Out-of-Sample Extension of the Fuzzy Transform

Giuseppe Patané

***Abstract*—This paper addresses the definition and computation of the out-of-sample membership functions and the resulting out-of-sample FT, which extend their discrete counterparts to the continuous case. Through the out-of-sample FT, we introduce a coherent analysis of the discrete and continuous FTs, which is applied to extrapolate the behaviour of the FT on new data and to achieve an accurate approximation of the continuous FT of signals on arbitrary data. To this end, we apply either an approximated approach, which considers the link between integral kernels and the spectrum of the corresponding Gram matrix, or an interpolation of the discrete kernel eigenfunctions with radial basis functions. In this setting, we show the generality of the proposed approach to the input data (e.g., graphs, 3D domains) and signal reconstruction.**

***Index Terms*—FT, Data-driven membership functions, Radial basis functions, Signal approximation, Data analysis**

I. Introduction

Due to the increasing availability of data, which is supported by ongoing technological advances in the acquisition, storage, and processing, several transformations (e.g., the Fourier transform, the Laplace transform, and the Fuzzy transform) have been proposed to solve problems that spread from signal analysis to the solution of partial differential equations, from the analysis to the approximation of signals, from fuzzy logic to fuzzy modelling. To deal with the definition of the Fuzzy transform (FT, for short) on arbitrary data (Sect. II), the data-driven FT [1] and the continuous FT [2] have been defined by applying the main concepts and results of spectral signal processing and manifold learning to the FT, e.g., dimensionality reduction [3] and kernels' eigenfunctions [4]. The data-driven and continuous FTs provide a link between the fuzzy theory and previous work on diffusion kernels [5] and wavelets [5]–[8], and geometric deep learning [9].

An input signal $f:\Omega\rightarrow\mathbb{R}$ is generally known at a set $\mathcal{P}$ of points and the corresponding FT $\mathcal{F}f$ is computable only at $\mathcal{P}$. As a result, we cannot estimate the value of $\mathcal{F}f$ at a new point not belonging to $\mathcal{P}$. Increasing the sampling of $\mathcal{F}f$, and consequently of the reconstructed signal $\mathcal{F}^{-1}(\mathcal{F}f)$, will require adding new points in $\mathcal{P}$ and evaluating the corresponding function values, the FT, and its inverse. This last option is not generally feasible (e.g., if the signal values have been experimentally measured or their evaluation is time-consuming) and requires computing the FT and its inverse again. Since the FT of a continuous signal is still continuous, we expect its behaviour to be accurately recovered at any point from a set of enough dense signal samples without further resampling the signal itself. The need to extrapolate the behaviour of the FT $\mathcal{F}f$ and of the reconstructed signal $\mathcal{F}^{-1}(\mathcal{F}f)$ out of the set of input samples is further justified by the observation that most of the signals exist outside the input domain. For instance, we can measure the heat values at a set of points on a thin plate (i.e., a bounded 2D domain) but recover the heat distribution on and around the plate.

For these reasons, we address the problem of defining the *out-of-sample membership functions* and the resulting *out-of-sample FT* (Sect. III), which extend their discrete counterparts to the continuous case. Through the out-of-sample FT, we introduce a coherent analysis of the discrete and continuous FTs, which extrapolate the behaviour of the FT on new data and achieve an accurate approximation of the continuous FT of signals on arbitrary data (Fig. 1). We also characterise the data-driven FT [1], [2] by representing the normalisation factor in terms of the input filter and the area/volume of the input domain (Sect. IV). As main properties of the out-of-sample extension of the FT, we discuss its linearity and continuity, self-adjointness and eigensystem, convergence and fast computation.

To this end, we apply (i) an *approximated approach* (Sects. V, VI) that considers the link between integral kernels and the spectrum of the corresponding Gram matrix. The out-of-sample membership functions and the related FT are more robust to noisy data and tailored to address signal denoising and multi-scale representations. Alternatively, we apply (ii) an *interpolation* of the discrete kernel eigenfunctions with radial basis functions, which is more accurate for "noise-free" data and useful for super-resolution and feature detection. Then, we characterise the relations between Fuzzy theory and manifold learning, explicitly focusing on integral and spectral kernels. The out-of-sample membership functions and FT on large data are efficiently computed by approximating the $\mathcal{L}^2(\Omega)$ scalar product with a *pseudoscalar product*, which encodes the weights associated with graph edges or the area/volume of Voronoi regions for data represented as surface/volume meshes and preserves the main properties of the out-of-sample extension of the data-driven FT. Finally, we discuss different tests on the out-of-sample FT and signal reconstruction in the experimental part.

II. Previous work

We review the FT (Sect. II-A), its extension to *continuous FT* [2] and the *data-driven FT* [1] (Sect. II-B).

A. Fuzzy tranform

Let us consider the space $\mathcal{L}^2(\Omega)$ of square integrable functions defined on a compact and connected domain Ω of $\mathbb{R}^n$, endowed with the $\mathcal{L}^2(\Omega)$ scalar product $\langle f,g\rangle_2:=\int_\Omega f(\mathbf{p})g(\mathbf{p})d\mathbf{p}$ and the corresponding norm

G. Patané is with CNR-IMATI, Consiglio Nazionale delle Ricerche, Istituto di Matematica Applicata e Tecnologie Informatiche Genova, Italy. E-mail: patane@ge.imati.cnr.it

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2023.3326657

$\|f\|_2^2 := \int_\Omega |f(\mathbf{p})|^2 d\mathbf{p}$. On the space $\mathcal{C}^0(\Omega)$ of continuous functions defined on Ω, we consider the $\mathcal{L}^2(\Omega)$ and the $\mathcal{L}^\infty(\Omega)$-norm $\|f\|_\infty := \max_{\mathbf{p}\in\Omega}\{|f(\mathbf{p})|\}$. Given a set $\mathcal{P} := \{\mathbf{p}_i\}_{i=1}^n$ of points in Ω, a family of functions $\mathcal{A} := \{A_i : \Omega \to [0,1]\}_{i=1}^n$ is a *fuzzy partition* of Ω if

- A_i is continuous, has its unique maximum 1 at $\mathbf{p}_i$;
- for all $\mathbf{p} \in \Omega$, $\sum_{i=1}^n A_i(\mathbf{p}) = 1$,

for each i. Then, the *FT* [10]–[13] of a function $f : \Omega \to \mathbb{R}$ is defined as the array $\mathbf{F}_n := (F_i)_{i=1}^n \in \mathbb{R}^n$ with components

$$F_i := \frac{\int_\Omega A_i(\mathbf{p}) f(\mathbf{p}) d\mathbf{p}}{\int_\Omega A_i(\mathbf{p}) d\mathbf{p}}, \qquad i = 1, \ldots, n. \tag{1}$$

Since the function f is known at a set of points $\mathcal{Q} := \{\mathbf{q}_i\}_{i=1}^s$ in Ω, the definition (1) is replaced by the *discrete FT* $\mathbf{F}_n := (F_i)_{i=1}^n \in \mathbb{R}^n$, whose components are

$$F_i := \frac{\sum_{j=1}^s A_i(\mathbf{q}_j) f(\mathbf{q}_j)}{\sum_{j=1}^s A_i(\mathbf{q}_j)}, \qquad i = 1, \ldots, n, \quad s \leq n.$$

The discrete FT is applied to recover an approximation $f_{F,n}$ of the function f underlying the set of values $(f(\mathbf{q}_i))_{i=1}^s$ through the *inverse FT* [10], which is defined as $f_{F,n}(\mathbf{p}) := \sum_{i=1}^n F_i A_i(\mathbf{p})$, $\mathbf{p} \in \mathbb{R}^d$.

FT as integral operator: Given a symmetric kernel $A : \Omega \times \Omega \to [0,1]$ (i.e., $A(\mathbf{p},\mathbf{q}) = A(\mathbf{q},\mathbf{p})$, $\mathbf{p},\mathbf{q} \in \Omega$), the *continuous FT* [1], [2] of $f : \Omega \to \mathbb{R}$ is defined as the continuous function $F : \Omega \to \mathbb{R}$

$$\begin{aligned} F(\mathbf{p}) &:= \frac{\int_\Omega A(\mathbf{p},\mathbf{q}) f(\mathbf{q}) d\mathbf{q}}{D(\mathbf{p})} = \int_\Omega K(\mathbf{p},\mathbf{q}) f(\mathbf{q}) d\mathbf{q} \\ &=: (\mathcal{L}_K f)(\mathbf{p}), \qquad D(\mathbf{p}) := \int_\Omega A(\mathbf{p},\mathbf{q}) d\mathbf{q}. \end{aligned} \tag{2}$$

Here, $\mathcal{L}_K$ is the *integral operator* induced by the *normalised kernel*

$$K : \Omega \times \Omega \to \mathbb{R}, \qquad K(\mathbf{p},\mathbf{q}) := \frac{A(\mathbf{p},\mathbf{q})}{D(\mathbf{p})}, \tag{3}$$

Noting that $F_i = F(\mathbf{p}_i) = (\mathcal{L}_K f)(\mathbf{p}_i)$ is the i-th component F_i of the FT of f, the continuous FT $F(\cdot)$ interpolates the values $\mathbf{F}_n := (F_i)_{i=1}^n$ of the discrete FT. According to Eq. (2), any FT associated with the membership function $A_\mathbf{p} : \Omega \to \mathbb{R}$ is re-written as an integral operator induced by the normalised kernel (3) with $A(\mathbf{p},\mathbf{q}) := A_\mathbf{p}(\mathbf{q})$. Viceversa, any symmetric kernel $K : \Omega \times \Omega \to \mathbb{R}$ can be considered as a membership function $A_\mathbf{p}(\mathbf{q}) : \Omega \to \mathbb{R}$, $A_\mathbf{p} := K(\mathbf{p},\mathbf{q})$, and the corresponding FT is equal to Eq. (2).

