Optimal Representation of Piecewise Hölder Smooth Bivariate Functions by the Easy Path Wavelet Transform

Gerlind Plonka[1], Armin Iske[2], and Stefanie Tenorth[1]

[1] Institute for Numerical and Applied Mathematics, University of Göttingen, Lotzestr. 16-18, 37083 Göttingen, Germany
{plonka,s.tenorth}@math.uni-goettingen.de

[2] Department of Mathematics, University of Hamburg, 20146 Hamburg, Germany
iske@math.uni-hamburg.de

Abstract

The *Easy Path Wavelet Transform* (EPWT) [26] has recently been proposed by one of the authors as a tool for sparse representations of bivariate functions from discrete data, in particular from image data. The EPWT is a locally adaptive wavelet transform. It works along pathways through the array of function values and it exploits the local correlations of the given data in a simple appropriate manner. In this paper, we aim to provide a theoretical understanding of the performance of the EPWT. In particular, we derive conditions for the path vectors of the EPWT that need to be met in order to achieve optimal N-term approximations for piecewise Hölder smooth functions with singularities along curves.

Key words. sparse data representation, wavelet transform along pathways, N-term approximation

AMS Subject classifications. 41A25, 42C40, 68U10, 94A08

1 Introduction

During the last few years, there has been an increasing interest in efficient representations of large multi-dimensional data, especially for signals. In the one-dimensional case, wavelets are particularly efficient to represent piecewise smooth signals with point singularities. In two dimensions, however, tensor product wavelet bases are no longer optimal for the representation of piecewise smooth functions with discontinuities along curves.

Within the last few years, more sophisticated methods were developed to design approximation schemes for efficient representations of two-dimensional data, in particular for images, where correlations along curves are essentially taken into account to capture the geometry of the given data. Curvelets [2, 3], shearlets [13, 14] and directionlets [38] are examples for non-adaptive highly redundant function frames with strong anisotropic directional selectivity.

For piecewise Hölder smooth functions of second order with discontinuities along C^2-curves, Candès and Donoho [2] proved that a best approximation f_N to a given function

f with N curvelets satisfies the asymptotic bound

$$\|f - f_N\|_2^2 \leq C\, N^{-2}\, (\log_2 N)^3,$$

whereas a (tensor product) wavelet expansion leads to asymptotically only $\mathcal{O}(N^{-1})$ [22]. Up to the $(\log_2 N)^3$ factor, this curvelet approximation result is optimal (see [7, Section 7.4]). A similar estimate has been achieved by Guo and Labate [13] for shearlet frames. These results, however, are not adaptive with respect to the assumed regularity of the target function, and so they cannot be applied to images of less regularity, i.e., images which are not at least piecewise C^2 with discontinuities along C^2-curves. Just recently, the N-term approximation properties of compactly supported shearlet frames have been studied also in the 3D case, [20]. Using two smoothness parameters α and β in $(1, 2]$ controlling the classical as well as the anisotropic smoothness, nearly optimal approximation rates are proven in [20]. In [12], a general framework called "parabolic molecules" has been presented covering most of the curvelet- and shearlet-like constructions and showing that these systems all provide the same approximation results.

In the bivariate case, for piecewise Hölder smooth functions of order $\alpha \neq 2$, one may rather adapt the approximation scheme to the image geometry instead of fixing a basis or a frame beforehand to approximate f. During the last few years, several different approaches were developed for doing so [1, 5, 6, 8, 9, 15, 19, 21, 23, 26, 27, 28, 30, 31, 35]. In [21], for instance, bandelet orthogonal bases and frames are introduced to adapt to the geometric regularity of the image. Due to their construction, the utilized bandelets are anisotropic wavelets that are warped along a geometrical flow to generate orthonormal bases in different bands. LePennec and Mallat [21] showed that their bandelet dictionary yields asymptotically optimal N-term approximations, even in more general image models, where the edges may also be blurred.

Further examples for geometry-based image representations are the nonlinear edge-adapted (EA) multiscale decompositions in [1, 15] (and references therein) based on ENO reconstructions. We remark that the resulting ENO-EA schemes lead to an optimal N-term approximation, yielding $\|f - f_N\|_2^2 \leq C\, N^{-2}$ for piecewise C^2-functions with discontinuities along C^2-curves. Moreover, unlike previous non-adaptive schemes, the ENO-EA multiresolution techniques provide optimal approximation results also for BV-spaces and L^p spaces, see [1].

In [26], a new locally adaptive discrete wavelet transform for sparse image representations, termed *Easy Path Wavelet Transform* (EPWT), has been proposed by one of the authors. The EPWT works along pathways through the array of function values, where it essentially exploits the local correlations of image values in an appropriate manner. We remark that the EPWT is not restricted to a regular (two-dimensional) grid of image pixels, but it can be extended, in a more general setting, to scattered data approximation in higher dimensions. In [27], the EPWT has been applied to data representations on the sphere. In the implementation of the EPWT, one needs to work with suitable data structures to efficiently store the path vectors that need to be accessed during the performance of the EPWT reconstruction. To reduce the resulting adaptivity costs, we have proposed a hybrid method for smooth image approximations in [28], where an efficient edge representation by the EPWT is combined with favorable properties of the biorthogonal tensor product wavelet transform.

In [29], we have proven optimal N-term approximation rates for the EPWT for piecewise Hölder continuous functions of order $\alpha \in (0, 1]$. Our proof in [29] is mainly based on an adaptive multiresolution analysis structure, which is only available using piecewise constant Haar wavelets along the pathways.

But the EPWT does not necessarily need to be restricted to Haar wavelets. In fact, the numerical results in [26, 33] show a much higher efficiency for smoother wavelet bases as e.g. Daubechies D4 filters and biorthogonal 7-9 filters. These observations motivate us to study the N-term approximation for Hölder smooth functions of order $\alpha > 1$ in this paper. In particular, we will derive conditions for the path vectors leading to optimal N-term approximation by the EPWT. More precisely, we present three conditions for the "optimal" path vectors that imply optimal N-term approximations of the form

$$\|f - f_N\|_2^2 \leq C\, N^{-\alpha} \tag{1.1}$$

for the application of the EPWT to piecewise Hölder smooth functions of order $\alpha > 1$, with allowing discontinuities along smooth curves of finite length. Unfortunately, the three conditions on the path vectors are very difficult to ensure. We regard these conditions as an idealized setting which heuristically explains the performance of the EPWT with higher order wavelets, and which give us a hint on how to construct suitable path vectors. The path constructions of the "relaxed EPWT" in [26] aimed at a good compromise to produce many small wavelet coefficients, on the one hand, and low costs for path coding, on the other hand. Interestingly enough, our results in this paper show that the path constructions in [26] yield already a fairly good tradeoff to meet the optimal path vector conditions. In our previous paper [16], we have applied the EPWT in image denoising, with path vectors that approximatively satisfy the optimal path vector conditions.

