to \mathbb{R}$, we define its *persistent diagram* $\mathrm{dgm}(\alpha)$ as the persistent diagram of the filtration $K_1 \subset K_2 \subset \cdots \subset K_r = \mathsf{X}$ where we let $K_i = f^{-1}(]-\infty, a_i])$, and $a_1 < a_2 < \cdots < a_r$ are the critical values of α (cf. [4]).

Definition 7. For two nonempty sets $X, Y \subset \mathbb{R}^2$ the *Hausdorff distance* and *bottleneck distances* are defined as

$$d_H(X,Y) = \max\Big\{\sup_{x\in X}\inf_{y\in Y}||x-y||_\infty, \sup_{y\in Y}\inf_{x\in X}||y-x||_\infty\Big\},$$
$$d_B(X,Y) = \inf_{\gamma}\sup_{x\in X}||x-\gamma(x)||_\infty$$

where we consider all possible bijections $\gamma : X \to Y$. Here, we use

$$||p-g||_\infty = \max\{|p_1-q_1|, |p_2-q_2|\}, \qquad \text{for } p,q \in \mathbb{R}^2.$$

We also have the inequality between these distances: $d_H(X,Y) \le d_B(X,Y)$ (see [7]).

Theorem 8. *Let X be a topological space with tame functions $\alpha, \beta : \mathsf{X} \to \mathbb{R}$. Then, the following stability property holds:*

$$d_B(dgm(\alpha), dgm(\beta) \le ||\alpha-\beta||_\infty. \tag{11}$$

4 Interactions with Persistent Homology

We finally present interactions between kernels and frames with persistent homology. We remark that these interactions provide basic concepts for exploring new possibilities on relating frames, kernels and persistent homology. We discuss two particular aspects: relations between kernels and persistent homology and relations between frames and persistent homology. Recent developments have uncovered specific interactions with novel ways for studying data and signal in different contexts.

4.1 Interactions between Kernels and Persistent Homology

In this section we describe recent developments for constructing a particular kernel for persistent diagrams which provides an interaction between kernel methods and persistent homology (see [22] for details).

We first remark that the bottleneck distance used in our previous discussions can be embedded in *p-Wasserstein distances*, which is a more general type of distance defined for any positive real number p as

$$d_{W,p}(F,G) = \left(\inf_{\gamma}\sum_{x\in F}||x-\gamma(x)||_\infty^p\right)^{1/p}.$$

Here, as before, γ ranges over all bijections from the elements of F to the elements of G. Notice that as $p \to \infty$, we have $d_{W,\infty} = d_B$.

We now define the particular kernel on the space of persistent diagrams described in [22]. The intuition behind these ideas is to use *scale-space theory* which considers a particular type of multi-scale decomposition of signals. This theory has been applied in image processing, and the idea is to apply these mechanisms to persistent diagrams seen as special type of images. This decomposition is given by the evolution derived by the partial differential equation for the heat diffusion problem. Each step in this evolution corresponds to one particular scale, and we construct a corresponding family of kernels for each scale, as indicated in the following definition.

Definition 8 ([22]). Let $\Omega = \{x = (x_1, x_2) \in \mathbb{R}^2, x_2 \geq x_1\}$ be the space above the diagonal, and let δ_p be a Dirac delta centered at the point p. For a persistent diagram D, we consider the solution $u : \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}$, $(x,t) \to u(x,t)$ of the partial differential equation $\Delta_x u = \partial_t u$ in $\Omega \times \mathbb{R}_{\geq 0}$, $u = 0$ on $\partial\Omega \times \mathbb{R}_{\geq 0}$, $u = \sum_{p \in D} \delta_p$ on $\partial\Omega \times \{0\}$. Using the set of persistent diagrams D, the feature map $\Omega_\sigma : \mathsf{D} \to L_2(\Omega)$ at scale σ is defined as $\Phi_\sigma(D) = u|_{t=\sigma}$, (namely $\Phi_\sigma(D)(x) = u(x,\sigma), x \in \Omega$). This map provides a kernel k_σ (the persistence scale space kernel) on D with

$$k_\sigma(F,G) = \langle \Phi_\sigma(F), \Phi_\sigma(G) \rangle_{L^2(\Omega)}.$$

With the concepts just introduced, we can now cite a main result of [22], where a stability property is described involving the application of the feature map to the persistent diagram. This presents a new type of interaction between persistent homology and kernel methods.

