Fractal properties of 4-point interpolatory subdivision schemes and wavelet scattering transform for signal classification

ARTICLE IN PRESS

Appl. Numer. Math. xxx (xxxx) xxx

Contents lists available at ScienceDirect

Applied Numerical Mathematics

journal homepage: www.elsevier.com/locate/apnum

Research Paper

Fractal properties of 4-point interpolatory subdivision schemes and wavelet scattering transform for signal classification

V. Bruni [a,*], F. Pelosi [b], D. Vitulano [a]

[a] Sapienza University of Rome, Department of Basic and Applied Sciences for Engineering, via Antonio Scarpa 16, Roma, 00161, Italy
[b] University of Siena - Department of Information Engineering and Mathematics, via Roma 56, Siena, 53100, Italy

ARTICLE INFO

Keywords:
Wavelet scattering transform
4-point subdivision scheme
Fractal dimension
Signal classification

ABSTRACT

Wavelet scattering is a recent time-frequency transform that shares the convolutional architecture with convolutional neural networks, but it allows for a faster training and it often requires smaller training sets. It consists of a multistage non-linear transform that allows us to compute the deep spectrum of a signal by cascading convolution, non-linear operator and pooling at each stage, resulting a powerful tool for signal classification when embedded in machine learning architectures. One of the most delicate parameters in convolutional architectures is the temporal sampling that strongly affects the computational load as well as the classification rate. In this paper the role of sampling in the wavelet scattering transform is studied for signal classification purposes. In particular, the role of subdivision schemes in properly compensating the information lost when using sampling at each stage of the transform is investigated. Preliminary experimental results show that, starting from coarse grids, interpolatory subdivision schemes reproduce copies of the original scattering coefficients at a fixed full grid that still represent distinctive features for signal classes. In fact, thanks to the ability of the scheme in reproducing similar fractal properties of the transform through an efficient iterative refinement procedure, the reproduced coefficients enable to obtain classification rates similar to those provided by the native wavelet scattering transform. The relationships between the tension parameter of the scheme and the fractal dimension of its limit curve are also investigated.

1. Introduction

Signal classification is assuming greater importance thanks to the increasing development of acquisition devices and the plethora of application fields that are based on signal processing [1–4]. Some examples are: *i*) scene recognition or air traffic control from acoustic signals in security and control systems; *ii*) arm motion recognition from electromyographic signals and heart anomalies recognition from ECG signals in medical applications; *iii*) human activity, spoken digit or music genre recognition as apps for smartphones — see Fig. 1. Signal classification mainly consists of two main steps: *feature extraction* and *model selection*. Feature extraction aims at selecting those relevant characteristics of the signal that make its class uniquely identified. Model selection consists in selecting the functions and the algorithms to use for discriminating and separating the features of signals belonging to different classes. Wavelet-inspired feature extraction methods are well suited for this task thanks to the nice and powerful function representation properties of the wavelet transform that made it a useful tool in a wider range of tasks and applications [5–11].

* Corresponding author.
E-mail address: vittoria.bruni@uniroma1.it (V. Bruni).

https://doi.org/10.1016/j.apnum.2024.09.022
Received 29 May 2024; Received in revised form 22 September 2024; Accepted 24 September 2024

Available online 2 October 2024

Please cite this article as: V. Bruni et al., *Appl. Numer. Math.*, https://doi.org/10.1016/j.apnum.2024.09.022

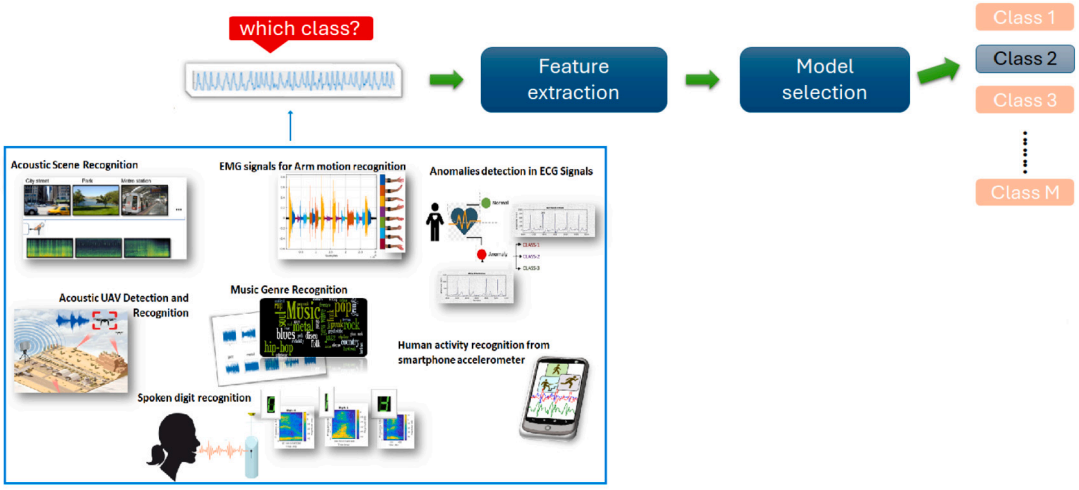

Fig. 1. Brief sketch of the signal classification process and some of related applications.

Wavelet Scattering Transform (WST) [12,13] is one of the most powerful tools [14,15] designed to emphasize signal features. Its basic scheme arises from the MEL spectrum decomposition, usually adopted to process audio/speech signals [16–18], and it has been successively generalized to higher dimensions, keeping the 1D time-frequency scattering properties [19].

Based on a recursive scheme, WST mainly consists of: i) a first stage of convolutions with wavelet kernels having constant-Q factor, ii) a pointwise non linear operator i.e., the absolute value of the output coefficients of the first phase, and iii) a further convolution step through a smoothing kernel — that mainly resembles a moving average applied to the output of the second step. It is worth highlighting that the algorithmic structure of WST is very similar to the well-known Convolutional Neural Networks [14], with the great advantage that convolution kernels do not need to be learnt from data. It turns out that its computation is very fast, avoiding the whole training process that usually is very time-consuming. Additional interesting properties of WST are temporal and frequency shift invariance, good frequency resolution and robustness to (moderate) deformations. It can also be proved [12,13] that WST is a contractive operator that leads to energy concentration in its first layers — especially the first one. However, despite the nice mathematical and computational properties of the transform, filters placement yields some redundancy of the transform so that suitable sampling is required for effective and feasible signal classification, especially when using a machine learning framework. In fact, in this context too little information leads to poor classification performance, so does too much information. The selection of the optimal number of coefficients that help the classifier to work correctly is then an open and delicate issue to be addressed, even for WST. The approaches related to this topic mainly exploit the non-expansion property of WST. In particular, they studied the correlation between WST layers to reduce its redundancy, as for instance in [12,20,21]. Other approaches have exploited more traditional tools suitably designed for this purpose [22–25]. More oriented approaches can be found in [21,26,27], where techniques like Principal Component Analysis (PCA), MultiDimensional Scaling (MDS) and Random Sampling have been adapted to lighten WST structure. All these approaches guarantee good classification results paying the price of an *a priori* tuning of the number of features to be preserved. An automatic scattering feature selection method has been presented in [28], where the Minimum Description Length principle is applied to reach a tradeoff between number of features and classification rate. However, due to the complexity and variability of the data and tasks, the problem still remains open.

As already outlined, feature reduction usually leads to an increasing and moderate loss in performance till an optimal point beyond which any further reduction gives definitively worst classification results — w.r.t. the application at hand. This ‘optimal features cardinality’ seems to be an insurmountable limitation for any classification process. In fact, if on one hand this choice leads to great savings in memory/processing/transmission for big data problems, on the other hand it pays the price of a sensible reduction in terms of classification accuracy [29,30]. In this paper we then try to address the problem of compensating for any severe feature reduction. In particular, we gave empirical evidence that a suitable interpolation of the reduced features can provide additional, data-driven and distinctive information which contributes to enhance the discrimination capabilities of a given machine-learning classifier, thus increasing the classification accuracy. It is worth observing that this result offers an additional advantage in terms of storage resources, as it paves the way to an optimal information storage with a few coefficients that can be conveniently ‘expanded’ whenever needed — a previous attempt in this sense can be found in [31]. Due to the variety of signals and types of features, in this paper we investigate the problem of WST-based classification of specific classes of signals, such as audio signals, that can be naturally represented in the time-frequency domain and are characterized by non-smooth components and rapid oscillations that make their compact representation a difficult task [32,33]. In fact, even their deep spectrum, which is provided by the WST, produces highly variable coefficients, making their sampling a delicate task, especially when defining a suitable input for classification tools, such as Support Vector Machine (SVM), and so on.