With using piecewise constant functions for the approximation of a bivariate function f, the EPWT yields an adaptive multiresolution analysis when relying on an adaptive Haar wavelet basis (see [26, 29]). If, however, smoother wavelet bases are utilized in the EPWT approach, such an interpretation is not obvious. In fact, while Haar wavelets admit a straightforward transfer from one-dimensional functions along pathways to bivariate Haar-like functions, we cannot rely on such simple connections between smooth one-dimensional wavelets (used by the EPWT) and a bivariate approximation of the "low-pass" function. Therefore, in this paper we will apply a suitable interpolation method, by using polyharmonic spline kernels, to represent the arising bivariate "low-pass" functions after each level of the EPWT. One key property of polyharmonic spline interpolation is polynomial reproduction of arbitrary order. We will come back to relevant approximation properties of polyharmonic splines in Section 2.

The outline of this paper is as follows. In Section 2, we first introduce the utilized function model, some issues on polyharmonic spline interpolation, and the EPWT algorithm. Then, in Section 3, we study the decay of EPWT-wavelet coefficients, where we will consider the highest level of the EPWT in detail. To achieve optimal decay results for the EPWT wavelet coefficients at all levels, we require three specific side conditions for the path vectors in the EPWT algorithm, the *region condition*, the *path smoothness condition* and the *diameter condition*, see Section 3.2. We also show, why the path smoothness condition is very restrictive. In fact, it cannot be met for the usual EPWT as described in [26] for $\alpha > 1$. Therefore, in this paper we introduce suitable *smooth path functions* in

order to derive the path vectors for the EPWT. Finally, Section 4 is devoted to the proof of asymptotically optimal N-term error estimates of the form (1.1) for piecewise Hölder smooth functions.

2 The EPWT and Polyharmonic Spline Interpolation

2.1 The Function Model

Suppose that $\Omega \subset [0,1]^2$ has a sufficiently smooth Lipschitz boundary. Further, let $F \in L^2(\Omega)$ be a piecewise smooth bivariate function, being smooth over a finite set of regions $\{\Omega_i\}_{1\le i\le K}$, where each region Ω_i has a sufficiently smooth boundary $\partial\Omega_i$. In other words, the boundary curves have bounded derivatives; in particular, they are Lipschitz and of finite length. Moreover, the set $\{\Omega_i\}_{1\le i\le K}$ is assumed to be a disjoint partition of Ω, so that

$$\bigcup_{i=1}^{K} \Omega_i = \overline{\Omega},$$

where each closure $\overline{\Omega}_i$ is a connected subset of $\overline{\Omega}$, for $i = 1, \dots, K$.

More precisely, we assume that F is Hölder smooth of order $\alpha > 0$ in each region Ω_i, $1 \le i \le K$, so that every μ-th derivative of F on Ω_i with $|\mu| = \lfloor\alpha\rfloor$ satisfies an estimate of the form

$$|F^{(\mu)}(x) - F^{(\mu)}(y)| \le C\,\|x-y\|_2^{\alpha-|\mu|} \qquad \text{for all } x, y \in \Omega_i.$$

Note that this assumption for F is equivalent to the condition that for each $x_0 \in \Omega_i$ there exists a bivariate polynomial q_α of degree $\lfloor\alpha\rfloor$ (usually the Taylor polynomial of F of degree $\lfloor\alpha\rfloor$ at $x_0 \in \Omega_i$) satisfying

$$|F(x) - q_\alpha(x-x_0)| \le C\|x-x_0\|_2^\alpha \tag{2.1}$$

for every $x \in \Omega_i$ in a neighborhood of x_0, where the constant $C > 0$ does not depend on x or x_0. But F may be discontinuous across the boundaries between adjacent regions. Note that the Hölder space $C^\alpha(\Omega_i)$ of order $\alpha > 0$, being equipped with the norm

$$\|F\|_{C^\alpha(\Omega_i)} := \|F\|_{C^{\lfloor\alpha\rfloor}(\Omega_i)} + \sum_{|\mu|=\lfloor\alpha\rfloor} \sup_{x\ne y} \frac{|F^{(\mu)}(x) - F^{(\mu)}(y)|}{\|x-y\|_2^{\alpha-\lfloor\alpha\rfloor}}$$

coincides with the Besov space $B^\alpha_{\infty,\infty}(\Omega_i)$, when α is not an integer. Here, we use the $C^{\lfloor\alpha\rfloor}(\Omega_i)$ norm

$$\|F\|_{C^{\lfloor\alpha\rfloor}(\Omega_i)} := \sup_{x\in\Omega_i} |F(x)| + \sum_{|\mu|=\lfloor\alpha\rfloor} \sup_{x\in\Omega_i} |\partial^\mu F(x)|,$$

see e.g. [4, Chapter 3.2]. In order to be able to apply the known error estimates and boundedness estimates for polyharmonic spline interpolation within the next sections, we slightly strengthen the smoothness condition for F by assuming the F on Ω_i is also contained in the Sobolev space $W_2^{\alpha+1}(\Omega_i)$, where for $\alpha = m \in \mathbb{N}$, $W_2^{m+1}(\Omega_i)$ consists of

all functions $f \in L^2(\Omega_i)$ having distributional derivatives $D^\mu f$, $|\mu| \leq m+1$, in $L^2(\Omega_i)$, and where for $\alpha \in (m, m+1)$ we use

$$\|f\|_{W_2^{\alpha+1}(\mathbb{R}^2)} = \left(\int_{\mathbb{R}^2} (1 + \|\omega\|_2^2)^{\alpha+1} |\hat{f}(\omega)|^2 d\omega \right)^2,$$

and

$$\|f\|_{W_2^{\alpha+1}(\Omega_i)} = \inf\{\|g\|_{W_2^{\alpha+1}(\mathbb{R}^2)} : g \in W_2^{\alpha+1}(\mathbb{R}^2),\, g|_{\Omega_i} = f\}.$$

Here the Fourier transform of f is given by $\hat{f}(\omega) := \int_{\mathbb{R}^2} f(x) e^{-i\omega^T x} dx$. The Sobolev embedding theorems state that $W_2^{\alpha+1} \subset C^\alpha$, see e.g. [34]. Moreover, the Sobolev embedding is usually sharp, i.e., the functions in $W_2^{\alpha+1}$ possessing a pointwise Hölder exponent greater than α form a so-called HR-small set of $W_2^{\alpha+1}$, where this set is both Haar null and of first Baire category, see [11].

Now by quasi-uniform sampling, the bivariate function F is assumed to be given by its function values taken at a finite grid I_{2J}. For a suitable integer $J > 1$, let $\{F(y)\}_{y \in I_{2J}}$ be the given samples of F, where $I_{2J} \subset \Omega$ with $\#I_{2J} = 2^{2J}$, and

$$\min_{y_1, y_2 \in I_{2J}} \|y_1 - y_2\|_2 \geq \kappa_1 2^{-J}, \qquad \sup_{x \in \Omega} \inf_{y \in I_{2J}} \|x - y\|_2 \leq 2^{-(J+1/2)} \tag{2.2}$$

with a positive constant κ_1 being independent of J. Moreover, let

$$\Gamma_i^{2J} := \{y \in I_{2J} : y \in \Omega_i\} \qquad \text{for } 1 \leq i \leq K \tag{2.3}$$

be the set of grid points that are contained in the regions Ω_i, for $1 \leq i \leq K$, where we assume J to be large enough, so that the sample points in Γ_i^{2J} are sufficiently dense, which in turn leads to sufficiently accurate approximations to F on the subdomains Γ_i by polyharmonic splines interpolation (see Subsection 2.2). Obviously,

$$\bigcup_{i=1}^{K} \Gamma_i^{2J} = I_{2J},$$

and for the size $\#\Gamma_i^{2J}$ of Γ_i^{2J} we have $\#\Gamma_i^{2J} \leq \#I_{2J} = 2^{2J}$ for every i with $1 \leq i \leq K$.