B. Data-driven Fuzzy transform

The data-driven membership functions and FT [1], [2] are defined in terms of the *Laplacian orthonormal eigensystem* $(\lambda_n, \phi_n)_{n=0}^{+\infty}$ [14], $\Delta\phi_n = \lambda_n \phi_n$, with $\lambda_0 = 0$ and $\lambda_n \leq \lambda_{n+1}$. Given a strictly, positive, continuous, and square-integrable filter $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$, we consider the power series $\varphi(s) = \sum_{n=0}^{+\infty} \alpha_n s^n$ and define the *data-driven FT*

$$\mathcal{L}_{K_\varphi} f := \varphi(\Delta) f := \sum_{n=0}^{+\infty} \varphi(\lambda_n) \langle f, \phi_n \rangle_2 \phi_n = \langle K_\varphi, f \rangle_2, \tag{4}$$

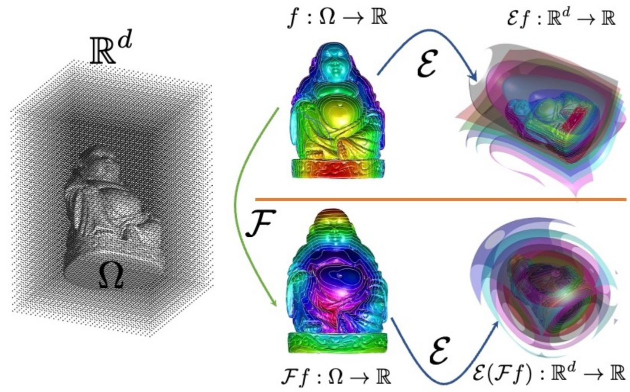

Fig. 1. Given a function $f : \Omega \to \mathbb{R}$, the scheme summarised the definition of the out-of-sample extensions $\mathcal{E}f$, $\mathcal{E}(\mathcal{F}f)$ of f and its FT $\mathcal{F}f$.

where $K_\varphi(\mathbf{p},\mathbf{q}) := \sum_{n=0}^{+\infty} \varphi(\lambda_n)\phi_n(\mathbf{p})\phi_n(\mathbf{q})$ is the *spectral kernel* and $A(\mathbf{p},\cdot) := K_\varphi(\mathbf{p},\cdot) = \varphi(\Delta)\delta_\mathbf{p}$ is the spectral membership function at $\mathbf{p}$. Selecting the filter function, we define different classes of membership functions; main examples are the *diffusion* ($\varphi_t(s) := \exp(-st)$), *poly-harmonic* ($\widetilde{\varphi}(s) := (s+\alpha)^{-k/2}$), and *commute-time* ($\widetilde{\varphi}(s) := (s+\alpha)^{-1/2}$) *filters* [9], [15], [16], $\alpha > 0$.

III. Out-of-sample Fuzzy transform

In a general setting, an input signal $f : \Omega \to \mathbb{R}$ is known at a set $\mathcal{P}$ of points and the corresponding FT $\mathcal{F}f$ is computable only at $\mathcal{P}$. As a result, we cannot estimate the value of $\mathcal{F}f$ at a new point not belonging to $\mathcal{P}$. Increasing the sampling of $\mathcal{F}f$, and consequently of the reconstructed signal $\mathcal{F}^{-1}(\mathcal{F}f)$, will require adding new points in $\mathcal{P}$ and evaluating the corresponding function values, the FT, and its inverse. This last option is not generally feasible (e.g., if the signal values have been experimentally measured or their evaluation is time-consuming) and requires computing the FT and its inverse again.

Since the FT of a continuous signal is still continuous, we expect that its behaviour can be accurately recovered at any point from a set $\mathbf{f}$ of dense signal samples. To this end, we address the out-of-sample extension of the corresponding FT $\mathcal{F}\mathbf{f} =: \mathbf{F}$, i.e. the computation of a function $\mathcal{E}(\mathcal{F}\mathbf{f}) : \mathbb{R}^d \to \mathbb{R}$ such that $\mathcal{E}(\mathcal{F}\mathbf{f})(\mathbf{p}_i) = (\mathcal{F}\mathbf{f})(\mathbf{p}_i) = F_i$, $i = 1, \ldots, n$. The out-of-sample extension of the FT is a natural way to recover the continuous formulation of the FT from its discrete version. More precisely we study the relationship between the discrete and continuous FTs through the *sampling operator* $\mathcal{R} : \mathcal{C}^0(\Omega) \to \mathbb{R}^s$, $f \mapsto \mathcal{R}f := f|_\mathcal{P} := (f(\mathbf{q}_i))_{i=1}^s = \mathbf{f}$, and the *out-of-sample operator* $\mathcal{E} : \mathbb{R}^s \to \mathcal{C}^0(\Omega)$, $\mathbf{f} = (f_i)_{i=1}^s \mapsto \mathcal{E}\mathbf{f}$, with $(\mathcal{E}\mathbf{f})(\mathbf{q}_i) = f_i$, $\forall i$. We require that the out-of-sample operator is linear: i.e., $\mathcal{E}(\alpha\mathbf{f} + \beta\mathbf{g}) = \alpha\mathcal{E}\mathbf{f} + \beta\mathcal{E}\mathbf{g}$, $\forall \alpha, \beta$, $\forall \mathbf{f}, \mathbf{g}$.

Applying the out-of-sample operator to the set $\mathbf{f} := (f(\mathbf{q}_i))_{i=1}^s$ of the f-values at $\mathcal{Q}$, we consider the diagram

$$\mathbf{f} \in \mathbb{R}^s \mapsto \mathcal{E}\mathbf{f} \in \mathcal{C}^0(\Omega) \mapsto (\mathcal{L}_K \mathcal{E}\mathbf{f}) \in \mathcal{C}^0(\Omega).$$

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2023.3326657

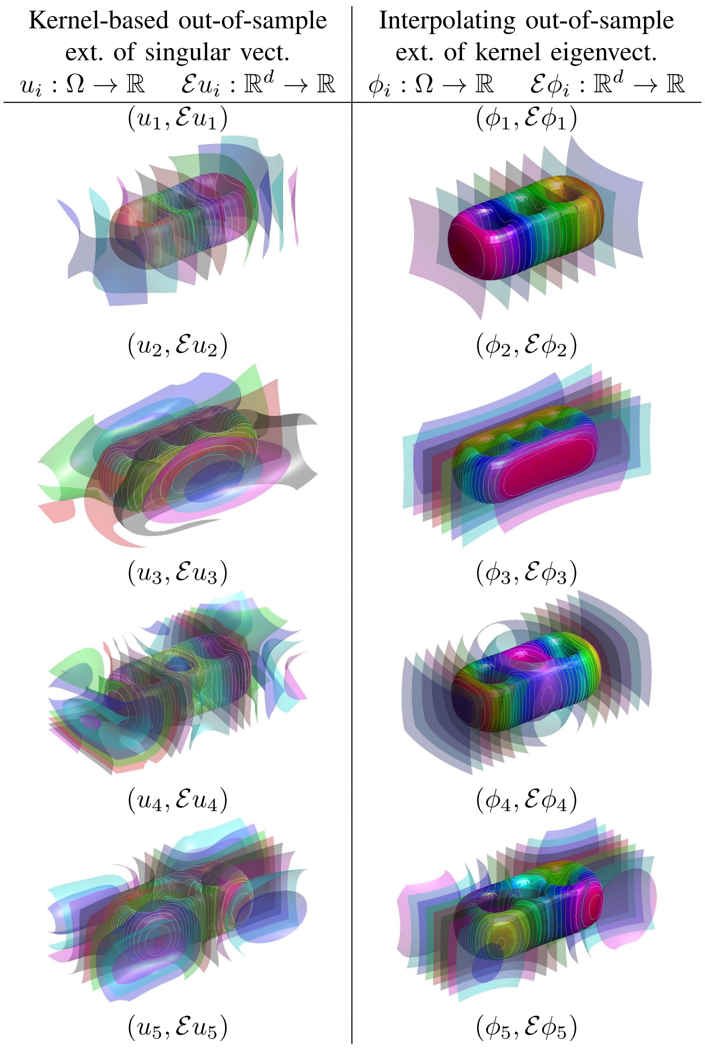

Fig. 2. Color-map and level-sets of the input singular vectors and Laplacian eigenvectors. Iso-surfaces of the kernel-based (Sect. IV) out-of-sample extension of the singular vectors and interpolating out-of-sample extension of the kernel eigenvectors (Sect. IV-B).

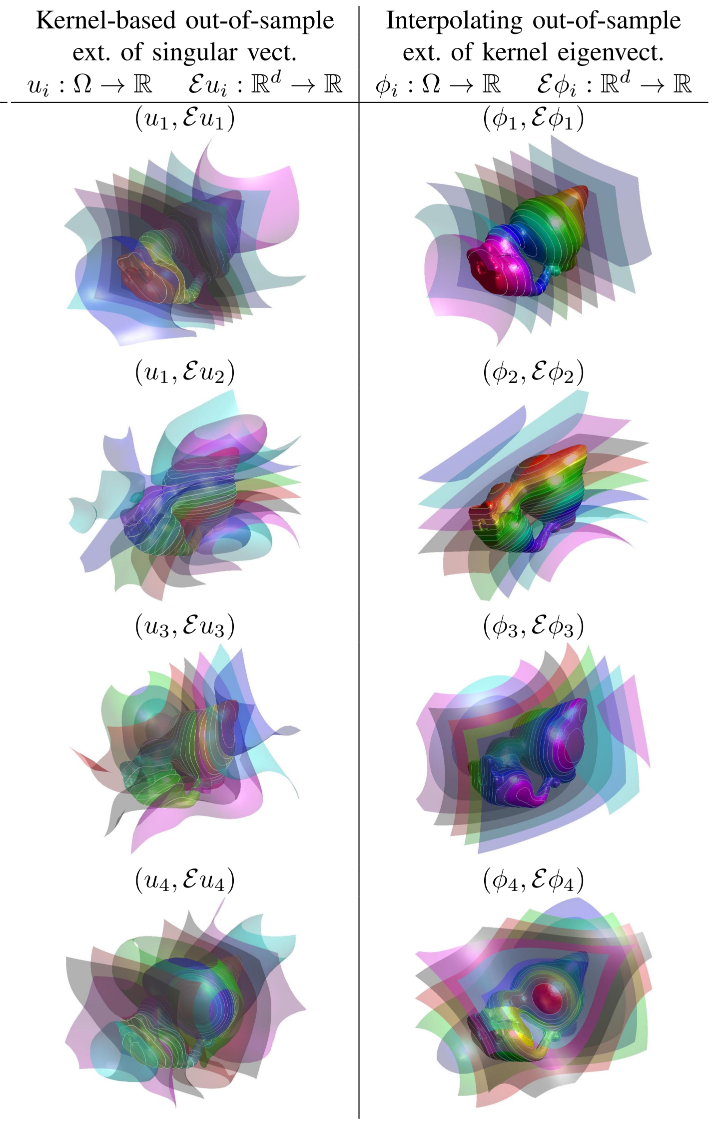

Fig. 3. Color-map and level-sets of the input singular vectors u_i and Laplacian eigenvectors ϕ_i. Iso-surfaces of the kernel-based (Sect. IV) out-of-sample extension $\mathcal{E}u_i$, $\mathcal{E}\phi_i$ of the singular vectors and interpolating out-of-sample extension of the kernel eigenvectors (Sect. IV-B).

In this setting, the error

$$\|\mathcal{L}_K f - \mathcal{L}_K \mathcal{E}\mathbf{f}\|_2 \leq \|K\|_2 \|f - \mathcal{E}\mathbf{f}\|_2 \tag{5}$$

between the data-driven FTs $\mathcal{L}_K f$ and $\mathcal{L}_K \mathcal{E}\mathbf{f}$ is guided by the accuracy of the out-of-sample extension $\mathcal{E}\mathbf{f}$ of f.