Theorem 9 ([22]). *The kernel k_σ is 1-Wasserstein stable, namely, for F, G two persistent diagrams, and a feature map Φ_σ, we have*

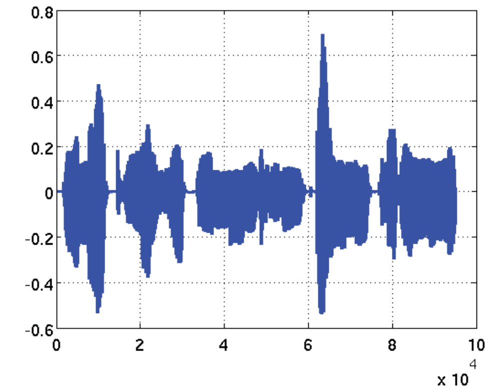
$$||\Phi_\sigma(F) - \Phi_\sigma(G)||_{L_2(\Omega)} \leq \frac{1}{\sigma\sqrt{8\pi}} d_{W,1}(F,G).$$

4.2 Interactions between Frames and Persistent Homology

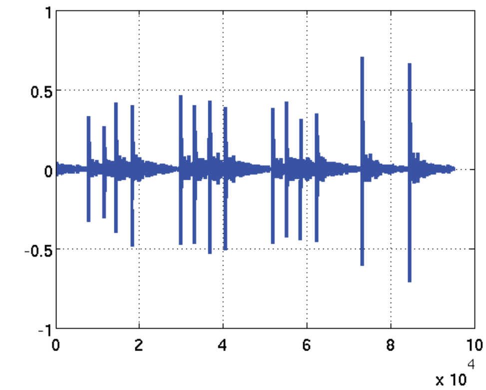
We now provide one particular interaction between frames and persistent homology, as described in more details in our previous work [1, 13]. Our basic result provides theoretical statements concerning stability properties of persistent diagrams of frame transforms $|Vf|$, when considering a frame decomposition $Vf \in L^2(\mathsf{X})$, where X is the parameter set of the frame $\{\psi_x\}_{x \in \mathsf{X}}$, and $f \in \mathsf{H}$.

Theorem 10 ([1]). *Let $f, g \in \mathsf{H}$ and assume $|Vf|$ and $|Vg|$ are tame functions with $V : \mathsf{H} \to L^2(\mathsf{X})$ a frame analysis operator, where we consider a discrete topological space X with a counting measure. Then, the following stability property holds.*

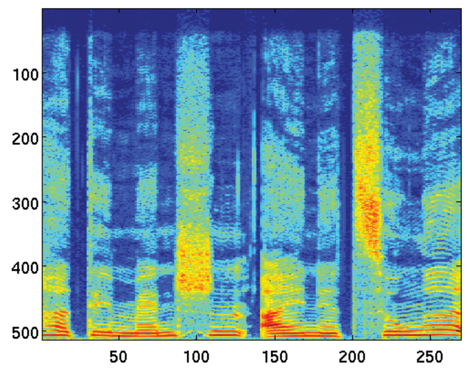
$$d_B(dgm(|Vf|), dgm(|Vg|)) \leq \sqrt{B}||f-g||_{\mathsf{H}}.$$