Remark. Note that our assumption concerning the Lipschitz smoothness of the subdomains' Ω_i boundaries is quite restrictive. Indeed, for image partitions, this condition is rather unlikely to be satisfied. In that case, the boundary curves of the subdomains are usually only piecewise smooth.

2.2 Polyharmonic Spline Interpolation

Next we compute a (piecewise) sufficiently smooth approximation to F from its given samples. To this end, we construct a suitable interpolant F^{2J} satisfying the interpolation conditions

$$F^{2J}(y) = F(y) \qquad \text{for all } y \in I_{2J}. \tag{2.4}$$

This is accomplished by the application of polyharmonic spline interpolation separately in each individual region Ω_i, in order to first obtain, for every index i, $1 \leq i \leq K$, a polyharmonic spline interpolant F_i^{2J} satisfying

$$F_i^{2J}(y) = F(y) \qquad \text{for all } y \in \Gamma_i^{2J} \tag{2.5}$$

at the interpolation points $\Gamma_i^{2J} \subset \Omega_i$. For the required (global) interpolant F^{2J}, we then let

$$F^{2J}(x) := \sum_{i=1}^{K} F_i^{2J}(x)\chi_{\Omega_i}(x) \qquad \text{for } x \in \Omega \tag{2.6}$$

to satisfy the interpolation conditions in (2.4), where χ_{Ω_i} is the characteristic function of Ω_i.

Recall that polyharmonic splines, due to Duchon [10], are suitable tools for *multivariate* interpolation from scattered data. For further details on polyharmonic splines, especially for relevant aspects concerning their local approximation properties, we refer to [18, Section 3.8]. We apply the polyharmonic spline interpolation scheme in two dimensions. Then the interpolant F_i^{2J} in (2.5) is assumed to have the form

$$F_i^{2J}(x) = \sum_{y \in \Gamma_i^{2J}} c_y^i \, \phi_\alpha \left(\|x - y\|_2 \right) + p_m^i(x), \tag{2.7}$$

where, for $\alpha \in (m, m+1)$, $m \in \mathbb{N}$, we use the fixed *polyharmonic spline* kernel $\phi_\alpha(r) = r^{2\alpha}$, and $p_m^i \in \mathcal{P}_m$ is a polynomial in the linear space $\mathcal{P}_m$ of all bivariate polynomials of degree at most m. By Theorem 8.16 in [37],

$$\hat{\phi}_\alpha(\omega) = \frac{2^{2\alpha+1}\Gamma(\alpha+1)}{\Gamma(-\alpha)} \|w\|_2^{-2\alpha-2},$$

and hence ϕ_α is of order $\lceil \alpha \rceil = m + 1$. For $\alpha = m \in \mathbb{N}$, we choose $\phi_\alpha(r) = r^{2m} \log(r)$, such that $\phi_\alpha = \phi_m$ is again of order $m + 1$ with the generalized Fourier transform

$$\hat{\phi}_m(\omega) = (-1)^{m+1} 2^{2m} \, \Gamma(m+1) \, m! \, \|w\|_2^{-2m-2},$$

see [37], Theorem 8.17.

Note that the interpolant F_i^{2J} in (2.7) has $\#\Gamma_i^{2J} + \dim(\mathcal{P}_m)$ parameters, i.e., the $\#\Gamma_i^{2J}$ coefficients c_n^i in its major part and another $\dim(\mathcal{P}_m) = (m+1)(m+2)/2$ parameters in its polynomial part. However, the interpolation problem (2.5) yields only $\#\Gamma_i^{2J}$ linear conditions on the $\#\Gamma_i^{2J} + (m+1)(m+2)/2$ parameters of F_i^{2J}. To eliminate additional degrees of freedom, we require another set of $(m+1)(m+2)/2$ linear constraints on the coefficients c_n^i, as given by the *vanishing moment conditions*

$$\sum_{y \in \Gamma_i^{2J}} c_n^i \, p(y) = 0 \qquad \text{for all } p \in \mathcal{P}_m. \tag{2.8}$$

To compute the coefficients of the polyharmonic spline interpolant F_i^{2J}, this then amounts to solving a square linear system with $\#\Gamma_i^{2J} + \dim(\mathcal{P}_m)$ equations, (2.5) and (2.8), for $\#\Gamma_i^{2J} + \dim(\mathcal{P}_m)$ unknowns, given by the parameters of F_i^{2J}.

According to the seminal work of Michelli [24] on (conditionally) positive definite functions, this square linear equation system has a unique solution, provided that the set of interpolation points are $\mathcal{P}_m$-regular, i.e., for $p \in \mathcal{P}_m$ we have the implication

$$p(y) = 0 \text{ for all } y \in \Gamma_i^{2J} \qquad \Longrightarrow \qquad p \equiv 0, \tag{2.9}$$

so that every polynomial in $\mathcal{P}_m$ can uniquely be reconstructed from its values on Γ_i^{2J}. In fact, this complies with earlier results in [10], where the well-posedness of the polyharmonic spline interpolation scheme is proven via a variational theory concerning (polyharmonic) splines minimizing rotation-invariant semi-norms in Sobolev spaces.

Therefore, we can conclude that the polyharmonic spline interpolant F_i^{2J} in (2.7), satisfying (2.5) and (2.8), is unique provided that the interpolation points Γ_i^{2J} satisfy the (rather weak) regularity conditions in (2.9). From now we will tacitly assume that the conditions in (2.9) are fulfilled. In this case, we can further conclude that the polyharmonic spline interpolation scheme achieves to reconstruct polynomials of degree m.

Concerning the *local* approximation order of polyharmonic spline interpolation, we refer to [17]. In the subsequent analysis in this paper, we require one specific approximation result for polyharmonic spline interpolation concerning its *global* approximation behaviour. From Lemma 4.1 in [25], we observe the estimate

$$\|F - F^{2J}\|_{L^2(\Omega)} \leq C_F \sum_{i=1}^{K} h_{\Omega_i}^{\alpha+1} \|F\|_{W_2^{\alpha+1}(\Omega_i)} \tag{2.10}$$

for the interpolation error in Sobolev spaces, where the *fill distance*

$$h_{\Omega_i} := \sup_{x \in \Omega_i} \inf_{y \in \Gamma_i^{2J}} \|x - y\|_2 \leq 2^{-J+1/2} \qquad \text{for } 1 \leq i \leq K$$

measures the density of the interpolation points in Ω_i. In particular, we also have

$$\|F - F^{2J}\|_{W_2^{\alpha+1}(\Omega_i)} \leq C \|F\|_{W_2^{\alpha+1}(\Omega_i)},$$

and hence

$$\|F^{2J}\|_{W_2^{\alpha+1}(\Omega_i)} \leq (C+1) \|F\|_{W_2^{\alpha+1}(\Omega_i)},$$

i.e., the interpolant F^{2J} is uniformly bounded in the $W_2^{\alpha+1}(\Omega_i)$-norm independently of $2J$, see [25], Corollary 4.3.. Finally, by a continuous embedding of $W_2^{\alpha+1}(\Omega_i)$ in $C^\alpha(\Omega_i)$, we ensure the boundedness of $\|F^{2J}\|_{C^\alpha}$ independently of J, see e.g. [32], Theorem 3.2.1.