In our examples, a signal $f : \Omega \to \mathbb{R}$ is represented through its colour-map and level-sets $\gamma_\alpha := \{\mathbf{p} \in \Omega : f(\mathbf{p}) = \alpha\}$, $\alpha \in \mathbb{R}$. Then, we represent the out-of-sample extension $\mathcal{E}f : \mathbb{R}^d \to \mathbb{R}$ through its colormap and isosurfaces $\Sigma_\alpha := \{\mathbf{p} \in \mathbb{R}^d : \mathcal{E}f(\mathbf{p}) = \alpha\}$; analogously, for the out-of-sample extension $\mathcal{E}(\mathcal{F}f)$ of the discrete generalised FT. The colour map varies the hue component of the hue-saturation-value colour model; the colours begin with red, pass through yellow, green, cyan, blue, and magenta, and return to red.

a) Kernel-based out-of-sample FT: Selecting s seed points over n input points and recalling Eq. (3), we evaluate the $n \times s$ kernel matrix $\mathbf{K} := (K(\mathbf{p}_i, \mathbf{q}_j))_{i=1,\dots,n}^{j=1,\dots,s}$, $K(\mathbf{p}_i, \mathbf{q}_j) := \frac{A(\mathbf{p}_i, \mathbf{q}_j)}{D(\mathbf{p}_i)}$, where $D(\mathbf{p}_i) := \int_\Omega A(\mathbf{p}_i, \mathbf{q}) d\mathbf{q}$ is approximated as $D(\mathbf{p}_i) \approx \sum_{j=1}^{s} A(\mathbf{p}_i, \mathbf{q}_j)$. Then, the *weighted generalised FT* is $\mathbf{F}_\mathbf{K} = \mathbf{K}\mathbf{f} = \mathbf{D}^{-1}\mathbf{A}\mathbf{f}$, where $\mathbf{A} := (A(\mathbf{p}_i, \mathbf{p}_j))_{i,j}$ and $\mathbf{D} := (\mathrm{diag}(D(\mathbf{p}_i)))_{i=1}^{n}$. In particular, $\mathbf{K}\mathbf{1} = \mathbf{1}$, i.e., $\mathbf{1}$ is a singular vector of $\mathbf{K}$ and the corresponding singular value is 1. We compute the *singular value decomposition* [17] (Ch. 2) $\mathbf{U}\mathbf{K}\mathbf{V} = \Gamma$ of the kernel matrix, with orthogonality conditions $\mathbf{U}^\top\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^\top\mathbf{V} = \mathbf{I}$, $\mathbf{U} \in \mathbb{R}^{n \times n}$, $\mathbf{V} \in \mathbb{R}^{s \times s}$. Here, $\Gamma \in \mathbb{R}^{n \times s}$ has $p := \min\{n, s\}$ non-null diagonal entries. According to the relations $\mathbf{K}\mathbf{V} = \mathbf{U}\Gamma$, $\mathbf{K}^\top\mathbf{U} = \mathbf{V}\Gamma^\top$, we rewrite the left and right eigenvectors as (Figs., 2, 3, 1^{st} column)

$$\begin{cases} u_i(\mathbf{p}_l) = \frac{1}{\sigma_i} \sum_{j=1}^{s} K(\mathbf{p}_l, \mathbf{q}_j) v_i(\mathbf{q}_j); \\ v_i(\mathbf{q}_l) = \frac{1}{\sigma_i} \sum_{j=1}^{n} K(\mathbf{p}_j, \mathbf{q}_l) u_i(\mathbf{p}_j); \end{cases} \quad i = 1, \dots, p.$$

Indeed, the functions

$$\begin{cases} u_i(\mathbf{p}) = \frac{1}{\sigma_i} \sum_{j=1}^{s} K(\mathbf{p}, \mathbf{q}_j) v_i(\mathbf{q}_j); \\ v_i(\mathbf{q}) = \frac{1}{\sigma_i} \sum_{j=1}^{n} K(\mathbf{p}_j, \mathbf{q}) u_i(\mathbf{p}_j); \end{cases}$$

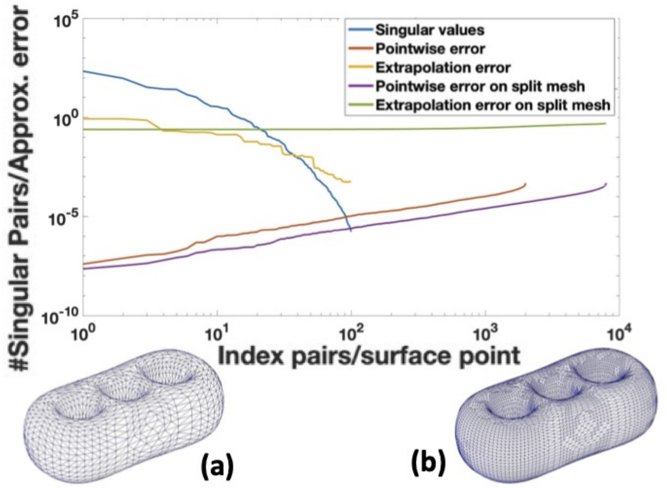

Fig. 4. Statistics of the kernel-based out-of-sample spectral FT on (a) an input and up-sampled domain Ω, represented by a point set $\mathcal{P}$ and its resampling $\mathcal{P}^{split}$. See also Fig. 5.

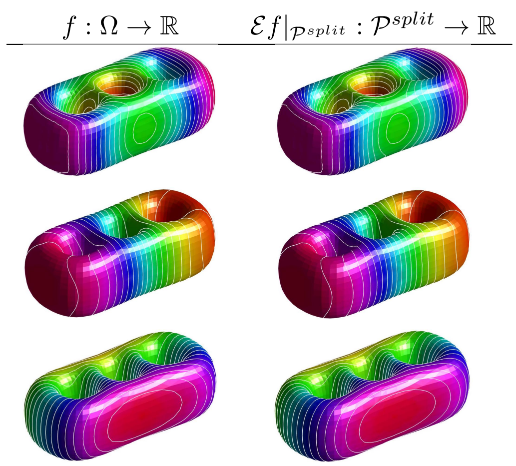

Fig. 5. With reference to Fig. 5, colour map and level sets of an input signal f and its out-of-sample extension $\mathcal{E}f:\mathbb{R}^d\rightarrow\mathbb{R}$ and its restriction to $\mathcal{P}^{split}$, $\mathcal{E}f|_{\mathcal{P}^{split}}:\mathcal{P}^{split}\rightarrow\mathbb{R}$, on an upsampling $\mathcal{P}^{split}$ of Ω. See also Fig. 7.

are the out-of-sample extensions of $\mathbf{u}_i$ and $\mathbf{v}_i$, respectively. Equivalently, $\mathcal{E}(\mathbf{u}_i)=u_i$ and $\mathcal{R}(u_i)=\mathbf{u}_i$. The out-of-sample kernel interpolates the entries of $\mathbf{K}$, i.e.,

$$\begin{aligned}K(\mathbf{p}_a,\mathbf{q}_b)&=\sum_{i=1}^{p}\sigma_i u_i(\mathbf{p}_a)v_i(\mathbf{q}_b)\\&=\mathbf{e}_a^\top\sum_{i=1}^{p}\sigma_i\mathbf{u}_i^\top\mathbf{v}_i\mathbf{e}_b.=K(a,b).\end{aligned}$$

Furthermore, the functions $(u_i(\cdot))_{i=1}^n$ are orthonormal with respect to the pointwise scalar product at $\mathcal{P}$, i.e.,

$$\langle u_a,u_b\rangle:=\sum_{i=1}^{n}u_a(\mathbf{p}_i)u_b(\mathbf{p}_i)=\mathbf{u}_a^\top\mathbf{u}_b=\delta_{ab},$$

as a consequence of the orthogonality of $\mathbf{U}$.

Through the out-of-sample extensions of the left and right-hand-side singular eigenvectors, we compute the out-of-sample extensions of the FT. Representing the dis-

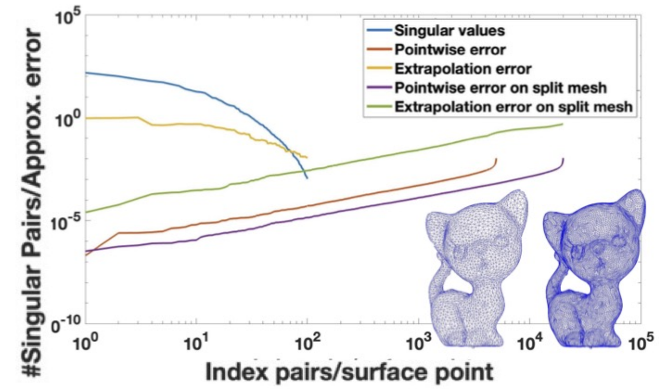

Fig. 6. Statistics of the kernel-based out-of-sample spectral FT on (a) an input and up-sampled domain Ω, represented by a point set $\mathcal{P}$ and its resampling $\mathcal{P}^{split}$. See also Fig. 7.

crete signal $\mathbf{f}=\sum_{i=1}^n\langle\mathbf{f},\mathbf{u}_i\rangle_2\mathbf{u}_i$ as a linear combination of the singular vectors $\mathcal{U}:=(\mathbf{u}_i)_{i=1}^n$, the function $(\mathcal{E}\mathbf{f})(\mathbf{p}):=\sum_{i=1}^n\langle\mathbf{f},\mathbf{u}_i\rangle_{\mathbf{D}}u_i(\mathbf{p})$ is the out-of-sample extension of $\mathbf{f}$, i.e.,

$$(\mathcal{E}\mathbf{f})(\mathbf{p}_j)=\sum_{i=1}^{n}\langle\mathbf{f},\mathbf{u}_i\rangle_2u_i(\mathbf{p}_j)=\sum_{i=1}^{n}\langle\mathbf{f},\mathbf{u}_i\rangle_2\mathbf{u}_i=f(\mathbf{p}_j).$$

Then, the *out-of-sample spectral membership functions* are

$$A(\mathbf{p}_j,\mathbf{q}):=K(\mathbf{p}_j,\mathbf{q}):=\sum_{i=1}^{p}\sigma_iu_i(\mathbf{p}_j)v_i(\mathbf{q}),$$

and the *out-of-sample FT* is

$$\mathcal{E}(\mathbf{L}_K\mathbf{f})(\mathbf{p})=\sum_{i=1}^{p}\sigma_i\langle\mathbf{f},\mathbf{u}_i\rangle_2\mathbf{v}_i(\mathbf{p}).\tag{6}$$

Eq. (6) can also be derived from Eq. (2) as follows

$$\begin{aligned}&\int_\Omega K(\mathbf{p},\mathbf{q})f(\mathbf{q})d\mathbf{q}=\sum_{i=1}^{p}\sigma_iv_i(\mathbf{p})\int_\Omega f(\mathbf{q})u_i(\mathbf{q})d\mathbf{q}\\&=\sum_{i=1}^{p}\sigma_i\langle f,u_i\rangle_2v_i(\mathbf{p})=\sum_{i=1}^{p}\sigma_i\langle\mathbf{f},\mathbf{u}_i\rangle_2\mathbf{v}_i(\mathbf{p})=\mathcal{E}(\mathbf{L}_K\mathbf{f})(\mathbf{p}).\end{aligned}$$