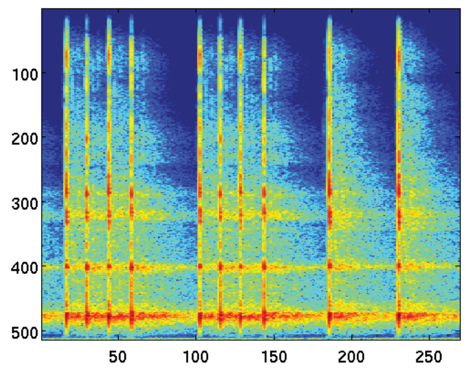


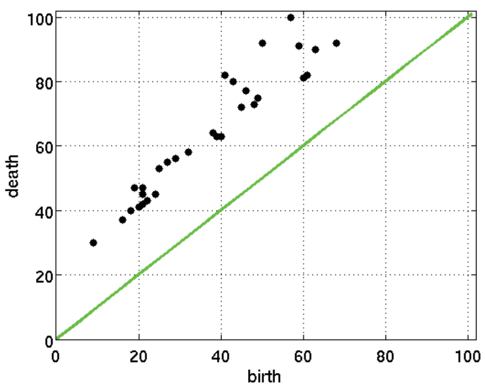
a) signal f_0



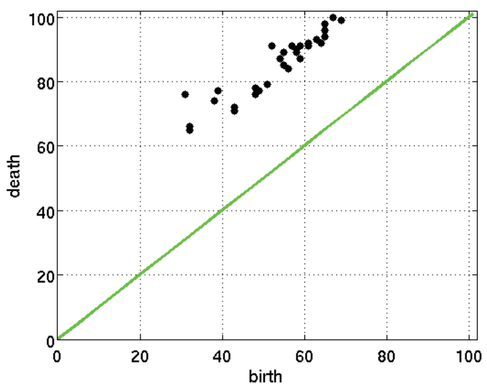
b) signal f_1


c) $|V f_0|$: spectrogram of f_0


d) $|V f_1|$: spectrogram of f_1



e) $\mathrm{dgm}(|V f_0|)$ (1-homology)



f) $\mathrm{dgm}(|V f_1|)$ (1-homology)

Fig. 1 Time-frequency plots and discriminative properties of persistence (experiment from [1]).

Proof. This is a consequence of the inequality (8) (the bounding of the norm of the frame operator) and the stability of the persistent diagrams described in (11):

$$
\begin{aligned}
&d_B(\mathrm{dgm}(|Vf|), \mathrm{dgm}(|Vg|)) \\
&\qquad \le ||\,|Vf| - |Vg|\,||_\infty \\
&\qquad \le ||Vf - Vg||_2 && \text{(counting measure property)} \\
&\qquad \le ||V||\;||f-g||_{\mathsf{H}} \\
&\qquad \le \sqrt{||V^*V||}||f-g||_{\mathsf{H}} && ||V||^2 = ||V^*V|| \\
&\qquad \le \sqrt{||S||}\;||f-g||_{\mathsf{H}} \\
&\qquad \le \sqrt{B}||f-g||_{\mathsf{H}}.
\end{aligned}
$$

□

This proposition is an initial step towards the integration of frame theory and persistent stability. We remark that new developments have been achieved in generalizing the work in [7], and the inequality (11), by avoiding the restrictions imposed by the functional setting and expressing the stability in a purely algebraic language (see [3, 4]). The usage of these more flexible and general stability properties is a natural future step in our program.

As an illustrative example, we present in Fig. 1 two acoustic signals f_0, f_1 and their corresponding spectrograms (STFT) $|Vf_0|$ and $|Vf_1|$. These represent a particular frame construction as required in the Theorem 10. In Fig. 1(e) and Fig.1(f) we display the persistent homology diagrams when considering $|Vf_0|$ and $|Vf_1|$ as two dimensional functions and analyzing the corresponding level sets as indicated in the definitions 5 and 6. In these persistent diagrams we have selected only the 30 most prominent 1-dimensional homological structures, displayed by the 30 dots with the largest distance to the diagonal (identity function) in the persistent diagram. These diagrams can be seen as homological fingerprints describing topological features of the corresponding spectrograms. These persistent diagrams can be used to identify and discriminate these spectrograms using a compact or sparse representation. We are then displaying a new type of characterization of time-frequency data using topological properties for identifying and discriminating signals.

References

1. H. Boche, M. Guillemard, K. G., and F. Philipp: Signal analysis with frame theory and persistent homology. In: *Proc. Sampling Theory and Applications (SampTA'13)*, 2013.
2. M.D. Buhmann: *Radial Basis Functions.* Cambridge University Press, Cambridge, UK, 2003.
3. G. Carlsson: Topology and data. *American Mathematical Society*, 46(2):255–308, 2009.
4. F. Chazal, V. de Silva, M. Glisse, and S. Oudot: The structure and stability of persistence modules. *arXiv:1207.3674*, 2012.
5. E.W. Cheney: *Introduction to Approximation Theory.* 2nd edition. McGraw Hill, New York, 1982.