2.3 The EPWT Algorithm

Now let us briefly recall the EPWT algorithm from our previous work [26]. To this end, let $\varphi \in C^\beta$, for $\beta \geq \alpha$, be a sufficiently smooth, compactly supported, one-dimensional scaling function, i.e., the integer translates of φ form a Riesz basis of the scaling space $V_0 := \mathrm{clos}_{L^2}\mathrm{span}\,\{\varphi(\cdot - k) : \ k \in \mathbb{Z}\}$. Further, let $\tilde{\varphi}$ be a corresponding biorthogonal and sufficiently smooth scaling function with compact support, and let ψ and $\tilde{\psi}$ be a corresponding pair of compactly supported wavelet functions. We refer to [4, Chapter 2] for a comprehensive survey on biorthogonal scaling functions and wavelet bases and summarize only the notation needed for the biorthogonal wavelet transform. For $j, k \in \mathbb{Z}$, let

$$\varphi_{j,k}(t) := 2^{j/2}\, \varphi(2^j t - k) \qquad \text{and} \qquad \psi_{j,k}(t) := 2^{j/2}\, \psi(2^j t - k),$$

likewise for $\tilde{\varphi}$ and $\tilde{\psi}$. The functions φ, $\tilde{\varphi}$ and $\psi, \tilde{\psi}$ are assumed to satisfy the refinement equations

$$\varphi(x) \ = \ \sqrt{2} \sum_n h_n \varphi(2x - n) \qquad \psi(x) = \sqrt{2} \sum_n q_n \varphi(2x - n)$$

$$\tilde{\varphi}(x) \;=\; \sqrt{2}\sum_n \tilde{h}_n\tilde{\varphi}(2x-n) \qquad \tilde{\psi}(x)=\sqrt{2}\sum_n \tilde{q}_n\tilde{\varphi}(2x-n)$$

with finite sequences of filter coefficients $(h_n)_{n\in\mathbb{Z}}$, $(\tilde{h}_n)_{n\in\mathbb{Z}}$ and $(q_n)_{n\in\mathbb{Z}}$, $(\tilde{q}_n)_{n\in\mathbb{Z}}$. By assumption, the polynomial reproduction property

$$\sum_k \langle p_m, \tilde{\varphi}_{j,k}\rangle\, \varphi_{j,k} = p_m \qquad \text{for all } j\in\mathbb{Z},$$

is satisfied for each polynomial p_m of degree less than or equal to $m=\lceil\alpha\rceil$, and so,

$$\langle p_m, \tilde{\psi}_{j,k}\rangle = 0 \qquad \text{for all } j,k\in\mathbb{Z}.$$

With these assumptions, $\{\psi_{j,k} : j,k\in\mathbb{Z}\}$ and $\{\tilde{\psi}_{j,k} : j,k\in\mathbb{Z}\}$ form biorthogonal Riesz bases of $L^2(\mathbb{R})$, i.e., for each function $f\in L^2(\mathbb{R})$, we have

$$f=\sum_{j,k\in\mathbb{Z}}\langle f,\psi_{j,k}\rangle\tilde{\psi}_{j,k}=\sum_{j,k\in\mathbb{Z}}\langle f,\tilde{\psi}_{j,k}\rangle\psi_{j,k}.$$

For a given univariate function f^j, $j\in\mathbb{Z}$, of the form $f^j(x)=\sum_{n\in\mathbb{Z}} c^j(n)\,\varphi_{j,n}$ one decomposition step of the discrete (biorthogonal) wavelet transform can be represented in the form

$$f^j(x)=f^{j-1}(x)+g^{j-1}(x),$$

where

$$f^{j-1}(x)=\sum_{n\in\mathbb{Z}} c^{j-1}(n)\,\varphi_{j-1,n} \quad \text{and} \quad g^{j-1}(x)=\sum_{n\in\mathbb{Z}} d^{j-1}(n)\,\psi_{j-1,n}$$

with

$$c^{j-1}(n)=\langle f^j,\tilde{\varphi}_{j-1,n}\rangle \qquad \text{and} \qquad d^{j-1}(n)=\langle f^j,\tilde{\psi}_{j-1,n}\rangle. \tag{2.11}$$

Conversely, one step of the inverse discrete wavelet transform yields for given functions f^{j-1} and g^{j-1} the reconstruction

$$f^j(x)=\sum_{n\in\mathbb{Z}} c^j(n)\,\varphi_{j,n} \quad \text{with } c^j(n)=\langle f^{j-1},\tilde{\varphi}_{j,n}\rangle+\langle g^{j-1},\tilde{\varphi}_{j,n}\rangle.$$

We recall that the EPWT is a wavelet transform that works along path vectors through point subsets of I_{2J}. For the characterization of suitable path vectors we first need to introduce *neighborhoods* of points in I_{2J}. For each point $y=(y_1,y_2)\in I_{2J}$, we define its neighborhood by

$$N(y):=\{x=(x_1,x_2)\in I_{2J}\setminus\{y\} : \|x-y\|_2\le 2^{-J+1/2}\},$$

where $\|x-y\|_2^2=(x_1-y_1)^2+(x_2-y_2)^2$.

Now the EPWT algorithm is performed as follows. For the application of the $2J$-th level of the EPWT we need to find a path vector $p^{2J}=(p^{2J}(n))_{n=0}^{2^{2J}-1}$ through the point set I_{2J}. This path vector is a suitable permutation of all points in I_{2J}, which can be determined by using the following strategy from [26]. Recall that $I_{2J}=\cup_{i=1}^K\Gamma_i^{2J}$, where Γ_i^{2J} determines the lattice points in Ω_i. Start with one point $p^{2J}(0)$ in Γ_1^{2J}. Now, for a given

n-th component $p^{2J}(n)$ being contained in the set Γ_i^{2J} in (2.3), for some $i \in \{1, \dots, K\}$, we choose the next component $p^{2J}(n+1)$ of the path vector p^{2J}, such that

$$p^{2J}(n+1) \in (N(p^{2J}(n)) \cap \Gamma_i^{2J}) \setminus \{p^{2J}(0), \dots, p^{2J}(n)\}, \tag{2.12}$$

i.e., $p^{2J}(n+1)$ should be a neighbor point of $p^{2J}(n)$, lying in the same set Γ_i^{2J}, that has not been used in the path, yet.