Based on these considerations, the first contribution of this paper is then to exploit the complexity of the deep spectrum of audio signals for guiding the feature expansion process. To this aim the fractal dimension of the wavelet scattering coefficients is evaluated bandwise. Fig. 2 shows a preliminary illustrative evidence of this statement, where the fractal dimension is measured through the

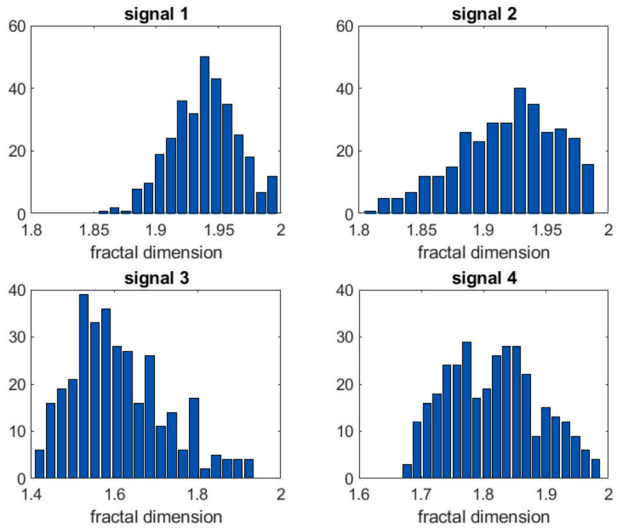

Fig. 2. Distribution of the fractal dimension of each band of the WST of four distinct audio signals from GTZAN [56] dataset. The fractal dimension has been measured using the Higuchi method, it falls in the interval $(1, 2)$ and it is greater than 1, as expected for 1D fractal functions.

Higuchi method [34]. As can be observed, the fractal dimension of scattering coefficients of audio signals is far from being equal to 1, unlike what would be expected for any conventional 1D signal.

Both theoretical and experimental analyses are then conducted to answer the following questions:

- Is it possible to interpolate critically sampled WST coefficients, having fractal nature, to the desired temporal grid while preserving/recovering the native fractal roughness?
- Which interpolation scheme is suitable for this purpose?
- Are the interpolated WST coefficients able to achieve the same classification rates of the original data on that grid?

Preliminary answers to these issues are mainly related to the already mentioned fractal nature of WST representations and have some innovative elements. Numerical evaluations will show that coefficients can be critically subsampled and successively 'recovered' by interpolation, preserving classification capabilities even for complicated signals. To this aim, suitable interpolation tools are subdivision schemes that converge to fractal curves; on the contrary, subdivision schemes converging to smooth curves often lead to worst classification results. In particular, the well known four-points subdivision scheme [35] is adopted in this paper since it allows us to interpolate WST coefficients, even in a fractal context (see [36]), through a simple parameter tuning, resulting computationally efficient.

In addition, bounds for the fractal dimension of the limit curve of the scheme are provided, highlighting an explicit dependence on the tension parameter of the scheme. The latter result represents a step forward toward the automatic selection of the best setting of the scheme parameters, depending on the range of the expected values for the fractal dimension to be reproduced. Apart from the literature related to the analysis of fractal or multifractal signals by means of the canonical wavelet transform [32,37,38], to the best of the authors' knowledge, this work is the first attempt to deal with wavelet scattering transform by adopting a methodology that is based on the properties of fractals. In addition, it represents a first step toward effective embeddings of subdivision schemes in machine learning processes.

The remainder of the paper is organized as follows. Next section provides a short overview of Wavelet Scattering Transform (WST) and classical four-points subdivision schemes (4SS). Section 3 presents the operational details of the proposed methodology in which WST and 4SS are combined to achieve an adequate signal representation that promotes good classification performance of a standard SVM classifier. Section 4 provides some experimental results and evaluation studies with respect to different tuning of the involved tools. Finally, Section 5 draws the conclusions and provides some guidelines for future work.

2. Mathematical background

In this section, a very short description of the two main elements of the proposed model will be offered: the Wavelet Scattering Transform [12] and the Subdivision Schemes (the reader may refer to [39], [40], [41] and references therein for a comprehensive treatment of univariate subdivision schemes). This will help the reader to better understand the sequel of this paper.

2.1. Wavelet scattering transform

Wavelet Scattering Transform (WST) [12] is a multi-layer non-linear time-frequency transform that allows us to compute the "deep spectrum" of a signal by cascading *convolutional*, *non-linear* and *pooling* operators at each layer except for the zero-th order

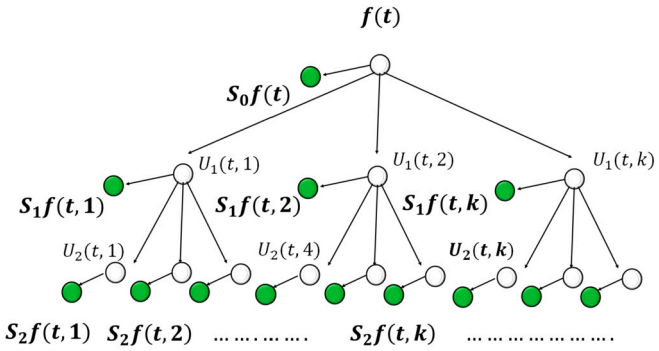

Fig. 3. Tree structure of the Wavelet Scattering Transform — green dots denote the output of each layer. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

layer. In particular, the transform is characterized by the tree structure represented in Fig. 3, where the 0-th layer consists of the convolution product of the original input signal f with a scaling function ϕ [32] that acts as a low pass filter. The successive m-th order wavelet scattering layers are defined by applying the regularizing function ϕ to the bands of a special constant-Q factor wavelet decomposition where each convolution with a band-pass filter is followed by the absolute value operator, that acts as a pointwise non-linear activation function.

More formally, let $f(t)$ with $t \in \Omega_t$ be a time-dependent signal defined in temporal interval $\Omega_t \subset \mathbb{R}$; and let $\phi(t)$ be a scaling function, $\{Q_m\}_{m=1,\dots,L}$ a sequence of integer Q-factors, and $\{\psi_{m,k}(t)\}_{k\in\mathbb{Z}}$ the constant Q_m-factor wavelet filter bank generated by the mother wavelet $\psi(t)$. Q_m represents the number of wavelets per octave, while

$$\psi_{m,k}(t) = \sqrt{\lambda_{m,k}}\psi(\lambda_{m,k}\, t), \tag{1}$$

where $\lambda_{m,k} = 2^{\frac{k}{Q_m}}$ is the dilation parameter and $k \in \mathbb{Z}$.

The 0-th order wavelet scattering layer $S_0 f$ is defined as

$$S_0 f(t) = f * \phi(t), \tag{2}$$

where $*$ denotes the convolution product along the temporal axis. Similarly, by denoting with I_m the set of admissible values for the parameter k at the m-th layer, the 1-st order wavelet scattering layer is

$$\mathbf{S}_1 f = [S_{1,k}(t)]_{k\in I_1 \subset \mathbb{Z}}, \tag{3}$$

where $[\cdot]$ stands for the collection $[S_{1,1}(t), S_{1,2}(t), \dots]$, with

$$S_{1,k}(t) = U_{1,k} * \phi(t),$$

and

$$U_{1,k}(t) = |f * \psi_{1,k}(t)|.$$

Accordingly, the m-th layer is defined as

$$\mathbf{S}_m f = [S_{m,k}(t)]_{k\in I_m \subset \mathbb{Z}}, \tag{4}$$

where

$$S_{m,k}(t) = U_{m,k} * \phi(t),$$

with

$$U_{m,k}(t) = |\mathbf{U}_{m-1} f * \psi_{m,k}(t)|,$$

where

$$\mathbf{U}_{m-1} f = [U_{m-1,k}(t)]_{k\in I_{m-1} \subset \mathbb{Z}}.$$

Definition 2.1. The *Wavelet Scattering Transform* (WST) $\mathbf{S}$ of the signal f is the collection of the $L+1$ layers, i.e.,

$$\mathbf{S} f = \{S_0 f, \mathbf{S}_1 f, \dots, \mathbf{S}_L f\}, \tag{5}$$

with $S_0 f$ as defined in eq. (2), and $\mathbf{S}_1 f, \dots, \mathbf{S}_L f$ as defined in eq. (4) with $m = 1, \dots, L$,

$\mathbf{S}f$ has a matrix form whose dimension depends on the properties of the constant Q-factor wavelet transform. In particular, the temporal dimension (*columns*) depends on the critical sampling that the transform allows to apply; the frequency dimension (*rows*) mainly depends on the number of octaves used in the transform at each layer. The redundancy of the transform thus depends on the set $\{Q_m\}_{m=1,\dots,L}$ of Q-factors selected at each layer.

WST has some nice properties that mainly inherits from the wavelet-based cascaded scheme [12,42]; they hold for $f \in L^2(\mathbb{R})$ and are listed below — the complete proofs can be found in [42], respectively Prop 2.9, Prop. 2.5 and Theorem 2.6.

Proposition 2.1 *(Time warping stability). Let denote $f_\tau(t) = f(t-\tau(t))$ and let T be the support of the scaling function ϕ, then $\exists$ a constant $C > 0$ such that $\forall\ \tau(t)$ such that $sup_t|\tau(t)| << T \quad and \quad |\tau'(t)| < 1$ it holds*

$$\|\mathbf{S}f_\tau - \mathbf{S}f\| \le C\, sup_t|\tau|\, \|f\|.$$

This property assures the stability of the transform to small deformations and variations $\tau(t)$ in the input signal f. This means that even when the input signal is distorted slightly, the scattering coefficients remain relatively unchanged. This property is valuable for tasks like object recognition or speech processing, where robustness to variations is essential. The property implies translation invariance as a particular case.