From the relation $\mathbf{K}:=\mathbf{D}^{-1}\mathbf{A}$, we have that $\mathbf{A}=\mathbf{DK}$ and the out-of-sample membership functions are

$$A(\mathbf{p}_j,\mathbf{q})=d(j)K(\mathbf{p}_j,\mathbf{q})=d(j)\sum_{i=1}^{n}\sigma_iu_i(\mathbf{p}_j)v_i(\mathbf{q}).$$

b) Interpolating out-of-sample FT: Selecting $s=n$ and $\mathbf{p}_i=\mathbf{q}_i$, $\forall i$, the singular value decomposition of the symmetric kernel matrix $\mathbf{K}\in\mathbb{R}^{n\times n}$ is replaced with its generalised eigendecomposition $\mathbf{K}\mathbf{x}_i=\lambda_i\mathbf{D}\mathbf{x}_i$, with $\mathbf{u}_i=\mathbf{v}_i=\mathbf{x}_i$ and $\sigma_i=\lambda_i$, $\mathbf{x}_i^\top\mathbf{D}\mathbf{x}_j=\delta_{ij}$, $\forall i,j$. In this case, the out-of-sample extension of the Laplacian eigenvectors minimises the energy

$$\mathcal{E}:=\sum_{i=1}^{n}\left[K(\mathbf{p},\mathbf{p}_i)-\sum_{l=1}^{n}\lambda_ld(l)\phi_l(\mathbf{p})\phi_l(\mathbf{p}_i)\right]^2.$$

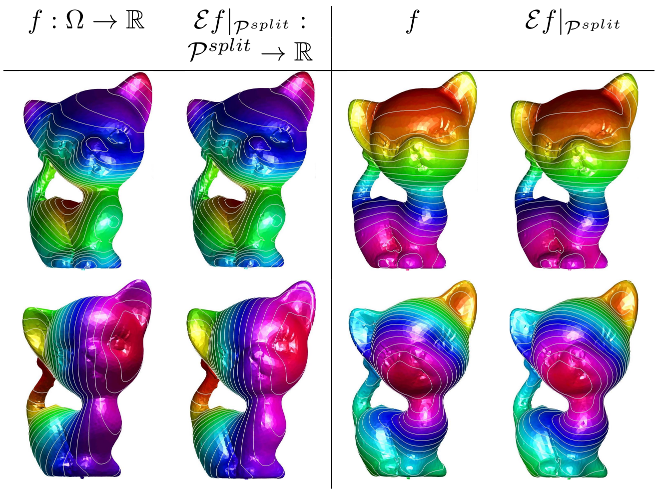

Fig. 7. With reference to Fig. 6, colour map and level sets of the input sognal f and its out-of-sample extension $\mathcal{E}f:\mathbb{R}^d\rightarrow\mathbb{R}$ and its restriction to $\mathcal{P}^{split}$, $\mathcal{E}f|_{\mathcal{P}^{split}}:\mathcal{P}^{split}\rightarrow\mathbb{R}$, on an upsampling $\mathcal{P}^{split}$ of Ω.

Firstly, we compute its derivatives

$$\partial_{\phi_k}\mathcal{E}=2\sum_{i=1}^{n}\left[K(\mathbf{p},\mathbf{p}_i)-\sum_{l=1}^{n}\lambda_l d(i)\phi_l(\mathbf{p})\phi_l(\mathbf{p}_i)\right]\lambda_k\phi_k(\mathbf{p}_i)$$
$$=2\lambda_k\sum_{i=1}^{n}K(\mathbf{p},\mathbf{p}_i)\phi_k(\mathbf{p}_i)-2\sum_{i,l=1}^{n}\lambda_l d(l)\phi_l(\mathbf{p})\phi_l(\mathbf{p}_i)\lambda_k\phi_k(\mathbf{p}_i).$$

Recalling that the Laplacian eigenvectors are orthonormal

$$\sum_{i=1}^{n}\phi_l(\mathbf{p}_i)\phi_k(\mathbf{p}_i)d(i)=\mathbf{x}_k^\top\mathbf{D}\mathbf{x}_l=\delta_{kl},$$

the condition $\partial_{\phi_k}\mathcal{E}=0$ reduces to the linear conditions $\sum_{i=1}^{n}K(\mathbf{p},\mathbf{p}_i)d(i)\phi_k(\mathbf{p}_i)-\phi_k(\mathbf{p})\lambda_k=0$, $\forall k$, i.e., (Figs. 2, 3 2^{nd} column)

$$\phi_k(\mathbf{p})=\frac{1}{\lambda_k}\sum_{i=1}^{n}d(i)K(\mathbf{p},\mathbf{p}_i)x_k(i).$$

In this case, the *out-of-sample FT* is (Figs. 4, 5)

$$\mathcal{E}(\mathbf{L}_K\mathbf{f})(\mathbf{q})=\sum_{i=1}^{n}\lambda_i d(i)\langle\mathbf{f},\mathbf{x}_i\rangle_2\phi_i(\mathbf{q}).$$

The approximation accuracy and convergence of $\mathcal{E}(\mathbf{L_K f})$ to $\mathcal{L}_K f$ will be discussed in Sect. V.

c) Experimental tests: For a given sampling $\mathcal{P}$ of Ω, we compute the generalised eigenvectors of $(\mathbf{K},\mathbf{D})$, where $\mathbf{K}$ is the Gram matrix $\mathbf{K}:=(K(\mathbf{p}_i,\mathbf{p}_j))_{i,j=1}^n$, induced by the Gaussian kernel $K(\mathbf{p},\mathbf{q}):=\exp(\|\mathbf{p}-\mathbf{q}\|_2/\sigma)$, σ is the width parameter, and $\mathbf{D}$ is the diagonal matrix whose entries are the sum of the rows of $\mathbf{K}$. The behaviour of the colour map, the shape and the distribution of the level sets show the analogous behaviour of the input and out-of-sample extension of the eigenfunctions, as confirmed by the low approximation error ϵ_∞. The out-of-sample extension has a generally smoother behaviour as a matter of its representation in terms of a smooth kernel (e.g., the Gaussian kernel).

To analyse the properties of the proposed approach (Figs., 4, 5; Figs., 6, 7), we plot the *singular values* of the Gram matrix $\mathbf{K}$ induced by the selected kernel. We select three samples $\mathcal{P}_{50K}$, $\mathcal{P}_{100K}$, and $\mathcal{P}_{150K}$ of a 3D domain Ω, represented as a triangle mesh. The samplings $\mathcal{P}_{100K}$, $\mathcal{P}_{150K}$ are achieved by splitting each triangle of $\mathcal{P}_{50K}$ into four sub-triangles by joining the midpoint of each edge; in this way, the new mesh has n_V+3n_T vertices, where n_V and n_T are the number of vertices and triangles of the input mesh, respectively. We consider the discrete signal $\mathbf{f}_{150}:=(f(\mathbf{p}_i))_{i=1}^{150K}$, achieved by sampling $f:\Omega\rightarrow\mathbb{R}$ at the point set $\mathcal{P}_{150K}$ of the domain Ω at the highest resolution with 150K samples. On these data sets, we evaluate the (a) *extrapolation* and (b) *maximum errors*

$$(a)\ \epsilon_i:=|f(\mathbf{p}_i)-\mathcal{E}\mathbf{f}_k(\mathbf{p}_i)|,\ k:=50,100,\ (b)\ \epsilon_\infty:=\max_i\{\epsilon_i\},$$

between the input f and the out-of-sample extension $\mathcal{E}\mathbf{f}$ evaluated at $\mathcal{P}_{50K}$, $\mathcal{P}_{100K}$, and $\mathcal{P}_{150K}$.

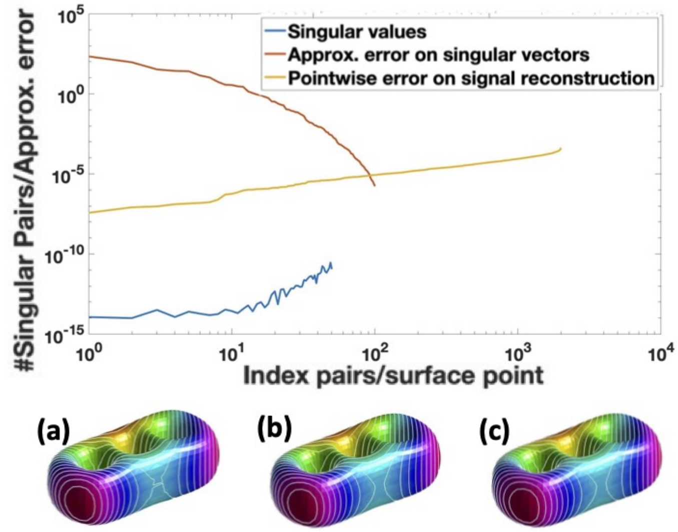

Fig. 8. Given a signal $f:\Omega\rightarrow\mathbb{R}$, we compute (a) its FT $\mathcal{F}f$, (b) its FT on an up-sampling $\mathcal{P}^{split}$ of Ω, and (c) the kernel-based out-of-sample FT $\mathcal{E}(\mathcal{F}f)$. We also report statistics on the singular values' distribution, approximation, and pointwise errors.

For the results in Figs., 4, 5, 6, 7, the singular values rapidly decrease to zero, the order of magnitude of the pointwise error between f at $\mathcal{P}_{150K}$ and $(\mathcal{E}\mathbf{f}_{50K})|_{\mathcal{P}_{150K}}$, $(\mathcal{E}\mathbf{f}_{100K})|_{\mathcal{P}_{150K}}$ remains lower than 1%. The order of magnitude of the approximation error of the kernel-based approach remains lower than 10^{-2}. According to Eq. (5) and assuming that the kernel matrix $\mathbf{K}$ is well conditioned, a low error in the approximation of the signal corresponds to an approximation of the corresponding generalised FT of the same order of magnitude.

Analogously to the previous tests, we compute the FT $\mathcal{F}\mathbf{f}_{150K}=\mathbf{F}_{150K}:=(F_i)_{i=1}^{150K}$ of the input signal $f:\Omega\rightarrow\mathbb{R}$ on the sampling $\mathcal{P}_{150K}$ of the domain Ω at the highest resolution with 150K samples: we consider $\mathbf{F}_{150K}$ as our ground-truth. We also compute the out-of-sample FTs $\mathcal{E}\mathbf{F}_{50K}$, $\mathcal{E}\mathbf{F}_{100K}:\mathbb{R}^3\rightarrow\mathbb{R}$ of the signal at the lower resolutions of Ω with 50K and 100K samples and evaluate these two out-of-sample FTs at $\mathcal{P}_{150K}$. Then, we measure the pointwise

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2023.3326657

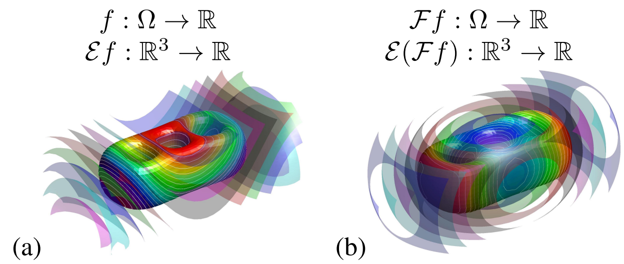

Fig. 9. (a) Color map and level sets of an input signal $f:\Omega\rightarrow\mathbb{R}$ and iso-surfaces of its out-of-sample extension $\mathcal{E}f:\mathbb{R}^d\rightarrow\mathbb{R}$, (b) colour map and level sets of the FT $\mathcal{F}f:\Omega\rightarrow\mathbb{R}$ and iso-surfaces of its kernel-based out-of-sample extension $\mathcal{E}(\mathcal{F}f):\mathbb{R}^d\rightarrow\mathbb{R}$.