In situations where $(N(p^{2J}(n)) \cap \Gamma_i^{2J}) \setminus \{p^{2J}(0), \dots p^{2J}(n)\}$ is an empty set, the path is *interrupted*, and we need to start a *new pathway* by choosing the next component $p^{2J}(n+1)$ from $\Gamma_i^{2J} \setminus \{p^{2J}(0), \dots, p^{2J}(n)\}$. If, however, this set is also empty, we choose $p^{2J}(n+1)$ from the set of remaining points $I_{2J} \setminus \{p^{2J}(0), \dots, p^{2J}(n)\}$. For a more detailed description of the path vector construction we refer to [26].

In particular, for a suitably chosen path vector p^{2J}, the number of interruptions can be bounded by $\tilde{K} = C_1 K$, where K is the number of regions, and where the constant C_1 does not depend on J but only on the shape of the regions Ω_i, see [29]. The so obtained vector p^{2J} is composed of *connected* pathways, i.e., each pair of consecutive components in these pathways is neighboring.

Now, we consider the data vector

$$\left(c^{2J}(\ell)\right)_{\ell=0}^{2^{2J}-1} := \left(F^{2J}\left(p^{2J}(\ell)\right)\right)_{\ell=0}^{2^{2J}-1}$$

and apply one level of a one-dimensional (periodic) wavelet transform to the function values of F^{2J} along the path p^{2J}. This yields the low-pass vector $(c^{2J-1}(\ell))_{\ell=0}^{2^{2J-1}-1}$ and the vector of wavelet coefficients $(d^{2J-1}(\ell))_{\ell=0}^{2^{2J-1}-1}$ according to the formulae in (2.11). Due to the piecewise smoothness of F^{2J} along the path vector p^{2J}, it follows that most of the wavelet coefficients in d^{2J-1} are small, whereas the wavelet coefficients corresponding to an interruption of the path (within one region or from one region to another) may possess significant amplitudes. We remark already here that, by contrast to the original EPWT algorithm described in this subsection, for the case of piecewise Hölder smooth images of order $\alpha > 1$, we will introduce a smooth *path function* $\tilde{p}^{2J}$ in Section 3.1. This path function will determine the components of the path vector p^{2J} and is assumed to have bounded derivatives almost everywhere. Note that for $\alpha > 1$, wavelet coefficients of (arbitrary) large magnitude may occur in situations where the derivatives of the path function are not uniformly bounded, as e.g. at tight turns of $\tilde{p}^{2J}$.

As regards the next level of the EPWT, the path vector p^{2J} yields a new subset of points

$$\Gamma^{2J-1} := \left\{p^{2J}(2\ell) : \ell = 0, \dots, 2^{2J-1} - 1\right\} = \bigcup_{i=1}^{K} \Gamma_i^{2J-1},$$

where $\Gamma_i^{2J-1} := \{p^{2J}(2\ell) : p^{2J}(2\ell) \in \Gamma_i^{2J}\}$. At level $j = 2J-1$, we first locate a second connected path vector $p^{2J-1} = (p^{2J-1}(\ell))_{\ell=0}^{2^{2J-1}-1}$ through Γ^{2J-1}, i.e., the entries of p^{2J-1} form a permutation of the points in Γ^{2J-1}. Similarly as before, we require that $p^{2J-1}(n)$ and $p^{2J-1}(n+1)$ are neighbors lying in the same point set Γ_i^{2J-1}. Here, $p^{2J-1}(n)$ and $p^{2J-1}(n+1)$ are said to be neighbors, i.e., $p^{2J-1}(n+1) \in N(p^{2J-1}(n))$, iff

$$\left\|p^{2J-1}(n) - p^{2J-1}(n+1)\right\|_2 \leq 2^{-J+1}.$$

Again, the number of path interruptions can be bounded by C_1K, where C_1 does not depend on J. Then we apply one level of the one-dimensional wavelet transform to the permuted data vector $(c^{2J-1}(per^{2J-1}(\ell)))_{\ell=0}^{2^{2J-1}-1}$, where the permutation per^{2J-1} is defined by $per^{2J-1}(\ell) := k$ iff $p^{2J-1}(\ell) = p^{2J}(2k)$ for $\ell, k = 0, \dots 2^{2J-1} - 1$. We obtain the low-pass vector $(c^{2J-2}(\ell))_{\ell=0}^{2^{2J-2}-1}$ and the vector $(d^{2J-2}(\ell))_{\ell=0}^{2^{2J-2}-1}$ of wavelet coefficients.

We continue by iteration over the remaining levels j for $j = 2J-2, \dots, 1$, where at every level j we first construct a path $p^j = (p^j(\ell))_{\ell=0}^{2^j-1}$ through the set

$$\Gamma^j := \{p^{j+1}(2\ell) : \ell = 0, \dots, 2^j - 1\} = \bigcup_{i=0}^{K} \Gamma_i^j$$

with applying similar strategies as described above. Here, $p^j(n)$ and $p^j(n+1)$ are called neighbors, iff

$$\left\|p^j(n) - p^j(n+1)\right\|_2 \leq D2^{-j/2}, \tag{2.13}$$

where $D \geq \sqrt{2}$ is a suitably determined constant (in the above description of p^{2J} and p^{2J-1} we have chosen $D = \sqrt{2}$). Then we apply the wavelet transform to the permuted vector $(c^j(per^j(\ell))_{\ell=0}^{2^j-1}$, yielding c^{j-1} and d^{j-1}.

Example. In this example, we explain the construction of the path vectors through the remaining data points with the low-pass values by a toy example. To this end, let $\Omega = [0,1)^2$ be divided into only two regions, Ω_1 and Ω_2. The function F is assumed to be Hölder continuous in each of these regions, but may be discontinuous across the curve separating the two regions Ω_1 and Ω_2. In our toy example, we have $J = 3$, i.e., an 8×8 image with 64 data values. At the highest level of the EPWT, we choose a path p^6 through the underlying point set $I_6 = \{(2^{-3}n_1, 2^{-3}n_2) : n_1, n_2 = 0, \dots, 7\} = \Gamma_1^6 \cup \Gamma_2^6$, such that each two consecutive components in the path are neighbors. We first pick all points in Γ_1^6, before jumping to Γ_2^6, see Figure 1(a). For the path construction at the next level, we first determine the set $\Gamma^5 = \Gamma_1^5 \cup \Gamma_2^5$ (containing only each second component of p^6), see Figure 1(b), before we construct a path according to the above description. Figures 1(c) and 1(d) show the sets Γ^4 and Γ^3 along with their corresponding path vectors. Besides, we tried to find the path vectors in a way such that the angles formed by the polygonal line of the path are as large as possible. In fact, the polygonal lines for p^6 and p^5 do not contain angles being smaller than 90 degrees, while the polygonal line of p^4 possesses two such angles, one in each region. In fact, locally straight polygonal lines of the path vectors give rise to obtain better smoothness bounds for the path functions $\tilde{p}^j$ that we will introduce in Section 3. In this example, we have

$$\begin{aligned} \|p^6(n+1) - p^6(n)\|_2 &\leq \sqrt{5} \leq D, & \|p^5(n+1) - p^5(n)\|_2 &\leq \sqrt{5} \leq \sqrt{2}D, \\ \|p^4(n+1) - p^4(n)\|_2 &\leq 4 \leq 2D, & \|p^3(n+1) - p^3(n)\|_2 &\leq \sqrt{10} \leq \sqrt{8}D \end{aligned}$$

(with one path interruption at each level for the jump from one region to the other), so that the path construction satisfies the above requirements with $D = \sqrt{5}$. This simple example also illustrates that the path construction leads at each level to point sets Γ_i^j with quasi-uniformly distributed points. See Section 3.2 for more explanation about the used quasi-uniformity.