Proposition 2.2 *(Contraction). $\forall\ f(t),\ g(t) \in L^2(\mathbb{R})$, it holds*

$$\|\mathbf{S}f - \mathbf{S}g\| \le \|f - g\|.$$

This property provides a non expansive transform that is stable to additive noise.

Proposition 2.3 *(Energy preservation). If each constant Q-factor wavelet transform, defined by the family $\{\psi_{m,k}(t)\}_{k\in\mathbb{Z}}$ as in eq. (1), is a tight frame, then*

$$\|f\|^2 = \|\mathbf{S}f\|^2 + \|\mathbf{U}_{L+1}f\|^2.$$

This property assures the preservation of the energy of the signal. In addition, under appropriate assumptions on the mother wavelet ψ, the high frequency component of the scheme, that is $\mathbf{U}_{L+1}$, goes to zero as the depth of the transform (L) increases. In fact, the modulus operator computes a smooth envelope, and pushes energy towards lower frequencies. Since the scattering transform is iterated on wavelet modulus operators $\mathbf{U}_m f$ (not on the output of the preceding layer $\mathbf{S}_{m-1}f$), it propagates all the energy of the signal towards lower frequencies, which is captured by the lowpass filter of scattering coefficients. Therefore, WST preserves signal energy and its depth can be reduced, since most of the signal energy is concentrated in the very first layers.

For the implementation details of the transform, refer to [12].

2.2. 4-point subdivision schemes

Subdivision schemes, introduced in 1970s (see [43–45] for the pioneer ideas and [39–41] for an exhaustive reading), are efficient tools for computer-aided curve and surface design. The basic idea is to define a curve or a surface out of an initial control polygon or control mesh by recursively subdividing them according to some refining rules. Due to its efficiency and convenience, subdivision is now an important tool in many applications in fields including Computer Graphics, Computer Aided Geometric Design, computer animation, surgical simulation and medical image processing.

According to whether the initial control points will be interpolated or not, subdivision schemes can be classified into two categories: interpolatory and approximate subdivision schemes. In general, approximating schemes give smoother curves or surfaces, but in the present we will focus on the interpolatory schemes to preserve the given data. The first and most popular interpolatory scheme is the called *4-point interpolatory scheme* (4SS), introduced by Dyn, Levin and Gregory in [35]. The recursive steps are based on a 4-point rule, as it follows.

Given the set of initial control points $\mathbf{P}^0 = \{\mathbf{P}_j^0 \in \mathbb{R}^d\}_{j=-2}^{n+1}$, let $\mathbf{P}^k = \{\mathbf{P}_j^k \in \mathbb{R}^d\}_{j=-2}^{2^k n+1}$ be the set of control points at level k ($k \ge 0,\ k \in \mathbb{Z}$), defines $\{\mathbf{P}_j^{k+1} \in \mathbb{R}^d\}_{j=-2}^{2^{k+1}n+1}$ recursively by the following 2 rules (binary scheme):

$$\begin{cases} \mathbf{P}_{2j}^{k+1} = \mathbf{P}_j^k, & -1 \le j \le 2^k n + 1, \\ \mathbf{P}_{2j+1}^{k+1} = \left(\frac{1}{2} + w\right)\left(\mathbf{P}_j^k + \mathbf{P}_{j+1}^k\right) - w\left(\mathbf{P}_{j-1}^k + \mathbf{P}_{j+2}^k\right), & -1 \le j \le 2^k n, \end{cases} \tag{6}$$

where w is a tension (shape) parameter. Each iteration of the binary scheme defined in eq. (6), can be rewritten as follows

$$\mathbf{P}^{k+1} = A\mathbf{P}^k,$$

where $A \in \mathbb{R}^{2^{k+1}n+4 \times 2^k n+4}$ is the subdivision matrix and has the following structure

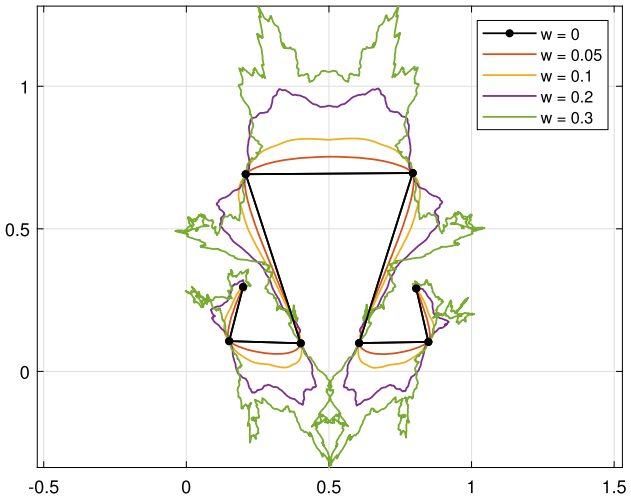

Fig. 4. Curves obtained after 6 steps of the subdivision scheme (6) on the initial points (black dots) for different values of the tension parameter $w = 0,\ 0.05,\ 0.1,\ 0.2,\ 0.3$.

$$A = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ -w & \frac{1}{2}+w & \frac{1}{2}+w & -w & 0 & 0 & 0 & \dots \\ \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & -w & \frac{1}{2}+w & \frac{1}{2}+w & -w & 0 & 0 & \dots \\ \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ \dots & 0 & -w & \frac{1}{2}+w & \frac{1}{2}+w & -w & 0 & \dots \\ \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ \dots & 0 & 0 & -w & \frac{1}{2}+w & \frac{1}{2}+w & -w & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{7}$$

It is worth observing that in literature there exist subdivision schemes with general m-arity, where m denotes the number of rules used to built new points and coincides with the factor by which the number of points is multiplied in each step (see [46] for ternary and [47] for quaternary schemes). For our approach we limit to binary scheme since it duplicates the number of control points, and thus it is in perfect matching with the upsampling/downsampling of a multiresolution setting.

By letting k tend to infinity, this process defines an infinite set of points in $\mathbb{R}^d$ whose smoothness properties depend on the values of the parameter w. In particular, it is proved that the 4-point interpolatory scheme is C^0-continuous for $-\frac{1}{4} < w < \frac{1}{4}$ and C^1-continuous for $0 < w < \frac{\sqrt{5}-1}{8}$ (see [35,48,49]). The 4-point scheme generates the piecewise linear interpolant in the initial control polygon for the trivial value $w = 0$, whereas, for $w = \frac{1}{16}$, the limit curve has the best possible Hölder regularity (i.e., it is almost C^2 with $R_H = 2 - \varepsilon$). Some example of continuous curves obtained with the scheme (6) can be observed in Fig. 4 for $w = 0,\ 0.05,\ 0.1,\ 0.2$.

The subdivision scheme method is not only an important tool for the fast generation of smooth geometric objects from initial control nets, but also an efficient tool for the fast generation of fractals. Actually the limit curve of the 4-point binary can be fractal for some special values of the corresponding parameters, as it is stated in the following theorem (see [36] for details).

Theorem 2.1. *For $-\frac{1}{2} < w < 0$ or $\frac{1}{4} \le w < \frac{1}{2}$, the limit curve of the 4-point scheme (6) is a fractal curve.*

An example of the fractal behavior of a 4-point subdivision scheme is shown in Fig. 4 for $w = 0.3$.

In particular, it can be proved that fractal limit curve of the scheme is the unique attractor of a specific Iterated Functions System (IFS) [50,51]. The i-th function of the IFS is

$$f_i(X) = X\, P^{-1} A_i P, \quad i = 1, \dots, \bar{d} \tag{8}$$

where the matrix A_i is obtained by breaking the subdivision matrix A into multiple square matrices, while P is the matrix of the initial control points suitably lifted to guarantee the equivalence with the scheme — hence, some of the rows of the matrices A_i may overlap (for more detail see [50,51]).

3. The proposed model

The proposed model deals with the $L+1$-layers WST, as defined in eq. (5), of an audio signal f. In order to make more readable the following formalism, the matrix form of the transform is considered. Therefore the WST $\mathbf{S}f$ of the signal f is denoted by $S_{M,N}$

in the sequel, where S is a rectangular matrix with (original) size $M \times N$: the rows denote the frequency bands of the wavelets used in the transform, while the columns are related to the time samples. As already outlined, $S_{M,N}$ is strongly redundant both on time axis and along scales. Hence, reduction is necessary. For an approach oriented to reduce information along rows, the reader can refer to [28]. Conversely, in this section we will deal with temporal sampling. The reduction problem for classification purposes can be formalized as follows.