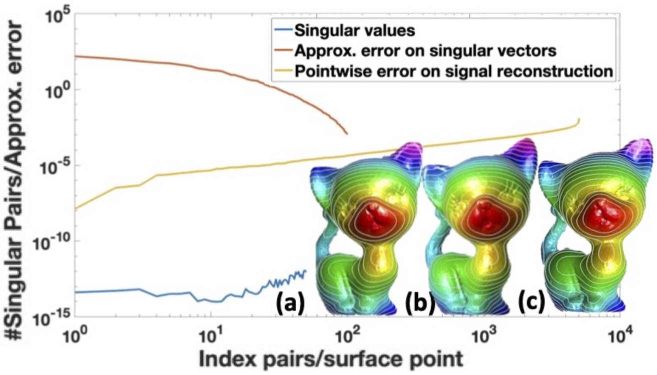

Fig. 10. Given a signal $f:\Omega\rightarrow\mathbb{R}$, we compute (a) its FT $\mathcal{F}f$, (b) its FT on an up-sampling $\mathcal{P}^{split}$ of Ω, and (c) the kernel-based out-of-sample FT $\mathcal{E}(\mathcal{F}f)$. We also report statistics on the singular values' distribution, approximation, and pointwise errors.

and ℓ_∞ errors between the resulting FTs $(\mathcal{E}\mathbf{F}_{50K})|_{\mathcal{P}_{150K}}$, $(\mathcal{E}\mathbf{F}_{150K})|_{\mathcal{P}_{150K}}$ and the ground-truth $\mathbf{F}_{150K}$, thus estimating the extrapolation capabilities of the out-of-sample FT. In this case, the order of magnitude of the pointwise error remains lower than 1%. This behaviour is consistent with the upper bound in (5) and the accuracy of the out-of-sample extension of the input signal.

IV. Out-of-sample data-driven Fuzzy transform

Firstly (Sect. IV-A), we express the data-driven FT in terms of the spectral filter and the area/volume of the input domain, thus avoiding the computation of integrals. Then, we introduce the out-of-sample extension of the data-driven FT (Sect. IV-B).

A. Data-driven Fuzzy transform: unified representation

We characterise the data-driven FTs [1], [2] by explicitly deriving the normalisation factor in terms of the input filter and the area/volume of the input domain. Recalling that the Laplacian eigenfunctions are orthonormal and $\phi_0=1$ is the eigenfunction associated with the null eigenvalue λ_0, we get the relation

$$\int_\Omega \phi_n(\mathbf{p})d\mathbf{p}=\begin{cases}0 & n\neq 0;\\ |\Omega| & n=0;\end{cases}$$

Choosing the corresponding membership functions $A_{\mathbf{p}}:\Omega\times\Omega\rightarrow\mathbb{R}$, centred at $\mathbf{p}$ and defined as $A_{\mathbf{p}}(\mathbf{q}):=K_\varphi(\mathbf{p},\mathbf{q})$, we get that

$$\begin{aligned}D(\mathbf{p}):&=\int_\Omega A_{\mathbf{p}}(\mathbf{q})d\mathbf{q}=\int_\Omega K_\varphi(\mathbf{p},\mathbf{q})d\mathbf{q}\\&=\sum_{n=0}^{+\infty}\varphi(\lambda_n)\phi(\mathbf{p})\int_\Omega\phi_n(\mathbf{q})d\mathbf{q}=\varphi(0)|\Omega|.\end{aligned}\tag{7}$$

Let us introduce the *normalised filter* $\widetilde{\varphi}(s):=\frac{\varphi(s)}{\varphi(0)|\Omega|}$, which is still positive and belongs to $\mathcal{L}^2(\mathbb{R}^+)$. Then, the *spectral membership function* centred at $\mathbf{p}$ is defined by the action of the data-driven FT on $\delta_{\mathbf{p}}$, i.e.,

$$\begin{aligned}\psi:&=\mathcal{L}_{K_{\widetilde{\varphi}}}\delta_{\mathbf{p}}=\widetilde{\varphi}(\Delta)\delta_{\mathbf{p}}\\&=K_{\widetilde{\varphi}}(\mathbf{p},\cdot)=\sum_{n=0}^{+\infty}\frac{\varphi(\lambda_n)}{\varphi(0)|\Omega|}\phi_n(\mathbf{p})\phi_n=\frac{1}{\varphi(0)|\Omega|}\mathcal{L}_{K_\varphi}\delta_{\mathbf{p}}.\end{aligned}$$

From Eq. (7), we rewrite Eq. (2) as

$$\begin{aligned}F(\mathbf{p})&=\frac{\int_\Omega A(\mathbf{p},\mathbf{q})f(\mathbf{q})d\mathbf{q}}{D(\mathbf{p})}=\frac{1}{\varphi(0)|\Omega|}\sum_{n=0}^{+\infty}\varphi(\lambda_n)\langle f,\phi_n\rangle_2\phi_n(\mathbf{p})\\&=\int_\Omega K_{\widetilde{\varphi}}(\mathbf{p},\mathbf{q})f(\mathbf{q})d\mathbf{q}=(\mathcal{L}_{K_{\widetilde{\varphi}}}f)(\mathbf{p}).\end{aligned}\tag{8}$$

The data-driven FT $\mathcal{L}_{K_{\widetilde{\varphi}}}$ is *linear* and *continuous* $\|\mathcal{L}_{K_{\widetilde{\varphi}}}f\|_2\leq\|K_{\widetilde{\varphi}}\|_2\|f\|_2\leq\frac{\|\varphi\|_2}{\varphi(0)|\Omega|}\|f\|_2$, *self-adjoint* $\mathcal{L}_{K_{\widetilde{\varphi}}}f=\frac{1}{\varphi(0)|\Omega|}\mathcal{L}_{K_\varphi}f$, and *positive-definite*, according to the relation

$$\langle\mathcal{L}_{K_{\widetilde{\varphi}}}f,f\rangle_2=\sum_{n=0}^{+\infty}\frac{1}{\varphi(0)|\Omega|}|\langle f,\phi_n\rangle_2|^2=\frac{1}{\varphi(0)|\Omega|}\|f\|_2^2\geq 0.$$

The data-driven FT is *injective* if and only if φ is strictly positive. In fact, $\mathcal{L}_{K_{\widetilde{\varphi}}}f=\mathcal{L}_{K_{\widetilde{\varphi}}}g$ if and only if

$$\sum_{n=0}^{+\infty}\widetilde{\varphi}(\lambda_n)\langle f-g,\phi_n\rangle_2=0\iff\langle f-g,\phi_n\rangle_2=0,\ \forall n,$$

i.e., $f=g$. Noting that

$$\mathcal{L}^{-1}_{K_{\widetilde{\varphi}}}f=\mathcal{L}_{K_{1/\widetilde{\varphi}}}f=\varphi(0)|\Omega|\mathcal{L}_{K_{1/\varphi}}f=\varphi(0)|\Omega|\mathcal{L}^{-1}_{K_\varphi}f,$$

the *inverse data-driven FT* is the integral operator induced by $K_{1/\widetilde{\varphi}}$. Equivalently, $g:=\mathcal{L}_{K_{\widetilde{\varphi}}}f=\widetilde{\varphi}(\Delta)f$ if and only if $f=\widetilde{\varphi}^{-1}(\Delta)g$. If $\widetilde{\varphi}(\Delta)$ is singular (i.e., $\varphi(\lambda_i)=0$, for some i), then we consider its peudoinverse $f=\widetilde{\varphi}^\dagger(\Delta)g$.

Discrete data-driven FT: Assuming that $\mathbf{W}$ is a weight matrix (e.g., the adjacency matrix), whose entry (i,j) is a strictly positive weight associated with the corresponding edge, the *Laplacian matrix* [18] is defined as $\tilde{\mathbf{L}}:=\mathbf{D}^{-1}\mathbf{L}$, where $\mathbf{L}:=\mathbf{D}-\mathbf{W}$ and $\mathbf{D}$ is the diagonal matrix whose entries are the sum of the rows of $\mathbf{W}$. The Laplacian matrix is $\mathbf{D}$-adjoint with respect to the scalar product $\langle\mathbf{f},\mathbf{g}\rangle_{\mathbf{D}}:=\mathbf{f}^\top\mathbf{D}\mathbf{g}$, $\mathbf{f}:=(f(\mathbf{p}_i))_{i=1}^n$, i.e., $\langle\tilde{\mathbf{L}}\mathbf{f},\mathbf{g}\rangle_{\mathbf{D}}=\langle\mathbf{f},\tilde{\mathbf{L}}\mathbf{g}\rangle_{\mathbf{D}}$, and its spectral decomposition is $\mathbf{L}\mathbf{X}=\mathbf{D}\mathbf{X}\Lambda$, $\mathbf{X}^\top\mathbf{D}\mathbf{X}=\mathbf{I}$, where $\mathbf{X}:=[\mathbf{x}_1,\dots,\mathbf{x}_n]$ is the eigenvectors' matrix and Λ is the diagonal matrix of the eigenvalues $(\lambda_i)_{i=1}^n$. The *discrete membership functions* are represented in matrix form as

$$\mathbf{K}_{\widetilde{\varphi}}:=\mathbf{X}\widetilde{\varphi}(\Gamma)\mathbf{X}^\top\mathbf{D}=\frac{1}{\varphi(0)|\Omega|}\mathbf{X}\varphi(\Gamma)\mathbf{X}^\top\mathbf{D},$$