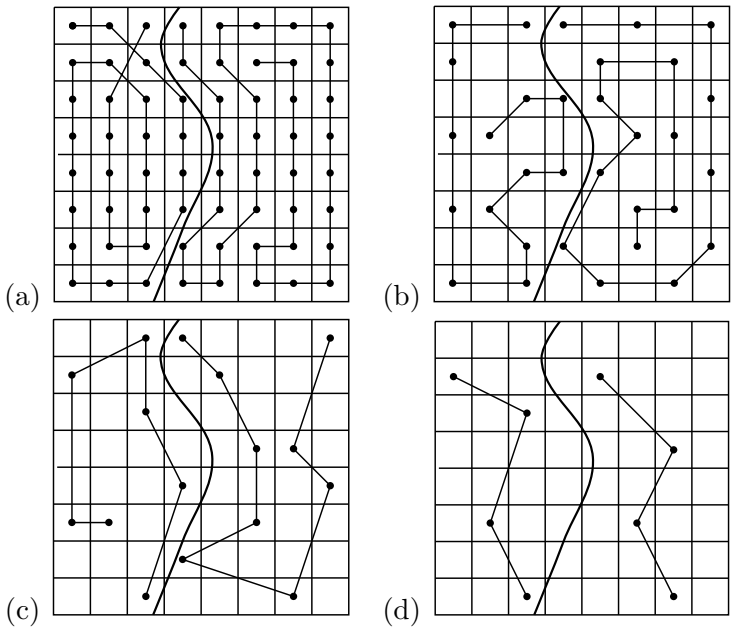

Figure 1: Path construction. (a) level six, (b) level five, (c) level four, (d) level three.

Remark 1. Considering the above strategy of the EPWT algorithm, it is heuristically clear that we are able to reduce the number of significant wavelet coefficients to a multiple of the number K of regions, where the target function F is smooth. Indeed, only when the path skips from one region to another, a finite number of significant wavelet coefficients will occur. This is in contrast to the usual tensor product wavelet transform, where the number of significant wavelet coefficients is usually related to the total length of the "smooth regions" boundaries and hence depends on the level j of the wavelet transform.

Remark 2. As shown in [26], the choice of the path vectors in each level is crucial for the performance of the approximation by means of the EPWT. In particular, it is not a good idea to take simply the same path vector $p^j(\ell) = p^{j+1}(2\ell)$ at each level, since one only exploits the data correlation along this one path and the EPWT is reduced to a one-dimensional transform. In Section 3.2, we will determine the so called *diameter condition* that requires the inequality (2.13) for neighboring components in the path construction thereby avoiding the above mentioned simple path choice. Our numerical experiments with the EPWT imply that the constant D should be chosen rather small, e.g. $D \approx 2$, see [33]. On the other hand, a too small D induces a small number of neighborhood points and hence yields a higher number of path interruptions. In Section 3.2, we will show that the inequality (2.13) for path construction implies a quasi-uniform distribution of values in Γ^j at each level of the EPWT.

Remark 3. Note that the implementation of the EPWT algorithm, as described above, is rather straightforward. In the original version of the EPWT, we first determine, at each level, a path vector p^j, before we apply a wavelet filter bank to the one-dimensional data vector that is ordered according to the path vector. In contrast to that strategy (from our previous work), we now consider using a slightly different approach to obtain the theoretical estimates of our present paper. To this end, we derive the path vector p^j

by first determining a piecewise smooth path function $\tilde{p}^j$, whose construction is detailed in the following of this paper. We also consider slightly different data vectors $c_p^j(\ell)$ in the theoretical estimates of Sections 3 and 4, where we will use L^2-projection operators determined by the dual scaling and wavelet functions, $\tilde{\varphi}$ and $\tilde{\psi}$. Further, we will employ polyharmonic spline interpolation for analytical purpose.

Remark 4. Note that the components of the path vector p^j lie in I_{2J} with containing 2d entries, i.e., $p^j \in \mathbb{R}^{I_{2J}\times 2}$. This is in contrast to the notation in [26]. Further, unlike in [26], we do no longer consider *index sets* but define a neighborhood of points by the Euclidean distance.

3 Decay of Wavelet Coefficients generated by the EPWT

Before we turn to the technical details, let us first sketch the basic ideas of the proof for optimal N-term approximations by the EPWT, where we use slightly changed notations.

As explained in the Subsection 2.2, we consider applying polyharmonic spline interpolation, from given image values $F(y)$, separately in the individual domains Ω_i, $i = 1, \dots, K$. We assume that F is Hölder smooth of order α on each Ω_i, so that the polyharmonic spline interpolant F^{2J} in (2.6), being Hölder smooth of order at least α, reconstructs bivariate polynomials of degree $m = \lceil \alpha \rceil$. Further, we apply a generalized notion of a path $\tilde{p}$ given as a sufficiently smooth planar parameter curve.

At the $2J$-th level of the EPWT, we first determine a (piecewise) smooth path function $\tilde{p}^{2J} : [0,1] \to \Omega$, such that each value $y \in I_{2J}$ can be approximated at sufficient accuracy by $\tilde{p}^{2J}(\frac{\ell}{2^{2J}})$, for some $\ell \in \{0, \dots, 2^{2J} - 1\}$. Consecutive values $\tilde{p}^{2J}(\frac{\ell}{2^{2J}})$ and $\tilde{p}^{2J}(\frac{\ell+1}{2^{2J}})$ should approximate neighboring values in

$$\tilde{\Gamma}^{2J} := I_{2J}$$

lying in the same set Ω_i (up to $\mathcal{O}(K)$ exceptions). Now, a path vector $p^{2J} \in \mathbb{R}^{2^{2J}\times 2}$ through all values of I_{2J} is determined by the permutation given by the order in which $\tilde{p}^{2J}(\ell/2^{2J})$, $\ell = 0, \dots, 2^{2J} - 1$, approximates the values in I_{2J}.

The path function $\tilde{p}^{2J}$ should be constructed such that only at most $C_1 K$ discontinuities (incurred by possible transitions from one region to another) may occur. In this way, we can bound the number of "interruptions" in the path vector p^{2J} by $C_1 K$.

Next, we consider a one-dimensional function $f^{2J}(t) = \sum_{k=0}^{2^{2J}-1} c_p^{2J}(k)\, \varphi_{2J,k}(t)$, $t \in [0,1)$, which approximates the piecewise smooth one-dimensional and scaled restriction of F resp. F^{2J} along the path $\tilde{p}^{2J}$ satisfying

$$|F(\tilde{p}^{2J}(2^{-2J}\ell)) - f^{2J}(2^{-2J}\ell)| \lesssim 2^{-J\alpha}.$$

Then, we apply one level of a smooth wavelet transform to f^{2J}.