Problem statement 1: wavelet scattering feature reduction for classification Let $\mathcal{F}$ be a suitable classification operator, $\mathcal{R}$ be the corresponding success score (*classification accuracy* and/or *computational load*— the higher the score the better the performance/efficiency of task) and let $S_{M,N}$ be the WST of a signal f. Define a suitable sampling operator $\mathcal{T}_{N1}: \quad S_{M,N1} = \mathcal{T}(S_{M,N}), \quad$ with $\quad N1 << N$ such that

$$\mathcal{R}[\mathcal{F}(S_{M,N})] \leq \mathcal{R}[\mathcal{F}(S_{M,N1})].$$

As mentioned in the introduction, $S_{M,N1}$ can be achieved using proper feature reduction methods, or restricting the type of sampling (e.g. uniform sampling) and using a 'trial and test' procedure: for each subsampling, the accuracy obtained with a predefined classification tool is evaluated. The procedure stops when the 'optimal' value $\overline{N}1$ is found i.e., when $S_{M,\overline{N}1}$ leads to the best (acceptable) classification score; more formally,

$$\overline{N}1 = argmax_{N1 \leq N} \mathcal{R}[\mathcal{F}(S_{M,N1})], \qquad \text{with} \quad S_{M,N1} = \mathcal{T}_{N1}(S_{M,N}). \tag{9}$$

In this paper $\mathcal{T}_{N1}$ is assumed to be known. In fact, independently of the procedure used for getting $\overline{N}1$, we are interested in whether there exists a coarser sub-sampling of $S_{M,N}$ that can guarantee the same optimal classification rate, as in eq. (9), whenever it is suitably interpolated to the finer $\overline{N}1$-points grid. More formally, the issue we are interested to address is the following.

Problem statement 2: wavelet scattering interpolation for classification Let $\overline{N}1$ be defined as in eq. (9) and $G_{\overline{N}1}$ the corresponding nodes grid, $S_{M,N2} = \mathcal{T}_{N2}(S_{M,N})$, with $N2 << \overline{N}1$ and $G_{N2} \subset G_{\overline{N}1}$, a coarser sub-sampled version of $S_{M,N}$, and let $\hat{S}^{N2}_{M,\overline{N}1}$ be a properly interpolated version of $S_{M,N2}$ to the finer grid $G_{\overline{N}1}$. Find $N2$ such that

$$\mathcal{R}[\mathcal{F}(\hat{S}^{N2}_{M,\overline{N}1})] \approx \mathcal{R}[\mathcal{F}(S_{M,\overline{N}1})]. \tag{10}$$

Solving the last problem requires answering two main questions:

- *Question 1* which property or constraint must be satisfied by the interpolated data;
- *Question 2* which interpolation procedure ensures that the constraint is met.

In fact, the rationale is to apply proper interpolatory schemes able to reproduce some of distinctive geometrical properties of the data rather than the data itself, even when a severe sampling has been applied to the original data.

Regarding *Question 1*, as already observed in the introduction, considering the nature of the analyzed signals, fractal dimension is considered as the feature to be reproduced. Fractal dimension has been introduced by Mandelbrot in 1967 [52–54] and provides a rational index of signal complexity that represents how its behavior changes along scales — signals having the same fractal dimension scale accordingly. It is usually computed via the Hausdorff-Besicovitch dimension. However, from a numerical point of view it is more convenient to consider the capacity dimension, which will be described in the sequel.

Let V be a bounded set in $\mathbb{R}^d$. Intuitively, one can measure V by means of the minimum number of balls $B_V(s)$, with radius s, required to completely cover V itself. Hence,

Definition 3.1. [32] Let V be a bounded set in $\mathbb{R}^d$ and let $B_V(s)$ denote the minimum number of balls with radius s covering V. The *capacity dimension* D_V of the set V is

$$D_V = -\liminf_{s \to 0} \frac{log\, B_V(s)}{log\, s}, \tag{11}$$

and the measure $\mathcal{M}_V$ of V can be approximated as

$$\mathcal{M}_V = -\limsup_{s \to 0} B_V(s) s^{D_V}. \tag{12}$$

Definition 3.1 introduces a measure $\mathcal{M}_V$ that can be finite or not and approximates the Hausdorff distance. Hence, the exponent D_V, as defined in eq. (11), will be considered as the fractal dimension of V in the sequel of paper.

It is worth observing that the fractal dimension of a feature set can be thought as a measure of the complexity, or equivalently the space-filling capacity of the data. In machine learning, this relates to the intrinsic dimensionality of the data set, which can be indicative of the complexity and variability within the data. Therefore, replicating the fractal dimension implies creating a new feature set that preserves the same intrinsic complexity as the original features. This does not necessarily mean having the same values, but the transformed features should have the same geometrical or topological properties in terms of their distribution.

Regarding the interpolation operator able to synthesize the missing WST information while reproducing the fractal nature of the data (*Question 2*), a subdivision scheme can be adopted with the shape parameter w suitably tuned, according to Theorem 2.1. The

subdivision scheme should be applied by starting from a $N2$-points regular grid and iterated in order to interpolate $S_{M,N2}$ row-wise and to achieve $\hat{S}^{N2}_{M,\overline{N1}}$ meeting the condition in eq. (10). However, since $\mathcal{F}$ has some internal optimization schemes that are hard to be formalized, in the next subsection we are interested in formally evaluating the ability of the subdivision scheme in reproducing the fractal nature of the original data, that is expected to reflect the complexity of the signal. In the section of experimental results we will then empirically prove that, thanks to the preserved complexity, scattering coefficients having the same fractal nature provide nearly the same clusters, i.e. the same separating hyperplanes in the case of Support Vector Machine (SVM).

3.1. Fractal dimension and tension parameter

To provide insights on the relationship between the fractal dimension of the limit curve and the tension parameter w of the 4-point subdivision scheme in eq. (6), we must recall the following property of fractals [55].

Theorem 3.1. *Let $\{f_1, f_2, ..., f_{\bar{d}}\}$ be an Iterated Function System (IFS) on $\mathbb{R}^N$ with the attractor $\mathcal{A}$ such that every function f_i has a contraction constant ρ_i. If the IFS is totally disconnected or just-touching, then the fractal dimension $D_{\mathcal{A}}$ of the attractor $\mathcal{A}$ is the unique positive and finite solution of the following equation*

$$\sum_{i=1}^{\bar{d}} \rho_i^{D_{\mathcal{A}}} = 1$$

If the IFS is overlapping, then $D_{\mathcal{A}} \leq \tilde{D}$, where $\tilde{D} \in [0, +\infty)$ is the solution of

$$\sum_{i=1}^{\bar{d}} \rho_i^{\tilde{D}} = 1 \tag{13}$$

It is worth observing for planar curves it is expected that $1 \leq D_{\mathcal{A}} \leq 2$ whenever $\mathcal{A}$ is a row of the scattering matrix $S_{M,\overline{N1}}$.

Hence, using Theorem 3.1 and the matrix representation of the subdivision scheme in eq. (7), the following Proposition holds true.

Proposition 3.1. *Let $\mathcal{A}$ be the limit curve of the 4-point subdivision scheme in eq. (6) with tension parameter $w \in \left(-\frac{1}{2}, 0\right) \cup \left[\frac{1}{4}, \frac{1}{2}\right)$, let $\{f_i\}_{i=1,...,\bar{d}}$ be the associated Iterated Functions System, as defined in eq. (8), and let $D_{\mathcal{A}}$ be the fractal dimension of $\mathcal{A}$, according to Definition 3.1, then*

$$(2 - \varepsilon) + log_2(\rho) \leq D_{\mathcal{A}} \leq -\frac{log_2(\bar{d})}{log_2(\rho)} \tag{14}$$

where $\rho = \begin{cases} 2w & w \in \left[\frac{1}{4}, \frac{1}{2}\right) \\ \frac{1}{4}(1 + \sqrt{1 - 16w}) & w \in \left(-\frac{1}{2}, 0\right) \end{cases}$ *and ε is a proper positive constant.*

The proof is in the Appendix.

Proposition 3.1 provides bounds for the fractal dimension of the limit curve of the 4-point subdivision scheme. It is then expected that a given fractal dimension $D_{\mathcal{A}} \in [1, 2]$ can be reached, with high accuracy, by the limit curve generated by the 4-point subdivision scheme for proper choices of the shape parameter $w \in \left(-\frac{1}{2}, 0\right) \cup \left[\frac{1}{4}, \frac{1}{2}\right)$.

Remark 1 It is worth observing that, under proper assumptions, there exists a tension parameter able to approximate the fractal dimension of the expected function better than the one of the linearly interpolated data. Specifically, let $\hat{S}^{N2}_{M,\overline{N1}}(w)$ be the interpolated $S_{M,N2}$, where the dependence on the shape parameter w of the 4SS has been made explicit, and let define

$$\bar{w} = argmin_{w \in \left(-\frac{1}{2}, 0\right) \cup \left[\frac{1}{4}, \frac{1}{2}\right)} |D_{S_{M,\overline{N1}}} - D_{\hat{S}^{N2}_{M,\overline{N1}}(w)}|$$

with $\bar{w} : 2 - \varepsilon + log_2 \rho(\bar{w}) \leq D_{\hat{S}^{N2}_{M,\overline{N1}}(\bar{w})}$, then

$$|D_{S_{M,\overline{N1}}} - D_{\hat{S}^{N2}_{M,\overline{N1}}(\bar{w})}| \leq |D_{S_{M,\overline{N1}}} - D_{S_{M,N2}}|. \tag{15}$$

In fact, it is sufficient to observe that the planar curve defined by $S_{M,N2}$ is $\hat{S}_{M,N2}(0)$, i.e. the piecewise linear function that is obtained from the subsampled scattering matrix $S_{M,N2}$ using the 4SS with shape parameter $w = 0$. The limit curve is expected to have dimension 1, as derives from Proposition 3.1.