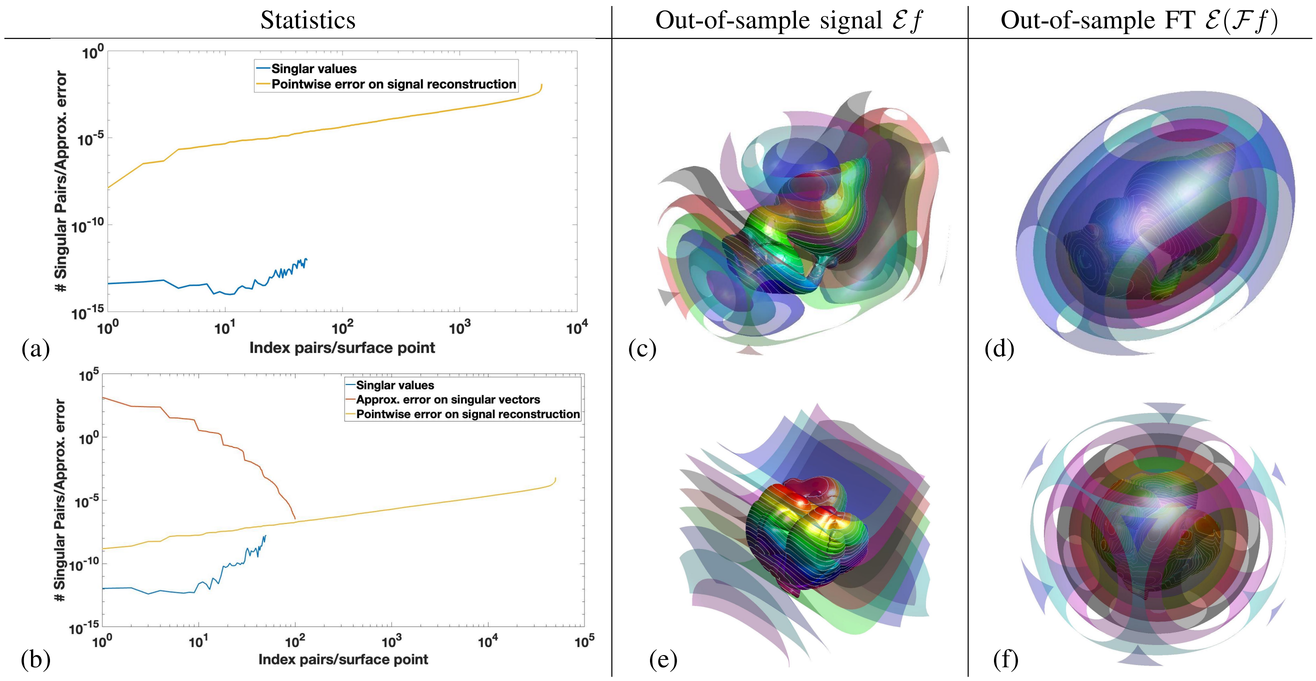

Fig. 11. (c,e) Color map and level sets of an input signal $f : \Omega \to \mathbb{R}$ and iso-surfaces of its kernel-based out-of-sample extension $\mathcal{E}f : \mathbb{R}^d \to \mathbb{R}$, (d,f) colour map and level sets of the FT $\mathcal{F}f$ and iso-surfaces of its out-of-sample extension $\mathcal{E}(\mathcal{F}f) : \mathbb{R}^d \to \mathbb{R}$.

and the *discrete data-driven FT* is

$$\mathbf{L}_{\mathbf{K}_{\widetilde{\varphi}}}\mathbf{f} =_{\text{Eq.}(8)} \sum_{i=1}^{n} \widetilde{\varphi}(\lambda_i)\langle \mathbf{f}, \mathbf{x}_i\rangle_{\mathbf{D}}\mathbf{x}_i = \widetilde{\varphi}(\widetilde{\mathbf{L}})\mathbf{f},$$

where $\widetilde{\varphi}(\widetilde{\mathbf{L}})$ is the filtered Laplacian matrix.

B. Out-of-sample data-driven Fuzzy transform

For each Laplacian eigenvector $\mathbf{x}_i := (x_i(j))_{j=1}^n$. we compute an *out-of-sample extension* $\psi_i : \Omega \to \mathbb{R}$, $\psi_i(\mathbf{p}_j) = x_i(j)$, through an implicit approximation $\psi_i(\mathbf{p}) := \sum_{j=1}^n \alpha_j^{(i)}\psi(\|\mathbf{p}-\mathbf{p}_j\|_2)$, $\alpha^{(i)} := (\alpha_j^{(i)})_{j=1}^n$, with radial basis functions (RBFs) $\Psi(\mathbf{p},\mathbf{p}_i) := \psi(\|\mathbf{p}-\mathbf{p}_i\|_2)$, $i = 1, \dots, n$, induced by the generating function $\psi : \mathbb{R} \to \mathbb{R}$. Imposing the interpolating conditions

$$x_i(a) = \psi_i(\mathbf{p}_a) := \sum_{j=1}^{n} \alpha_j^{(i)}\psi(\|\mathbf{p}_a - \mathbf{p}_j\|_2) = \sum_{j=1}^{n} \Psi(a,j)\alpha_j^{(i)},$$

$\Psi(i,j) := \psi(\|\mathbf{p}_i - \mathbf{p}_j\|_2) = \Psi(\mathbf{p}_i, \mathbf{p}_j)$, we get the $n \times n$ linear system $\mathbf{\Psi}\alpha^{(i)} = \mathbf{x}_i$. The coefficient matrix is computed only once and applied to evaluate all the out-of-sample extensions of the Laplacian eigenvectors. In particular, $\mathcal{E}(\mathbf{x}_i) = \psi_i$ and $\mathcal{R}(\psi_i) = \mathbf{x}_i$. Rewriting the discrete signal $\mathbf{f} := (f_i)_{i=1}^n$ in terms of the Laplacian eigenvectors as $\mathbf{f} = \sum_{i=1}^n \langle \mathbf{f}, \mathbf{x}_i\rangle_{\mathbf{D}}\mathbf{x}_i$, the function

$$f(\mathbf{p}) := \sum_{i=1}^{n} \langle \mathbf{f}, \mathbf{x}_i\rangle_{\mathbf{D}}\psi_i(\mathbf{p}) \tag{9}$$

is the out-of-sample extension of $\mathbf{f}$; in fact,

$$(f(\mathbf{p}_j))_{j=1}^n = \left[\sum_{i=1}^{n} \langle \mathbf{f}, \mathbf{x}_i\rangle_{\mathbf{D}}\psi_i(\mathbf{p}_j)\right]_{j=1}^n = \sum_{i=1}^{n} \langle \mathbf{f}, \mathbf{x}_i\rangle_{\mathbf{D}}\mathbf{x}_i = \mathbf{f}.$$

Once we have computed the out-of-sample extension of the Laplacian eigenvectors, the spectral membership function at $\mathbf{p}_j$ is evaluated as $K_{\widetilde{\varphi}}(\mathbf{p}_j, \mathbf{q}) := \sum_{i=1}^n \widetilde{\varphi}(\lambda_i)\psi_i(\mathbf{p}_j)\psi_i(\mathbf{q})$, $\mathbf{p} \in \Omega$, and the corresponding *out-of-sample spectral FT* is (Figs. 8, 9, 10, 11)

$$(\mathcal{L}_{K_{\widetilde{\varphi}}}f)(\mathbf{p}) = \langle K_{\widetilde{\varphi}}(\mathbf{p}, \cdot), f\rangle_2 = \sum_{i=1}^{n} \widetilde{\varphi}(\lambda_i)\langle f, \psi_i\rangle_2\psi_i(\mathbf{p}), \tag{10}$$

which is analogous to Eq. (4). In a similar way,

$$\begin{aligned} A_i(\mathbf{p}) &= \int_{\Omega} K_{\widetilde{\varphi}}(\mathbf{p},\mathbf{q})d\mathbf{q} = \sum_{i=1}^{n} \widetilde{\varphi}(\lambda_i)\psi_i(\mathbf{p}) \int_{\Omega} \psi_i(\mathbf{q})d\mathbf{q} \\ &= \sum_{i=1}^{n} \widetilde{\varphi}(\lambda_i)\psi_i(\mathbf{p}) \int_{\Omega} \psi_i(\mathbf{q})d\mathbf{q} \\ &= \sum_{i=1}^{n} \widetilde{\varphi}(\lambda_i)\psi_i(\mathbf{p}) \int_{\Omega} \sum_{j=1}^{n} \alpha_j^{(i)}\psi(\|\mathbf{q}-\mathbf{p}_j\|_2) \\ &= \sum_{i,j=1}^{n} \alpha_j^{(i)}\widetilde{\varphi}(\lambda_i)\psi_i(\mathbf{p}) \int_{\Omega} \psi(\|\mathbf{q}-\mathbf{p}_j\|_2)d\mathbf{q}, \end{aligned}$$

where the last integral is easily computed as the generating function ψ is 1D.

Computation: For the out-of-sample extension with RBFs, the evaluation of Eq. (10) requires computing the component

$$\langle f, \psi_i\rangle_2 = \langle f, \sum_{j=1}^{n} \alpha_j^{(i)}\psi(\|\mathbf{p}-\mathbf{p}_i\|_2)\rangle_2$$

$$= \sum_{i=1}^{n} \alpha_j^{(i)}\langle f, \psi(\|\mathbf{p}-\mathbf{p}_i\|_2)\rangle_2 = \sum_{i=1}^{n} \alpha_j^{(i)} \int_{\Omega} f(\mathbf{p})\psi(\|\mathbf{p}-\mathbf{p}_i\|_2)d\mathbf{p},$$

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2023.3326657

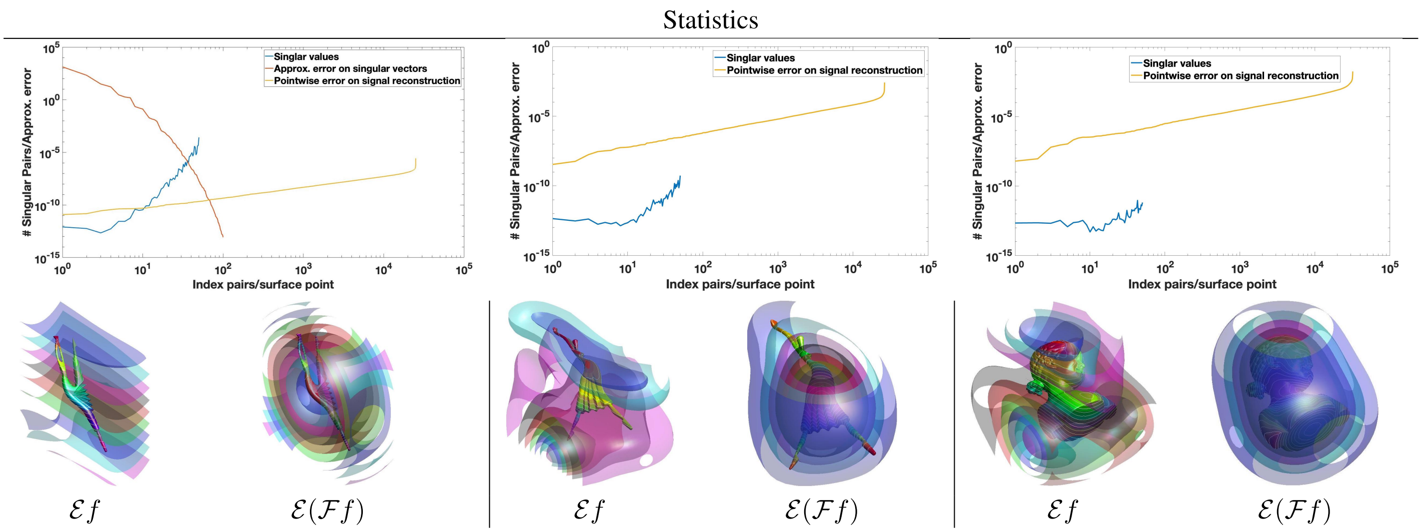

Fig. 12. Kernel-based out-of-sample extension $\mathcal{E}f:\mathbb{R}^d\rightarrow\mathbb{R}$ of the input signal and its kernel-based out-of-sample extension $\mathcal{E}(\mathcal{F}f):\mathbb{R}^d\rightarrow\mathbb{R}$ of the FT on 3D domains.

through quadrature rules with RBFs. Increasing the sampling density of Ω (Fig. 13), the behaviour of the out-of-sample extensions of the input signal and the corresponding generalised FT remain almost unchanged and coherent in terms of the variation of the colourmap and distribution/shape of the iso-surfaces. Furthermore, the maximum variation of $\mathcal{E}f$, $\mathcal{F}f$, and $\mathcal{E}(\mathcal{F}f)$ on the different samplings is lower than 10^{-2}.