6. W. Cheney and W. Light: *A Course in Approximation Theory*. Brooks/Cole Publishing, Pacific Grove, 2000.
7. D. Cohen-Steiner, H. Edelsbrunner, and J. Harer: Stability of persistence diagrams. *Discrete and Computational Geometry*, 37(1):103–120, 2007.
8. J. Dieudonné: *Éléments D'Analyse. Tome 1*. Paris, Gauthier-Villars, 1971.
9. H. Edelsbrunner and J. Harer: Persistent homology - a survey. In: *Surveys on Discrete and Computational Geometry: Twenty Years Later*, AMS-IMS-SIAM Joint Summer Research Conference, June 18-22, 2006, Snowbird, Utah, vol. 453, page 257. American Mathematical Society, 2008.
10. H. Edelsbrunner, D. Letscher, and A. Zomorodian: Topological persistence and simplification. In: *Proc. 41st Ann. IEEE Sympos. Found Comput. Sci.*, pages 454–463, 2000.
11. M. Fornasier and H. Rauhut: Continuous frames, function spaces, and the discretization problem. *J. of Fourier Anal and Appl.*, 11(3):245–287, 2005.
12. J. González and M. Guillemard: Algunas aplicaciones de la topología algebráica. In: *Aportaciones Matemáticas, Sociedad Matemática Mexicana*, 42:153–170, 2011.
13. M. Guillemard and A. Iske: Signal filtering and persistent homology: an illustrative example. In: *Proc. Sampling Theory and Applications (SampTA'11)*, 2011.
14. A. Iske: *Multiresolution Methods in Scattered Data Modelling*. Springer, 2004.
15. A.A. Kirillov and A.D. Gvishiani: *Theorems and Problems in Functional Analysis*. Springer, 1982.
16. H. Meschkowski: *Hilbertsche Räume mit Kernfunktion*. Springer, Berlin, 1962.
17. H. Minh, P. Niyogi, and Y. Yao: Mercer's theorem, feature maps, and smoothing. In: *Proceedings of the 19th Annual Conference on Learning Theory (COLT)*, 2006.
18. P. Niyogi, S. Smale, and S. Weinberger: Finding the homology of submanifolds with high confidence from random samples. *Discrete and Computational Geometry*, 39(1):419–441, 2008.
19. R. Opfer: Tight frame expansions of multiscale reproducing kernels in Sobolev spaces. *Applied and Computational Harmonic Analysis*, 20:357–374, 2006.
20. M.J.D. Powell: *Approximation Theory and Methods*. Cambridge University Press, Cambridge, UK, 1981.
21. A. Rakotomamonjy and S. Canu: Frames, reproducing kernels, regularization and learning. *Journal of Machine Learning Research*, 06:1485–1515, September 2005.
22. J. Reininghaus, S. Huber, U. Bauer, and R. Kwitt: A stable multi-scale kernel for topological machine learning. In: *2015 (CVPR '15)*, Boston, MA, USA, June 2015.
23. R. Schaback: Native Hilbert spaces for radial basis functions I. *International Series of Numerical Mathematics*, 132:255–282, 1999.
24. R. Schaback: A unified theory of radial basis functions: Native Hilbert spaces for radial basis functions II. *Journal of Computational and Applied Mathematics*, 121(1-2):165–177, 2000.
25. R. Schaback: *Kernel-Based Meshless Methods*. Lecture Notes, Universität Göttingen, 2007.
26. A. Shadrin: *Approximation Theory*. Lecture Notes, DAMTP, University of Cambridge, 2005.
27. I. Steinwart, D. Hush, and C. Scovel: An explicit description of the reproducing kernel Hilbert spaces of Gaussian RBF kernels. *IEEE Transactions on Information Theory*, 52(10):4635–4643, 2006.
28. M. van der Laan: *Signal Sampling Techniques for Data Acquisition in Process Control*. PhD thesis, University of Groningen, The Netherlands, 1995.
29. H. Wendland: *Scattered Data Approximation*. Cambridge University Press, Cambridge, UK, 2005.
30. A. Zomorodian and G. Carlsson: Computing persistent homology. *Discrete Comput. Geom.*, 33(2):249–274, 2005.