Remark 2 The numerical evaluation of the fractal dimension of a sequence requires the use of proper estimators, like the Higuchi estimator which provides reasonable estimations of the fractal dimension [34]. It depends on a free parameter k_{max} whose value is greater than 2. In the case of 1D sequences, k_{max} denotes the value of the maximum allowed sampling step. More precisely, let

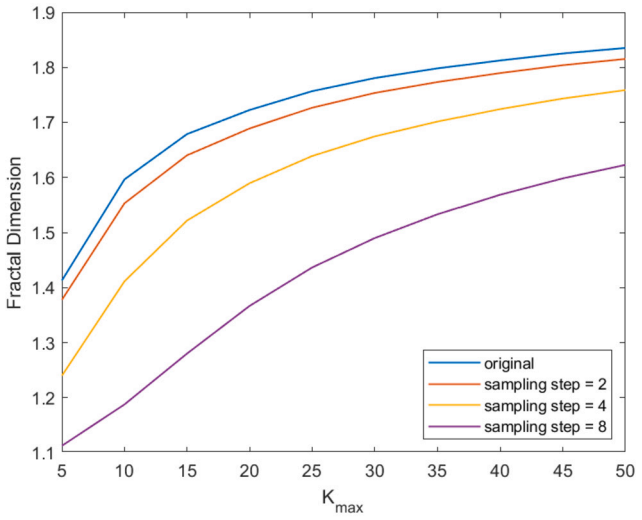

Fig. 5. Average fractal dimension of WST coefficients of 800 audio signals: different subsampling steps have been used for the WST matrix: fractal dimension decreases as subsampling step grows.

$X_i : i \in \{1, ..., N\}$ be the values of a given function computed over a uniform grid of knots (for example, a row of the scattering matrix $S_{M,N}$). Hence, $\forall\, k \in \{1, .., k_{max}\}$ and $m \in \{1, .., k\}$, it is possible to define the following lengths

$$L_m(k) = \frac{N-1}{\lfloor \frac{N-m}{k} \rfloor k^2} \sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |X_{m+ik} - X_{m+(i-1)k}|$$

and their averages with respect to m, i.e.,

$$L(k) = \frac{1}{k} \sum_{m=1}^{k} L_m(k).$$

According to eq. (11), $L(k)$ is expected to satisfy a negative power law with respect to its argument k, whose exponent represents the fractal dimension $D_{\{X_i\}}$ of the sequence $\{X_i\}_{1 \le i \le N}$. A simple linear regression of the set of data $\left\{ \left(\log \frac{1}{k}, \log L(k) \right) \right\}_{k=1,...,k_{max}}$ then provides an estimate of $D_{\{X_i\}}$ as the slope of the fitted linear function. The latter will be adopted in the experimental results to give evidence of the fractal nature of WST coefficients.

4. Experimental results

The proposed approach has been tested on the well-known dataset GTZAN [56]. The latter is usually adopted for machine learning applications and involves different music genre signals. Specifically, it is composed of 10 genres, where each one contains 100 clips of 30 seconds and sampled at 22050 Hz. The original set of 1000 signals has been split into 800 (80%) signals for the training set and 200 (20%) signals for the test set. Obviously, both subsets are composed of signals randomly selected from the whole initial dataset, where classes are equally represented. All simulations have been performed in Matlab environment. A two-layer WST ($L = 2$) but using different couples of Q factors has been employed in all tests. Most of the results presented in this section refers to the couple of Q-factors (Q_1, Q_2), with $Q_1 = 8$ and $Q_2 = 1$. With these settings, WST provides the matrix $S_{M,N} = S_{334,256}$ for each signal in the training set; hence, our starting data have size $334 \times 256 \times 800$ (800 will be omitted in order to make easier the notation), where 334 are the WST bands, 256 are the time samples of each band and 800 are the number of training set signals. An increasing and uniform subsampling has been performed on $S_{334,256}$ along the original time dimension (256 samples): from $N = 256$ to $N2 = 32$ samples.

Test 1 The first test aims at showing that WST of audio signals have a fractal nature that becomes weaker as time sampling becomes more severe, as expected. Fig. 5 shows the results related to signals in the training set. As can be observed, increasing subsampling factors lead to different behaviors in the fractal dimension with respect to the parameter k_{max} employed in the Higuchi method. The fractal dimension deviates from its original one as much as the subsampling increases. In particular, the stronger subsampling leads to the matrix $S_{334,32}$ that loses most of the fractal data information.

Test 2 The second test sets $\overline{N}1 = N = 256$ and aims at verifying the condition in Problem statement 2, with decreasing values of $N2$ starting from $N2 = \overline{N}1$ to $N2 = 32$. Hence, the sequence of sampled WST i.e., $\hat{S}^{N2}_{M,\overline{N}1}$, with $N2 = 256, 128, 64, 32$, have been generated and then interpolated to get the sequence of WST matrices $\hat{S}^{N2}_{M,\overline{N}1}$, with $N2 = 256, 128, 64, 32$, each having the starting $\overline{N}1 = N = 256$ temporal samples. The interpolated matrices have been used as input for a classifier and classification accuracy has

Table 1
Classification accuracy, measured in terms of percentage (%) of correct assignment, achieved by subsampling ($N2 = 32$) and then interpolating a (8,1)-twolayer WST using a 4-point subdivision scheme with different values for the shape parameter w. Accuracies equal to or greater than the one achieved by using the original WST matrix $S_{334,256}$ (*last column*) are in bold.

	$\hat{S}^{32}_{334,256}$										$S_{334,256}$
w	-0.5	-0.25	-0.1	0	0.1	0.2	0.25	0.3	0.4	0.5	
%	86.5	87	87.5	87.5	87.5	87.5	87.5	**88.5**	88	**88.5**	*88.5*

Table 2
Classification accuracy (measured as percentage of correct assignments) achieved by subsampling ($N2 = 32$) and then interpolating a twolayer WSTs with different couples of Q factors (in the brackets) using a 4-point interpolatory subdivision scheme with different values of the shape parameter w. The last column contains the accuracy achieved using the original WST matrix $S_{M,N1}$. Accuracies equal to or greater than the one achieved using $S_{M,\overline{N}1}$ are in bold.

	$\hat{S}^{N2}_{M,\overline{N}1}$										$S_{M,\overline{N}1}$
w (Q_1, Q_2)	-0.5	-0.25	-0.1	0	0.1	0.2	0.25	0.3	0.4	0.5	
(4,3)	89.5	88.5	89.0	89.5	88.5	89.5	89.5	89.5	**90.5**	**90.0**	*90.0*
(4,2)	89.0	88.0	88.5	88.0	89.5	89.0	87.5	89.0	90.0	90.0	*90.5*
(3,2)	88.5	88.0	88.0	87.0	88.5	88.5	87.0	87.5	87.0	86.0	*89.0*

Table 3
Fractal dimension of the original matrix of WST coefficients (average on 800 training set signals), its subsampled version, the interpolated version via the subdivision scheme with $w = 0.3$ and the interpolated version via the subdivision scheme with $w = 0.5$.

	$S_{334,256}$	$S_{334,32}$	$\hat{S}^{32}_{334,256}(0.3)$	$\hat{S}^{32}_{334,256}(0.5)$
FD	1.8352	1.6229	1.7317	1.8163

been evaluated. As mentioned in the previous section, the 4-point interpolatory scheme in [35] has been adopted as a representative stationary subdivision scheme that is able to reproduce fractal curves; however, it is expected that any subdivision scheme working in fractal modality can be effectively adopted to interpolate the above dataset — or similar datasets containing complicated signals. With regard to the classifier, a classical SVM with polynomial kernel has been employed for simplicity.

Some results are in Table 1, where different values of the shape parameter w of the subdivision scheme have been considered. The values of w fall in the range of values for which the convergence of the subdivision scheme is guaranteed. In particular, Table 1 refers to $N2 = 32$ and reports the classification accuracy provided by $\hat{S}^{32}_{334,256}$ using different shape parameters of the adopted interpolatory scheme. The last column contains the classification accuracy using the original $S_{334,256}$ as input of the classifier. As it can be observed, the interpolated WST matrices that provide the same classification score of the original WST matrix are those generated using w values in the range that guarantees convergence to fractal curves (in particular, [.25 .5)). This result confirms that it is possible to store just a subset of data information, even for complicated signals, from which recovering a proper input for a classifier by adopting a suitable interpolation scheme. The interpolated data guarantee the same classification rate of the original data, even though they do not exactly reproduce them, as clearly shown in Fig. 6, but just some specific properties. Fig. 7 compares the fractal dimensions of the original data ($S_{334,256}$), the subsampled ones ($S_{334,32}$) along with the interpolated ones $\hat{S}^{32}_{334,256}$ with two values of the shape parameter w: 0.3 and 0.5. As can be observed, the synthesized signals have a fractal dimension very close to that of the original coefficients.