V. Approximated out-of-sample data-driven FT

Since the out-of-sample extensions $(\psi_i)_{i=1}^n$ are not orthonormal, we introduce a pseudoscalar product that approximates the $\mathcal{L}^2(\Omega)$ product and makes the out-of-sample eigenfunctions orthonormal (Sect. V-A). This pseudoscalar product also induces an approximation of the out-of-sample FT that is independent of the evaluation of the integral $\int_\Omega f(\mathbf{p})\psi(\|\mathbf{p}-\mathbf{p}_i\|_2)d\mathbf{p}$ and converges to (10), as the sampling density of Ω increases. The *approximated out-of-sample data-driven FT* still satisfies the main properties of the FT, such as linearity and continuity, self-adjointness and eigensystem, surjectivity, convergence and fast computation (Sect. V-B).

A. Pseudoscalar product

Given $f,g\in\mathcal{C}^0(\Omega)$ and $\mathcal{P}:=\{\mathbf{p}_i\}_{i=1}^n$ samples of Ω, let us consider the *pseudoscalar product*

$$\langle f,g\rangle_\star:=\langle \mathbf{f},\mathbf{g}\rangle_{\mathbf{D}}=\mathbf{f}^\top\mathbf{D}\mathbf{g}, \tag{11}$$

where $\mathbf{f}:=(f(\mathbf{p}_i))_{i=1}^n$, $\mathbf{g}:=(g(\mathbf{p}_i))_{i=1}^n$, and $\mathbf{D}$ is a positive definite matrix (e.g., $\mathbf{D}:=\mathbf{I}$ or the mass matrix in Sect. IV-A). Since $\|\cdot\|_{\mathbf{D}}$ is a norm, the pseudo norm

- satisfies the *triangle inequality*: $\|f+g\|_\star\leq\|f\|_\star+\|g\|_\star$;
- is *absolutely homogeneous*: $\|\alpha f\|_\star=|\alpha|\|f\|_\star$ and satisfies the *scalar multiplication*: $\langle\alpha f,\beta g\rangle_\star=\alpha\beta\langle f,g\rangle_\star$, $\alpha,\beta\in\mathbb{R}$;
- is *positive* but does not satisfy the nullity condition; $\|f\|_\star\geq 0$ and $\|f\|_\star=0$ implies only that $f(\mathbf{p}_i)=0$, $i=1,\dots,n$;
- is *distributive* $\langle f,(g+h)\rangle_\star=\langle f,g\rangle_\star+\langle f,h\rangle_\star$ and *commutative*: $\langle f,g\rangle_\star=\langle g,f\rangle_\star$;
- is *convergent*: increasing the sampling density of Ω, we get that the pseudo inner product

$$\mathbf{f}^\top\mathbf{D}\mathbf{g}=\sum_{i,j=1}^n D(i,j)f(\mathbf{p}_i)g(\mathbf{p}_j)=\langle f,g\rangle_2$$

provides an increasingly more accurate approximation of the $\mathcal{L}^2(\Omega)$ scalar product.

Furthermore, the function $f_{\text{approx}}:=\sum_{i=1}^n\langle\mathbf{f},\mathbf{x}_i\rangle_{\mathbf{D}}\phi_i$ is the best least-squares approximation of f in the space $\mathcal{S}:=\text{span}\{\phi_i\}_{i=1}^n$ with respect to the pseudo-norm $\|\cdot\|_\star$; in fact, the minimum of the error

$$\left\|f-\sum_{i=1}^n\alpha_i\phi_i\right\|_\star^2=\left\|\mathbf{f}-\sum_{i=1}^n\alpha_i\mathbf{x}_i\right\|_{\mathbf{D}}^2=\sum_{i=1}^n|\alpha_i-\langle\mathbf{f},\mathbf{x}_i\rangle_{\mathbf{D}}|^2,$$

is achieved for $\alpha_i:=\langle\mathbf{f},\mathbf{x}_i\rangle_{\mathbf{D}}$. Indeed, among all the interpolating functions of f the function f_{approx} is unique and minimises the $\|\cdot\|_\star$ norm.

Comparison between generalised and out-of-sample extension of the FT: Let us now estimate the difference between the generalised FT $\mathcal{F}f$ and the out-of-sample extension of the discrete FT $\mathcal{E}(\mathbf{L}_{\mathbf{K}}\mathbf{f})$. From the definition of the generalised FT and the spectral representation of the our-of-sample kernel, we get that

$$\mathcal{F}f(\mathbf{p})=\int_\Omega\sum_{i=1}^n\lambda_i\phi_i(\mathbf{q})\phi_i(\mathbf{p})f(\mathbf{q})d\mathbf{q}=\sum_{i=1}^n\lambda_i\langle f,\phi_i\rangle_2\phi_i(\mathbf{p}).$$

Indeed,

$$\begin{aligned}&\|\mathcal{F}f(\mathbf{p})-d(i)(\mathcal{E}\mathbf{L}_{\mathbf{K}}\mathbf{f})(\mathbf{p})\|_2^2\\&\approx\left\|\sum_{i=1}^n\varphi(\lambda_i)\left[\langle f,\phi_i\rangle_2\phi_i-d(i)\langle\mathbf{f},\mathbf{x}_i\rangle_2\right]\phi_i\right\|_\star^2\\&=\sum_{i=1}^n|\varphi(\lambda_i)|^2|\langle f,\phi_i\rangle_2-d(i)\langle\mathbf{f},\mathbf{x}_i\rangle_2|^2,\end{aligned}$$

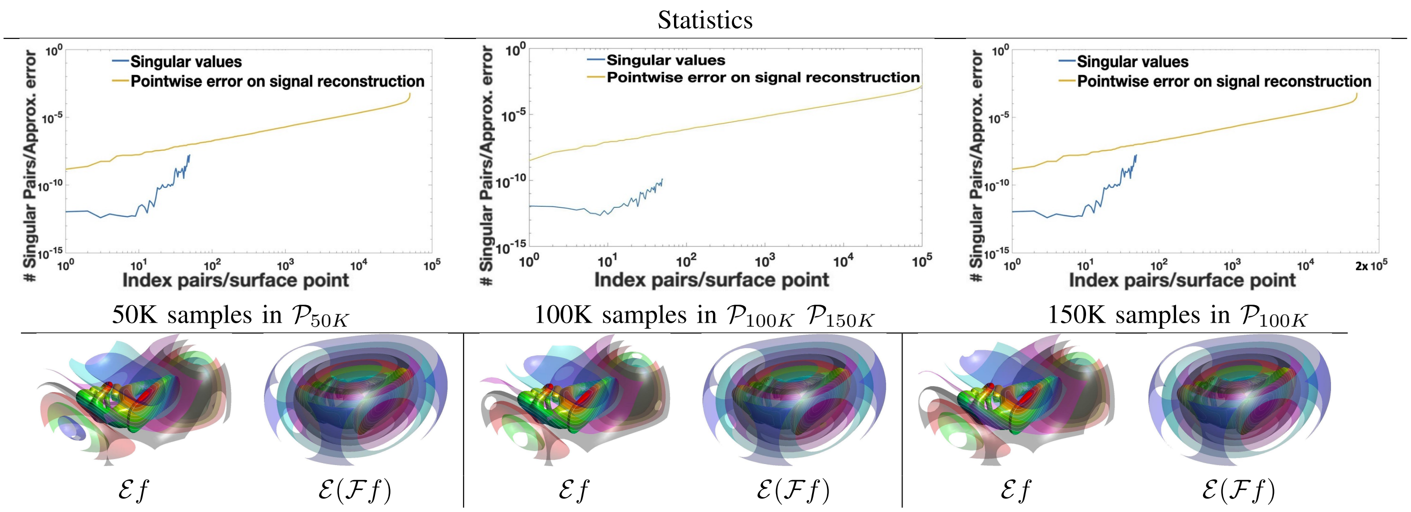

Fig. 13. Out-of-sample extension $\mathcal{E}f:\mathbb{R}^d\rightarrow\mathbb{R}$ of the input signal and its out-of-sample extension $\mathcal{E}(\mathcal{F}f):\mathbb{R}^d\rightarrow\mathbb{R}$ of the FT on a 3D domain with an increasing sampling density.

which converges to zeros as the sampling density increases; in fact, $\lambda_i\rightarrow 0$ and $d(i)\langle\mathbf{f},\mathbf{x}_i\rangle_2\rightarrow\langle f,\phi_i\rangle_2$, as $i\rightarrow+\infty$.

Approximated out-of-sample data-driven Fuzzy transform

From Eq. (11), we have that

$$\begin{aligned}\langle f,\phi_i\rangle_\star&=\langle\sum_{j=1}^{n}\langle\mathbf{f},\mathbf{x}_j\rangle_\mathbf{D}\phi_j,\phi_i\rangle_\star=\sum_{i,j=1}^{n}\langle\mathbf{f},\mathbf{x}_j\rangle_\mathbf{D}\langle\phi_i,\phi_j\rangle_\star\\&=\sum_{i,j=1}^{n}\langle\mathbf{f},\mathbf{x}_j\rangle_\mathbf{D}\langle\mathbf{x}_i,\mathbf{x}_j\rangle_\mathbf{D}=\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}.\end{aligned}\tag{12}$$

Recalling Eq. (4), we define the *approximated out-of-sample FT* $\mathcal{L}^{\text{approx}}_{K_\varphi}:\mathcal{C}^0(\Omega)\rightarrow\mathcal{C}^0(\Omega)$ as

$$\begin{aligned}&(\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}f)(\mathbf{p}):=\langle f,K_{\widetilde{\varphi}}(\mathbf{p},\cdot)\rangle_\star=\langle f,\sum_{i=1}^{n}\widetilde{\varphi}(\lambda_n)\phi_i(\mathbf{p})\phi_i\rangle_\star\\&=\sum_{i=1}^{n}\widetilde{\varphi}(\lambda_i)\langle f,\phi_i\rangle_\star\phi_i(\mathbf{p})=_{\text{Eq.}(12)}\sum_{i=1}^{n}\widetilde{\varphi}(\lambda_i)\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}\phi_i(\mathbf{p}).\end{aligned}\tag{13}$$

Then, Eq. (13) is efficiently computed as the term $\langle f,\phi_i\rangle_2$ in Eq. (10) is now replaced by $\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}$.