In this way, significant wavelet coefficients may only occur at a finite number of locations on the interval $[0,1)$ that correspond to discontinuities or to tight turns of the path function $\tilde{p}^{2J}$. However, the number of such significant coefficients does not depend on 2^J but on the number of regions, K. Therefore, with the performance of one level of the (periodic) wavelet transform, we will find that most of the wavelet coefficients of $f^{2J} = f^{2J-1} + g^{2J-1}$ occurring in the wavelet part g^{2J-1}, are small.

At the next j-th level of the EPWT, we consider the set

$$\tilde{\Gamma}^j := \{\tilde{p}^{j+1}(2^{-j}\ell) : \ell = 0, \dots, 2^j - 1\} = \cup_{i=1}^K \tilde{\Gamma}_i^j.$$

By construction, we have $\#\tilde{\Gamma}^j = 2^j$. Observe that $\tilde{\Gamma}^j$ are point sets in Ω. Again, we then construct a smooth path function $\tilde{p}^j$ that approximates the values in $\tilde{\Gamma}^j$ in a suitable order and thus determines the path vector p^j whose components are the values in $\tilde{\Gamma}^j$ in the order in which they are approximated by $\tilde{p}^j$. To obtain sufficiently accurate polyharmonic spline interpolations F^j to F with $F^j(y) = F(y)$ for $y \in \tilde{\Gamma}^j$, and to apply the same arguments as at the highest level, three specific conditions on the path functions $\tilde{p}^j$ need to be satisfied. We can briefly explain these conditions, termed *region condition*, *path smoothness condition* and *diameter condition*, as follows (for more details on these conditions we refer to Subsection 3.2).

Firstly, the *region condition* requires that the path function should prefer to approximate all values of one region set $\tilde{\Gamma}_i^j$, before proceeding with the values of the next region. In other words, the path vector should prefer to traverse the points belonging to one region set $\tilde{\Gamma}_i^j$, before "jumping" to another region, where jumps within one region should be avoided. Assuming further that the path function is smooth with suitably bounded derivatives (*path smoothness condition*), we can optimally exploit the smoothness of the function F along the path. Finally, the *diameter condition* ensures a quasi-uniform distribution of remaining points in each $\tilde{\Gamma}^j$, and therefore leads to a sufficiently accurate polyharmonic spline interpolation F^j at each level of the EPWT.

We remark that the region condition and the diameter condition can be forced by using the strategies for the path construction as proposed in Subsection 2.2. The path smoothness condition cannot be satisfied by the original path construction methods in [26]. But since the numerical evaluations contain only a few levels of the EPWT, the path smoothness condition can be forced by avoiding "small angles" in the path vector and by preferring "path snakes", see e.g. examples in [26]. The three path conditions will allow us to estimate the EPWT wavelet coefficients similarly as for one-dimensional piecewise smooth functions with a finite number of singularities to finally obtain an optimal N-term approximation using only the N most significant EPWT wavelet coefficients for the image reconstruction.

3.1 The Highest Level of the EPWT

Let us now explain the $2J$-th level of the EPWT in detail. The performance of the subsequent levels of the EPWT and the corresponding estimates are then derived in a similar manner.

We consider a sufficiently smooth parameter curve $\tilde{p}^{2J}(t)$, $t \in [0,1]$, through Ω approximating the values in $\tilde{\Gamma}^{2J}$ in a certain order and the corresponding path vector p^{2J} of length 2^{2J} with values in I_{2J}, where we assume that

$$|\tilde{p}^{2J}(2^{-2J}\ell) - p^{2J}(\ell)| \leq \mu 2^{-J} \qquad \ell = 0, \dots, 2^{2J} - 1, \tag{3.1}$$

and where μ is a fixed small constant (independent of J) with assuming $\tilde{p}^{2J}(t) \in \Omega_i$, for $t \in [\ell/2^{2J}, (\ell+1)/2^{2J}]$, if $\tilde{p}^{2J}(2^{-2J}\ell)$ and $\tilde{p}^{2J}(2^{-2J}(\ell+1))$ are in Ω_i. Now, we regard the function $\tilde{f}^{2J}$ that is defined by the one-dimensional restriction of F^{2J} along the curve $\tilde{p}^{2J}$,

$$\tilde{f}^{2J}(t) = F^{2J}\left(\tilde{p}^{2J}(t)\right) \qquad \text{for } t \in [0,1].$$

If the path function $\tilde{p}^{2J}$ crosses over from one region Ω_i to another, there may occur a discontinuity of $\tilde{f}^{2J}$ caused by a discontinuity of F^{2J}, i.e., there are indices ℓ and $\ell+1$ in $\{0,1,\ldots,2^{2J}-1\}$, where we have a discontinuity between $F^{2J}(\tilde{p}^{2J}(2^{-2J}\ell))$ and $F^{2J}(\tilde{p}^{2J}(2^{-2J}(\ell+1)))$, i.e., a discontinuity of $\tilde{f}^{2J}$ in the subinterval $[\ell/2^{2J},(\ell+1)/2^{2J}]$. Without loss of generality, we assume that we have only discontinuities of $\tilde{f}^{2J}$ caused by jumps from one region to another.

Recall that the trace theorem for Hölder resp. Besov spaces (see [34]) implies that for $F^{2J}|_{\Omega_i} \in B^{\alpha}_{\infty,\infty}(\Omega_i)$, the (scaled) restriction $\tilde{f}^{2J}(t)$ along the curve $\tilde{p}^{2J}(t)$ is again Hölder smooth of order α in each subinterval of $[0,1]$, determined by

$$T_i = \{t \in [0,1] : \tilde{p}^{2J}(t) \in \Omega_i\}$$

and with assuming that the corresponding path vector p^{2J} has no interruptions in Ω_i. In particular, for every $\alpha \in (0,1)$, we find the estimate

$$\begin{aligned}
|\tilde{f}^{2J}|_{C^\alpha(T_i)} &= \sup_{\substack{t_1\neq t_2\\ t_1,t_2\in T_i}} \frac{|F^{2J}(\tilde{p}^{2J}(t_1))-F^{2J}(\tilde{p}^{2J}(t_2))|}{|t_1-t_2|^\alpha} \\
&= \sup_{\substack{t_1\neq t_2\\ t_1,t_2\in T_i}} \frac{|F^{2J}(\tilde{p}^{2J}(t_1))-F^{2J}(\tilde{p}^{2J}(t_2))|}{\|\tilde{p}^{2J}(t_1)-\tilde{p}^{2J}(t_2)\|_2^\alpha} \frac{\|\tilde{p}^{2J}(t_1)-\tilde{p}^{2J}(t_2)\|_2^\alpha}{|t_1-t_2|^\alpha} \\
&\leq \|F^{2J}\|_{C^\alpha(\Omega_i)} \|\tilde{p}^{2J}\|^\alpha_{C^1(T_i)}.
\end{aligned}$$