To empirically show that inequality in eq. (15) is satisfied, the fractal dimension of the original matrix of WST coefficients $S_{334,256}$, its subsampled version $S_{334,256}$, the interpolated version $\hat{S}^{32}_{334,256}(0.3)$ via the subdivision scheme with $w = 0.3$ and the interpolated version $\hat{S}^{32}_{334,256}(0.5)$ via the subdivision scheme with $w = 0.5$ has been computed and reported in Table 3. As can be observed,

$$|D(S_{334,256}) - D(\hat{S}^{32}_{334,256}(.3))| = 0.1035 < |D(S_{334,256}) - D(S_{334,32})| = 0.2123$$
$$|D(S_{334,256}) - D(\hat{S}^{32}_{334,256}(.5))| = 0.0189 < |D(S_{334,256}) - D(S_{334,32})| = 0.2123$$

and shape parameters $w > .3$ provides consistent replicates of the fractal dimension of the original data.

Test 3 This test aims at studying the dependency of the results on WST parameters settings. In particular, Table 2 contains the classification accuracy achieved using different couples of Q-factors. Comments similar to those made for the couple $(8, 1)$ still hold. However, it is worth observing that similar accuracies can be obtained also using shape parameters w not belonging to the fractal range. It is more evident when using smaller Q-factors whose wavelet kernels induce increased smoothness, causing the loss of the fractal nature of WST.

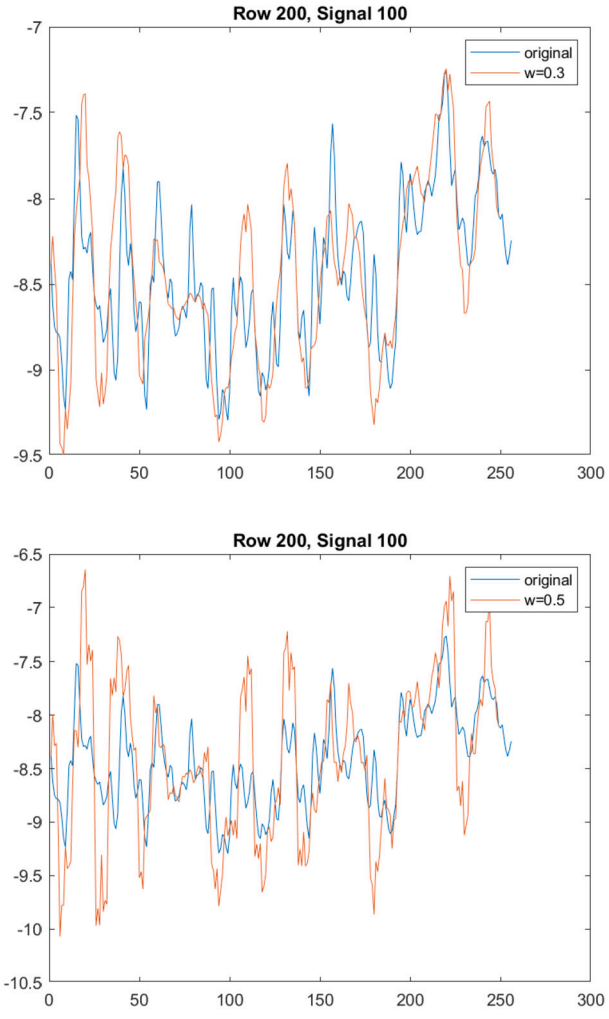

Fig. 6. Detail of original $S_{334,256}$ coefficients and synthesized $\hat{S}^{32}_{334,256}$ via the adopted subdivision scheme with $w = .3$ (*top*) and $w = .5$ (*bottom*).

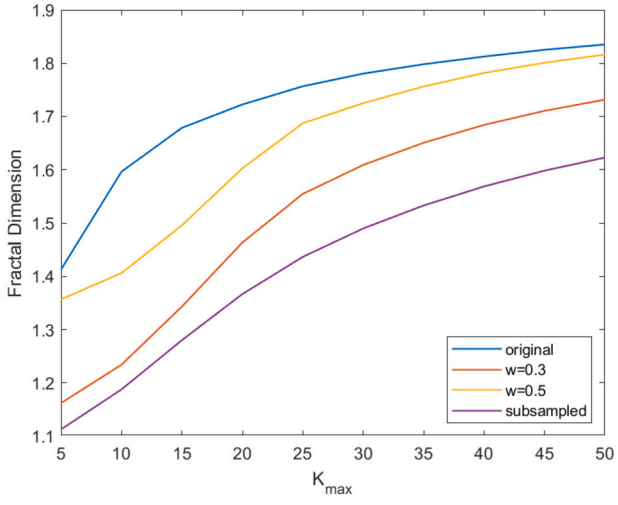

Fig. 7. Average fractal dimension of WST coefficients of audio signals in the training set versus the parameter k_{max} of the Higuchi method (the larger K_{max} the better the estimation of the fractal dimension of the curves): Original ($S_{334,256}$), subsampled by 8 ($S_{334,32}$) and interpolated ($\hat{S}^{32}_{334,256}$) with $w = .3$ and $w = .5$.

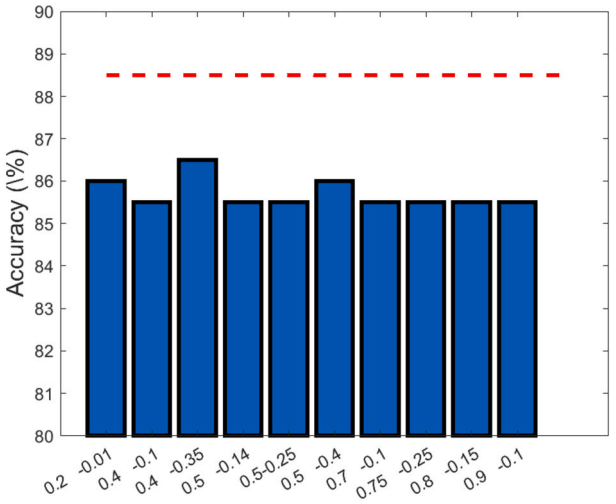

Fig. 8. Classification accuracy (%) for the GTZAN dataset using interpolated (8,1)-wavelet scattering coefficients by the generalized Chaikin corner-cutting subdivision scheme — different arbitrary corner-cutting parameters (α, β) assuring convergence to fractal curves have been considered; the dashed line indicates the expected reference classification accuracy.

Test 4 This test aims at evaluating the performance of an approximate subdivision scheme. The generalized Chaikin corner-cutting subdivision scheme with two arbitrary parameters [57] has been selected. Given the set of initial control points $\mathbf{P}^0 = \{\mathbf{P}_j^0 \in \mathbb{R}^d\}_{j=0}^{n+1}$, let $\mathbf{P}^k = \{\mathbf{P}_j^k \in \mathbb{R}^d\}_{j=0}^{2^k n+1}$ be the set of control points at level k $(k \geq 0,\ k \in \mathbb{Z})$, the scheme defines $\{\mathbf{P}_j^{k+1} \in \mathbb{R}^d\}_{j=0}^{2^{k+1} n+1}$ recursively by the following rule:

$$\begin{cases} \mathbf{P}_{2j}^{k+1} = \alpha \mathbf{P}_j^k + (1-\alpha)\mathbf{P}_{j+1}^k, & 0 \leq j \leq 2^k n, \\ \mathbf{P}_{2j+1}^{k+1} = \beta \mathbf{P}_j^k + (1-\beta)\mathbf{P}_{j+1}^k, & -1 \leq j \leq 2^k n, \end{cases}$$

where α and β are two arbitrary corner-cutting parameters. In particular, if $(\alpha, \beta) : \beta < 0,\ (\alpha - \beta) < 1,\ (\alpha + \beta) > 0$, the scheme converges to a fractal curve.

Fig. 8 shows the classification results achieved using different couples of corner-cutting parameters satisfying previous constraints. As can be observed, the classification accuracy ranges from 85.5% to 86.5%. These values are far from the expected classification accuracy (88.5%). This result highlights that interpolation condition plays a crucial role in the feature expansion process and in the reproduction of the expected fractal dimension, thanks to the preservation of original information of the data, although severely sampled. On the contrary, approximation schemes, leaving the expansion method less constrained, can cause the reproduction of similar shapes for classes of signals that can be more easily confused.

Even though future research will be devoted to investigate further this topic, these results definitively show that a 4-point interpolatory subdivision scheme is able to effectively interpolate WST coefficients having a fractal nature and paves the way to new machine learning approaches based on subdivision schemes. Based on these considerations, non-stationary schemes [58–60] will be investigated in future research; in fact, thanks to their level-dependent refinement rule, they exhibit flexibility and easier adaptation to the data and the specific task, while maintaining parametric representations. In addition, an in-depth study will be conducted on the optimal parameters setting, such as the sequence of Q-factors of the WST and the optimal interpolation grid, depending of the specific task and signal types, while also providing tighter bounds for the fractal dimension of the limit curve of the subdivision scheme. The latter study will have a twofold benefit: it will enable us to define an automatic and flexible tool for signal classification; it will contribute to providing insights into the fractal properties of subdivision schemes and related applications.