B. Properties

a) Linearity & continuity: The approximated out-of-sample FT is *linear* $\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}(\alpha f+\beta g)=\alpha\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}f+\beta\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}g$, and *"pseudo" continuous*, according to the following uppper bound

$$\begin{aligned}&\|\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}f\|_\star^2=\left\|\sum_{i=1}^{n}\widetilde{\varphi}(\lambda_i)\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}\phi_i\right\|_\star^2\\&=\left\|\sum_{i=1}^{n}\widetilde{\varphi}(\lambda_i)\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}\mathbf{x}_i\right\|_\mathbf{D}^2=\sum_{i=1}^{n}|\widetilde{\varphi}(\lambda_i)|^2|\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}|^2\\&\leq\|\mathbf{f}\|_\mathbf{D}^2\sum_{i=1}^{n}|\widetilde{\varphi}(\lambda_i)|^2\leq\|\mathbf{f}\|_\mathbf{D}^2\|\widetilde{\varphi}\|_2^2=\|f\|_\star^2\|\widetilde{\varphi}\|_2^2.\end{aligned}$$

Furthermore, $\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}$ is *linear with respect to the input filter*, i.e., $\mathcal{L}^{\text{approx}}_{K_{\alpha\widetilde{\varphi}_1+\beta\widetilde{\varphi}_2}}=\alpha\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_1}}+\beta\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_2}}$.

b) Self-adjointness and spectrum: The approximated out-of-sample FT is *self-adjoint* with respect to the pseudoscalar product, i.e.,

$$\langle\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}f,g\rangle_\star=\sum_{n=0}^{+\infty}\widetilde{\varphi}(\lambda_n)\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}\langle\mathbf{g},\mathbf{x}_i\rangle_\mathbf{D}=\langle f,\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}g\rangle_\star.$$

The *eigensystem* of $\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}$ is $(\widetilde{\varphi}(\lambda_n),\phi_n)_{n=0}^{+\infty}$, i.e.,

$$\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}\phi_j=\sum_{i=0}^{n}\widetilde{\varphi}(\lambda_i)\langle\mathbf{x}_j,\mathbf{x}_i\rangle_\mathbf{D}\phi_i=\widetilde{\varphi}(\lambda_j)\phi_j.$$

In particular, $\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}1=0$.

c) Convergence: Let us now estimate the approximation error between the out-of-sample FT and its approximation,

$$\begin{aligned}\|\mathcal{L}_{K_{\widetilde{\varphi}}}f-\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}}}f\|_2^2&=\left\|\sum_{i=1}^{n}\widetilde{\varphi}(\lambda_i)\left(\langle\phi_i,f\rangle_2-\langle\mathbf{x}_i,\mathbf{f}\rangle_\mathbf{D}\right)\phi_i\right\|_2^2\\&=\sum_{i=1}^{n}|\widetilde{\varphi}(\lambda_i)|^2\left|\langle\phi_i,f\rangle_2-\langle\mathbf{x}_i,\mathbf{f}\rangle_\mathbf{D}\right|^2,\end{aligned}$$

which converges to zero as $\langle\mathbf{x}_i,\mathbf{f}\rangle_\mathbf{D}$ converges to $\langle\phi_i,f\rangle_2$, as the sampling density of Ω increases. The same result applies when considering the norm $\|\cdot\|_\star$.

d) Fast computation: According to the upper bound

$$\begin{aligned}&\|\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_1}}f-\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_2}}f\|_\star^2=\|\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_1-\widetilde{\varphi}_2}}f\|_\star^2\\&=\left\|\sum_{i=1}^{n}(\widetilde{\varphi}_1(\lambda_i)-\widetilde{\varphi}_2(\lambda_i))\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}\phi_i\right\|_\star^2\\&=\left[\sum_{i=1}^{n}|\widetilde{\varphi}_1(\lambda_i)-\widetilde{\varphi}_2(\lambda_i)|^2|\langle\mathbf{f},\mathbf{x}_i\rangle_\mathbf{D}|^2\right]^2\leq\|\mathbf{f}\|_\mathbf{D}^2\|\widetilde{\varphi}_1-\widetilde{\varphi}_2\|_2^2,\end{aligned}$$

a good approximation $\widetilde{\varphi}_2$ of $\widetilde{\varphi}_1$ guarantees that $\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_2}}$ is a good approximation of $\mathcal{L}^{\text{approx}}_{K_{\widetilde{\varphi}_1}}$. For instance, let p be a polynomial (or rational polynomial) approximation of $\widetilde{\varphi}$ such that $r_\infty := \|\widetilde{\varphi} - p\|_\infty$. Then, $\|\mathcal{L}_{K_{\widetilde{\varphi}}} f - \mathcal{L}_{K_p} f\|_\star^2 \leq \|\widetilde{\varphi} - p\|_\infty \|\mathbf{f}\|_{\mathbf{D}}^2 \leq r_\infty \|\mathbf{f}\|_{\mathbf{D}}^2$. Indeed, the filter $\widetilde{\varphi}$ is approximated with a polynomial (or a rational polynomial) p, which allows us to evaluate $(\mathcal{L}_{\widetilde{\varphi}}^{\text{approx}} f)(\mathbf{p})$ without computing the Laplacian spectrum, which is generally unfeasible in terms of computational cost and storage overhead.

e) Restriction and out-of-sample extensions: Applying the linearity of the out-of-sample extension and restriction operators, and the relations $\mathcal{E}(\mathbf{x}_i) = \phi_i$, $\mathcal{R}(\phi_i) = \mathbf{x}_i$, we have that $\mathcal{L}_{K_{\widetilde{\varphi}}}$ is the out-of-sample extension of the discrete spectral operatir $\mathbf{L}_{\widetilde{\varphi}}$, i.e., $\mathcal{L}_{K_{\widetilde{\varphi}}} f = \mathcal{E}(\mathbf{L}_{\widetilde{\varphi}} \mathbf{f})$. In fact, $(\mathcal{L}_{K_{\widetilde{\varphi}}} f)(\mathbf{p}_i) = (\mathbf{L}_{\widetilde{\varphi}} \mathbf{f})(i)$, $i = 1, \ldots, n$. According to the previous relations, the pseudoscalar product and the approximated out-of-sample FT provide a coherent approximation of the FT. Similarly, we derive the relations $\mathcal{R}(\mathcal{L}_{K_{\widetilde{\varphi}}} f) = \mathbf{L}_{\widetilde{\varphi}} \mathbf{f}$ and $\mathcal{L}_{K_{\widetilde{\varphi}}}(\mathcal{E}\mathbf{f}) = \mathcal{L}_{K_{\widetilde{\varphi}}} f$.

f) Experimental tests: Given a discrete function $\mathbf{f}$ (Figs. 12, 13), we compute its out-of-sample extension according to Eq. (9), its discrete FT $\mathbf{F} := \mathbf{L}_\mathbf{K} \mathbf{f}$, and the corresponding out-of-sample extension $\mathcal{E}(\mathbf{F})$. The behaviour of the discrete FT $\mathbf{F}$ is represented through its colour map and level sets; the behaviour of the out-of-sample extension $\mathcal{E}(\mathbf{F}) : \mathbb{R}^d \rightarrow \mathbb{R}$ is represented through its iso-surface $\Sigma_\alpha := \{\mathbf{p} \in \mathbb{R}^d : \mathcal{E}\mathbf{F}(\mathbf{p}) = \alpha\}$. We notice the consistency between the behaviour of $\mathbf{F}$ and $\mathcal{E}\mathbf{F}$, where each level set γ_α on Ω corresponds to an iso-surface Σ_α of $\mathcal{E}(\mathbf{F})$.

VI. Conclusions and future work

This paper has presented the definition and computation of the out-of-sample membership functions and the resulting out-of-sample FT, which extend their discrete counterparts to the continuous case. Through the out-of-sample FT, we have achieved a coherent analysis of the discrete and continuous FTs, which is applied to extrapolate the behaviour of the FT on new data and to achieve an accurate approximation of the continuous FT of signals on arbitrary data. As main future work, we plan to investigate (i) the class of functions used for the out-of-sample FT and (ii) the definition of the generating kernel from the input data better to adapt the out-of-sample approximation of the input signal and of the FT to the data itself, e.g., generating these functions and kernels from the input data through learning and manifold learning.

Acknowledgements: We thank the reviewers for their thorough review and constructive comments, which helped us to improve the technical part and presentation of the revised paper. This work has been partially supported by the European Commission, NextGenerationEU, Missione 4 Componente 2, "*Dalla ricerca all'impresa*", Innovation Ecosystem RAISE "*Robotics and AI for Socio-economic Empowerment*", ECS00000035. Data and code will be available upon request.

References

[1] G. Patanè, "Data-driven fuzzy transform," *IEEE Trans. on Fuzzy Systems*, vol. 30, no. 9, pp. 3774–3784, 2022.
[2] ——, "Continuous fuzzy transform as integral operator," *IEEE Trans. on Fuzzy Systems*, vol. 29, no. 10, pp. 3093–3102, 2021.
[3] P. Berard, G. Besson, and S. Gallot, "Embedding Riemannian manifolds by their heat kernel," *Geometric and Functional Analysis*, vol. 4, no. 4, pp. 373–398, 1994.
[4] S. Ghosh, N. Das, T. Gonzalves, P. Quaresma, and M. Kundu, "The journey of graph kernels through two decades," *Computer Science Review*, vol. 27, pp. 88 – 111, 2018.
[5] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computations*, vol. 15, no. 6, pp. 1373–1396, 2003.
[6] S. Lafon, Y. Keller, and R. R. Coifman, "Data fusion and multicue data matching by diffusion maps," *IEEE Trans. on Pattern Analysis Machine Intelligence*, vol. 28, no. 11, pp. 1784–1797, 2006.
[7] A. Singer, "From graph to manifold Laplacian: the convergence rate," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 128 – 134, 2006.
[8] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011.
[9] M. Bronstein and A. Bronstein, "Shape recognition with spectral distances," *Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 1065 –1071, 2011.
[10] I. Perfilieva, "Fuzzy transforms: Theory and applications," *Fuzzy Sets and Systems*, vol. 157, no. 8, pp. 993 – 1023, 2006.
[11] I. Perfilieva, V. Novák, and A. Dvořák, "Fuzzy transform in the analysis of data," *Int. Journal of Approximate Reasoning*, vol. 48, no. 1, pp. 36–46, 2008.
[12] I. Perfilieva and B. D. Baets, "Fuzzy transforms of monotone functions with application to image compression," *Information Sciences*, vol. 180, no. 17, pp. 3304–3315, 2010.
[13] I. Perfilieva, P. Hodakova, and P. Hurtik, "Differentiation by the F-transform and application to edge detection," *Fuzzy Sets and Systems*, vol. 288, pp. 96 – 114, 2016.
[14] S. Rosenberg, *The Laplacian on a Riemannian Manifold*. Cambridge University Press, 1997.
[15] R. M. Rustamov, "Multiscale biharmonic kernels," *Computer Graphics Forum*, vol. 30, no. 5, pp. 1521–1531, 2011.
[16] M. Aubry, U. Schlickewei, and D. Cremers, "The wave kernel signature: a quantum mechanical approach to shape analysis," in *Computer Vision Workshops*, 2011, pp. 1626–1633.
[17] G. Golub and G. VanLoan, *Matrix Computations*. John Hopkins University Press, 2nd Edition, 1989.
[18] F. R. K. Chung, *Spectral graph theory*. American Mathematical Society, 1997.