5. Conclusions

In this paper a first study concerning the joint use of sampling and interpolation of time-frequency features for signal classification purposes has been presented. To this aim wavelet scattering transform and subdivision schemes have been considered. In particular, the role of subdivision schemes in properly compensating the information lost when using sampling at each stage of the transform has been investigated. The results of the study can be summarized in a twofold contribution. First, it has been shown that, for specific signals and sequences of Q-factors, WST coefficients can have a fractal nature. This consideration is completely new and paves the way to novel approaches that deal with this kind of signals. Secondly, it has been shown that a classical 4-point subdivision scheme can be used to interpolate strongly undersampled versions of signal WST; the latter can be then used as input to a standard classifier, achieving classification accuracy comparable to that of the original data. Obviously, the value of the shape parameter w of the 4-point subdivision scheme must be such that the scheme converges to curves of fractal type. This result is relatively important as it connects

the subdivision world to the machine learning one and opens up several ways for combining them into effective future applications. Future research will be oriented to theoretically exploring the aforementioned link, in order to formalize and generalize some existing theoretical results on subdivision schemes. Specifically, non-stationary schemes will be studied due to their level dependent refinement rule and parametric representation.

CRediT authorship contribution statement

V. Bruni: Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **F. Pelosi:** Writing – review & editing, Software, Methodology, Conceptualization. **D. Vitulano:** Writing – review & editing, Software, Methodology, Conceptualization.

Acknowledgements

This work has been partially supported by PNRR-CN1 SPOKE 6 - Multiscale modeling engineering applications B83C22002940006 under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU. This research has been accomplished within RITA (Research ITalian network on Approximation) and the Italian national research group GNCS (INdAM).

Appendix A. Proof of Proposition 3.1

Convergent subdivision schemes define IFSs whose attractor is $\mathcal{A}$ and having contractive functions $f_i(X) = X P^{-1} A_i P, \quad i = 1, ..., \bar{d}$ [51]. The matrices A_i have eigenvalues of the form $1 \geq \lambda_1 \geq \lambda_2 ....$, where 1 corresponds to the translational component of the transformation matrix, while

$$\lambda_2 = \frac{1}{2},\ \lambda_3 = 2w,\ \lambda_4 = \frac{1}{4}(1 + \sqrt{1 - 16w}),\ \lambda_5 = \frac{1}{4}(1 - \sqrt{1 - 16w}).$$

In addition, the eigenvalues of $P^{-1} A_i P$ are equal to the eigenvalues of A_i. As a result, each f_i is contractive with contractive constant equal to the spectral radius ρ_i of the matrix A_i. In particular, the maximum non unitary eigenvalue of A_i is

$$\rho_i = \begin{cases} \lambda_3 & \text{if} \quad \frac{1}{4} \leq w < \frac{1}{2} \\ \lambda_4 & \text{if} \quad -\frac{1}{2} \leq w < 0. \end{cases}$$

Since the subdivision matrix A can be partitioned into $\bar{d}$ overlapping matrices that have the same eigenvalues by construction, i.e. $\rho_1 = \rho_2 = ... = \rho_{\bar{d}}$, then eq. (13) holds. As a result, $D_{\mathcal{A}} \leq \tilde{D}$, with

$$\sum_{i=1}^{\bar{d}} \rho_i^{\tilde{D}} = 1 \quad \Leftrightarrow \quad \bar{d}\, \rho_1^{\tilde{D}} = 1 \quad \Leftrightarrow \quad \tilde{D} = -\frac{log_2(\bar{d})}{log_2(\rho_1)}.$$

To prove the lower bound, let consider the explicit form of the small edge vectors $\mathbf{E}_j^k = \mathbf{P}_j^k - \mathbf{P}_{j-1}^k$ generated at the k-th iteration of the 4SS between two initial points, as proved in [31]. In particular,

$$\mathbf{E}_j^k = \begin{cases} \mathbf{a}_{1j} \gamma_1'^k + \mathbf{a}_{2j} \gamma_2'^k + \mathbf{a}_{3j} (2w)^k + \mathbf{a}_{4j} \left(\frac{1}{2}\right)^k & w \in \left(-\frac{1}{2}, 0\right) \\ \mathbf{b}_{1j} \gamma_1''^k + \mathbf{b}_{2j} \gamma_2''^k + \mathbf{b}_{3j} (2w)^k + \mathbf{b}_{4j} \left(\frac{1}{2}\right)^k & w \in \left(\frac{1}{4}, \frac{1}{2}\right) \\ \mathbf{c}_{1j} \gamma_1'''^k + \mathbf{c}_{2j} \gamma_2'''^k + \mathbf{c}_{3j} \left(\frac{1}{2}\right)^k + \mathbf{c}_{4j} k \left(\frac{1}{2}\right)^k & w = \frac{1}{4} \end{cases}$$

where $\mathbf{a}_{lj}, \mathbf{b}_{lj}$ and $\mathbf{c}_{lj}, \quad l = 1, ...4$ are non-zero and independent of k, while depend on the tension parameter w and the initial control points. Moreover,

$$\begin{cases} \gamma'_{1,2} = \frac{1}{4}(1 \pm \sqrt{1 - 16w}),\ \frac{1}{2} < \gamma'_2 < 1,\ |\gamma'_1| < |\gamma'_2|,\ |\gamma'_2| > |2w| \\ \gamma''_{1,2} = \frac{1}{4}(1 \pm i\sqrt{1 - 16w}), \quad |\gamma''_1| < 2w \quad |\gamma''_2| < 2w \\ |\gamma_1'''| = |\gamma_2'''| = \frac{1}{2} \end{cases}$$

It is worth observing that $|\mathbf{E}_j^k|$ can be rewritten in the form $|\mathbf{E}_j^k| = \alpha^k |\tilde{\mathbf{E}}_j^k|$, i.e.,

$$|\mathbf{E}_j^k| = \begin{cases} |\gamma_2'|^k \left| \mathbf{a}_{1j} \frac{\gamma_1'^k}{\gamma_2'^k} + \mathbf{a}_{2j} + \mathbf{a}_{3j} \frac{(2w)^k}{\gamma_2'^k} + \mathbf{a}_{4j} \left(\frac{1}{2\gamma'_2}\right)^k \right| & w \in \left(-\frac{1}{2}, 0\right) \\ |2w|^k \left| \mathbf{b}_{1j} \left(\frac{\gamma''_1}{2w}\right)^k + \mathbf{b}_{2j} \left(\frac{\gamma''_2}{2w}\right)^k + \mathbf{b}_{3j} + \mathbf{b}_{4j} \left(\frac{1}{4w}\right)^k \right| & w \in \left(\frac{1}{4}, \frac{1}{2}\right) \\ \left(\frac{1}{2}\right)^k \left| \mathbf{c}_{1j} (2\gamma_1''')^k + \mathbf{c}_{2j} (2\gamma_2''')^k + \mathbf{c}_{3j} + \mathbf{c}_{4j} k \right| & w = \frac{1}{4} \end{cases}$$

The proof follows by observing that the number of boxes $B_{\mathcal{A}}(s)$, with $s = \frac{1}{2^k}$, needed to recover the graph of the attractor $\mathcal{A}$ of the IFS (limit curve generated by the 4SS) is such that

$$B_{\mathcal{A}}\left(\frac{1}{2^k}\right) \geq 2^k \sum_{j=1}^{2^k n} |\mathbf{E}_j^k|$$

where for simplicity $\mathbf{E}_j^k$ stands for the edge vector generated between all the initial control points. Hence

$$\lim_{k\to+\infty} -\frac{log_2 B_{\mathcal{A}}\left(\frac{1}{2^k}\right)}{log_2\left(\frac{1}{2^k}\right)} \geq \lim_{k\to+\infty} \frac{log_2(2^{2k} n\, n_k\, |\alpha|^k \min_j \tilde{\mathbf{E}}_j^k)}{k} \to (2-\varepsilon_k) + log_2|\alpha|,$$

where $\varepsilon = \lim_{k\to+\infty} \frac{n_k}{k}$ accounts for eventual zero-length edge vectors ($n \cdot n_k$ is the number of non-zero edge vectors). •

References

[1] H.M. Cortés Campos, J.F. Gomez-Aguilar, C.J. Zuniga-Aguilar, L.F. Avalos-Ruiz, J.E. Lavin-Delgado, Application of fractional-order integral transforms in the diagnosis of electrical system conditions, Fractals 32 (3) (2024) 1–30.
[2] D. Klepl, M. Wu, F. He, Graph neural network-based EEG classification: a survey, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 493–503.
[3] M. Valtierra-Rodriguez, D. Granados-Lieberman, J.E. Torres-Fernandez, J.R. Rodríguez-Rodríguez, J.F. Gómez-Aguilar, A new methodology for tracking and instantaneous characterization of voltage variations, IEEE Trans. Instrum. Meas. 65 (7) (2016) 1596–1604.
[4] K. Zaman, M. Sah, C. Direkoglu, M. Unoki, A survey of audio classification using deep learning, IEEE Access 11 (2023) 106620–106649.
[5] G. Ala, M.L. Di Silvestre, E. Francomano, A. Tortorici, Wavelet-based efficient simulation of electromagnetic transients in a lightning protection system, IEEE Trans. Magn. 39 (3) (2003) 1257–1260.
[6] D.L. Donoho, De-noising by soft-thresholding, IEEE Trans. Inf. Theory 42 (3) (1995) 613–627.
[7] I.W. Selesnick, Wavelets, a modern tool for signal processing, Phys. Today 60 (10) (2007) 78–79.
[8] R. Soleymani, I.W. Selesnick, D.M. Landsberger, SEDA: a tunable Q-factor wavelet-based noise reduction algorithm for multi-talker babble, Speech Commun. 96 (2018) 102–115.
[9] V. Bruni, M. Cotronei, F. Pitolli, A family of level-dependent biorthogonal wavelet filters for image compression, J. Comput. Appl. Math. 367 (2020).
[10] J.R. Razo-Hernandez, A. Mejia-Barron, D. Granados-Lieberman, M. Valtierra-Rodriguez, J.F. Gomez-Aguilar, A new phasor estimator for PMU applications: P class and M class, J. Mod. Power Syst. Clean Energy 8 (1) (2020) 55–66.
[11] T. Marchand, M. Ozawa, G. Biroli, S. Mallat, Multiscale data-driven energy estimation and generation, Phys. Rev. X 13 (4) (2023).
[12] J. Anden, S. Mallat, Deep scattering spectrum, IEEE Trans. Signal Process. 62 (16) (2014) 4114–4128.
[13] J. Bruna, S. Mallat, Invariant scattering convolution networks, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1872–1886.
[14] Charu C. Aggarwal, Neural Networks and Deep Learning, A Textbook, 2019.
[15] M.P. Deisenroth, A. Faisal, C. Soon Ong, Mathematics for Machine Learning, 2020.
[16] M. Xu, L. Duan, J. Cai, L. Chia, C. Xu, Q. Tian, HMM-based audio keyword generation, in: Kiyoharu Aizawa, Yuichi Nakamura, Shinichi Satoh (Eds.), Advances in Multimedia Information Processing – PCM 2004: 5th Pacific Rim Conference on Multimedia, 2004.
[17] Md. Sahidullah, G. Saha, Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition, Speech Commun. 54 (4) (2012) 543–565.
[18] A. Abdulsatar, V.V. Davydov, V.V. Yushkova, A.P. Glinushkin, V.Yu. Rud, Age and gender recognition from speech signals, J. Phys. Conf. Ser. 1410 (1) (2019).
[19] J. Anden, V. Lostanlen, S. Mallat, Joint time–frequency scattering, IEEE Trans. Signal Process. 67 (14) (2019) 3704–3718.
[20] J. Anden, S. Mallat, Multiscale scattering for audio classification, in: Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011, Miami, Florida, USA, October 24-28, 2011.
[21] V. Lostanlen, A. Cohen-Hadria, J. Pablo Bello, One or two frequencies? The scattering transform answers, in: 2020 28th European Signal Processing Conference (EUSIPCO), 2021.
[22] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40 (2014) 16–28.
[23] M. Cox, T. Cox, Multidimensional scaling, in: Handbook of Data Visualization, Springer Handbooks Comp. Statistics, Springer, Berlin, Heidelberg, 2008.
[24] A.J. Ferreira, M.A.T. Figueiredo, Efficient feature selection filters for high-dimensional data, Pattern Recognit. Lett. 33 (13) (2012) 1794–1804.
[25] I. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philos. Trans. A 374 (2016).
[26] Z. Liu, G. Yao, Q. Zhang, J. Zhang, X. Zeng, Wavelet scattering transform for ECG beat classification, Comput. Math. Methods Med. (2020).
[27] F. Rodriguez-Algarra, B.L. Sturm, Re-evaluating the scattering transform, in: Proc. of 16th International Society for Music Information Retrieval Conference (ISMIR 2015), Late-Breaking and Demo Session at: Malaga Spain, 2015.
[28] V. Bruni, M.L. Cardinali, D. Vitulano, An MDL-based wavelet scattering features selection for signal classification, Axioms 11 (8) (2022).
[29] V. Bruni, M.L. Cardinali, D. Vitulano, A short review on minimum description length: an application to dimension reduction in PCA, Entropy 24 (2) (2022).
[30] V. Bruni, G. Monteverde, D. Vitulano, An entropy-based speed up for hyperspectral data classification via CNN, in: Proc. of the 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS 2022), Rome, Italy, 2022.
[31] Y. Zhang, C. Shang, Q. Shen, Interpolating destin features for image classification, in: Proc. of the 13th UK Workshop on Computational Intelligence (UKCI), Guildford, UK, 2013.
[32] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 2008.
[33] V. Bruni, S. Marconi, D. Vitulano, Time-scale atoms chains for transients detection in audio signals, IEEE Trans. Audio Speech Lang. Process. 18 (3) (2010) 420–433.
[34] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, Phys. D: Nonlinear Phenom. 31 (2) (1988) 277–283.
[35] N. Dyn, D. Levin, J.A. Gregory, A 4-point interpolatory subdivision scheme for curve design, Comput. Aided Geom. Des. 4 (4) (1987) 257–268.
[36] H. Zheng, Z. Ye, Y. Lei, X. Liu, Fractal properties of interpolatory subdivision schemes and their application in fractal generation, Chaos Solitons Fractals 32 (2007) 113–123.
[37] Y. Laksari, H. Aubert, D.L. Jaggard, J.-Y. Tourneret, Lacunarity of fractal superlattices: a remote estimation using wavelets, IEEE Trans. Antennas Propag. 53 (4) (2005) 1358–1363.
[38] H. Aubert, D.L. Jaggard, Wavelet analysis of transients in fractal superlattices, IEEE Trans. Antennas Propag. 50 (3) (2002).
[39] M. Sabin, Recent progress in subdivision: a survey, in: N.A. Dodgson, M.S. Floater, M.A. Sabin (Eds.), Advances in Multiresolution for Geometric Modelling. Mathematics and Visualization., Springer, Berlin, Heidelberg, 2005.

[40] M. Sabin, Analysis and Design of Univariate Subdivision Schemes, Springer, 2010.
[41] J. Warren, H. Weimer, Subdivision Methods for Geometric Design: A Constructive Approach, Morgan Kaufmann Publishers Inc., San Francisco, 2001.
[42] S. Mallat, Group invariant scattering, Commun. Pure Appl. Math. 65 (10) (2012) 1331–1398.
[43] G. Chaikin, An algorithm for high speed curve generation, Comput. Graph. Image Process. 3 (1974) 346–349.
[44] E. Catmull, J. Clark, Recursively generated B-spline surfaces on arbitrary topological meshes, Comput. Aided Des. 10 (6) (1978) 350–355.
[45] D. Doo, M. Sabin, Behaviour of recursive division surfaces near extraordinary points, Comput. Aided Des. 10 (6) (1978) 356–360.
[46] G. Mustafa, F. Khan, A new 4-point quaternary approximating subdivision scheme, Abstr. Appl. Anal. 1 (14) (2009).
[47] K.P. Ko, B.G. Lee, G.J. Yoon, A ternary 4-point approximating subdivision scheme, Appl. Math. Comput. 190 (2) (2007) 1563–1573.
[48] N. Dyn, J.A. Gregory, D. Levin, A butterfly subdivision scheme for surface interpolation with tension control, ACM Trans. Graph. 9 (2) (1990) 160–169.
[49] N. Dyn, Subdivision schemes in computer-aided geometric design, in: W. Light (Ed.), Advances in Numerical Analysis, vol. 2, 1992, pp. 36–104.
[50] Y. Hu, H. Zheng, J. Geng, Calculation of dimensions of curves generated by subdivision schemes, Int. J. Comput. Math. 96 (6) (2019) 1278–1291.
[51] S. Schaefer, D. Levin, R. Goldman, Subdivision schemes and attractors, in: Proc. of the Eurographics Symposium on Geometry Processing, Vienna, Austria, 2005.
[52] B. Mandelbrot, How long is the coast of Britain? Statistical self-similarity and fractional dimension, Science 156 (3775) (1967) 636–638.
[53] B. Mandelbrot, The Fractal Geometry of Nature, Macmillan, 1983.
[54] K. Falconer, Fractal Geometry, Wiley, 2003.
[55] M. Barnsley, Fractals Everywhere, Academic Press, 1993.
[56] G. Tzanetakis, P. Cook, Music genre classification of audio signals, IEEE Trans. Speech Audio Process. 10 (5) (2002) 293–302.
[57] J. Wang, H. Zheng, F. Xu, D. Liu, Fractal properties of the generalized Chaikin corner-cutting subdivision scheme, Comput. Math. Appl. 61 (2011) 21197.
[58] C. Conti, N. Dyn, Non-stationary subdivision schemes: state of the art and perspectives, in: G.E. Fasshauer, M. Neamtu, L.L. Schumaker (Eds.), Approximation Theory XVI. AT 2019, in: Springer Proceedings in Mathematics & Statistics, vol. 336, 2021.
[59] M. Charina, C. Conti, N. Guglielmi, v. Protasov, Limits of level and parameter dependent subdivision schemes: a matrix approach, Appl. Math. Comput. 272 (2016) 20–27.
[60] C. Conti, L. Romani, J. Yoon, Approximation order and approximate sum rules in subdivision, J. Approx. Theory 207 (2016) 380–401.