For the case $\alpha \in (1,2)$, the chain rule yields

$$\begin{aligned}
|\tilde{f}^{2J}|_{C^\alpha(T_i)} &= \sup_{\substack{t_1\neq t_2\\ t_1,t_2\in T_i}} \frac{|\nabla F^{2J}(\tilde{p}^{2J}(t_1))(\tilde{p}^{2J})'(t_1)-\nabla F^{2J}(\tilde{p}^{2J}(t_2))(\tilde{p}^{2J})'(t_2)|}{|t_1-t_2|^{\alpha-1}} \\
&\leq \sup_{\substack{t_1\neq t_2\\ t_1,t_2\in T_i}} \Big(\frac{\|\nabla F^{2J}(\tilde{p}^{2J}(t_1))-\nabla F^{2J}(\tilde{p}^{2J}(t_2))\|_2}{\|\tilde{p}^{2J}(t_1)-\tilde{p}^{2J}(t_2)\|_2^{\alpha-1}} \frac{\|\tilde{p}^{2J}(t_1)-\tilde{p}^{2J}(t_2)\|_2^{\alpha-1}}{|t_1-t_2|^{\alpha-1}} \|(\tilde{p}^{2J})\|_{C^1(T_i)} \\
&\qquad + \|F^{2J}\|_{C^1(\Omega_i)} \frac{(\tilde{p}^{2J})'(t_1)-(\tilde{p}^{2J})'(t_2)}{|t_1-t_2|^{\alpha-1}} \Big) \\
&\leq \|F^{2J}\|_{C^\alpha(\Omega_i)} \|\tilde{p}^{2J}\|^\alpha_{C^1(T_i)} + \|F^{2J}\|_{C^1(\Omega_i)} \|\tilde{p}^{2J}\|_{C^\alpha(T_i)}.
\end{aligned}$$

In general, we have

Lemma 3.1 *Suppose that the parameter curve $\tilde{p}^{2J} : [0,1] \to \Omega$ satisfies in each subinterval T_i the smoothness condition*

$$\|\tilde{p}^{2J}\|_{C^\beta(T_i)} \leq C_2\, 2^{J\beta} \tag{3.2}$$

for all $\beta \leq \max\{\alpha,1\}$ and with a constant C_2 being independent of J, i.e., $\tilde{p}^j \in C^\alpha(T_i)$ for $\alpha > 1$ and $\tilde{p}^j \in C^1(T_i)$ for $\alpha \leq 1$, $i = 1,\ldots,K$, with strictly bounded derivatives. Then for the one-dimensional restriction $\tilde{f}^{2J}(t) = F^{2J}(\tilde{p}^{2J}(t))$ along the curve $\tilde{p}^{2J}$ we have

$$\|\tilde{f}^{2J}\|_{C^\alpha(T_i)} \leq \tilde{C} 2^{J\alpha} \|F^{2J}\|_{C^\alpha(\Omega_i)}, \tag{3.3}$$

where $\|F^{2J}\|_{C^\alpha(\Omega_i)}$ is bounded by a constant which is independent of J.

Proof. For $\alpha < 2$ the assertion already follows from the previous observations. Let $m := \lfloor \alpha \rfloor$. Further, we use the abbreviations $F = F^{2J}$, $\tilde{p} = \tilde{p}^{2J}$. In the general case the formula by Faà di Bruno implies that $D^m \tilde{f}^{2J}(t) := \frac{d^m}{dt^m} \tilde{f}^{2J}(t) = D^m(F(\tilde{p}(t)))$ is a finite linear combination of terms

$$S_{\nu,J}(t) := D^\nu F(x)|_{x=\tilde{p}(t)} \; (D^{m_1}\tilde{p}_1(t))^{b_1} \cdots (D^{m_k}\tilde{p}_k(t))^{b_k} \,,$$

where $\tilde{p} = \tilde{p}^{2J} = (\tilde{p}_1, \tilde{p}_2)^T$, $m_1, \cdots, m_k, b_1, \cdots, b_k \in \mathbb{N}$ with $m_1 b_1 + \cdots + m_k b_k = m$, and with $\nu \in \mathbb{N}_0^2$, $|\nu| \le m$. In the following estimate, limit our attention to the case $k = 2$, in order to alleviate notations and without loss of generality. For $t_1, t_2 \in T_i$ with $t_1 \neq t_2$ we now obtain with (3.2) the estimate

$$\begin{aligned}
&\tfrac{S_{\nu,J}(t_1) - S_{\nu,J}(t_2)}{|t_1 - t_2|^{\alpha - m}} \\
\le\; & \left(\tfrac{|(D^\nu F)(\tilde{p}(t_1)) - (D^\nu F)(\tilde{p}(t_2))|}{\|\tilde{p}(t_1) - \tilde{p}(t_2)\|_2^{\alpha-m}} \right) \left(\tfrac{\|\tilde{p}(t_1) - \tilde{p}(t_2)\|_2^{\alpha-m}}{|t_1 - t_2|^{\alpha-m}} \right) |(D^{m_1}\tilde{p}_1(t_1))^{b_1} (D^{m_2}\tilde{p}_2(t_1))^{b_2}| \\
& + |(D^\nu F)(\tilde{p}(t_2))| \left(\tfrac{|(D^{m_1}\tilde{p}_1(t_1))^{b_1}(D^{m_2}\tilde{p}_2(t_1))^{b_2} - (D^{m_1}\tilde{p}_1(t_2))^{b_1}(D^{m_2}\tilde{p}_2(t_2))^{b_2}|}{|t_1 - t_2|^{\alpha-m}} \right) \\
\le\; & |F|_{C^{\alpha+|\nu|-m}(\Omega_i)} \, \|\tilde{p}\|_{C^1(T_i)}^{\alpha-m} \, \|\tilde{p}\|_{C^{m_1}(T_i)}^{b_1} \, \|\tilde{p}\|_{C^{m_2}(T_i)}^{b_2} + \|F\|_{C^{|\nu|}(\Omega_i)} \Big(b_1 \|\tilde{p}\|_{C^{m_2}(T_i)}^{b_2} \\
& \times \|\tilde{p}\|_{C^{m_1}(T_i)}^{b_1 - 1} \|\tilde{p}\|_{C^{m_1+\alpha-m}(T_i)} + b_2 \|\tilde{p}\|_{C^{m_1}(T_i)}^{b_1} \|\tilde{p}\|_{C^{m_2}(T_i)}^{b_2 - 1} \|\tilde{p}\|_{C^{m_2+\alpha-m}(T_i)} \Big) \\
\le\; & |F|_{C^{\alpha+|\nu|-m}(\Omega_i)} \, 2^{J(\alpha-m)} \, 2^{J(m_1 b_1)} \, 2^{J(m_2 b_2)} \\
& + \|F\|_{C^{|\nu|}(\Omega_i)} (b_1 2^{J(m_2 b_2 + m_1 b_1 - m_1 + m_1 + \alpha - m)} + b_2 2^{J(m_1 b_1 + m_2 b_2 - m_2 + m_2 + \alpha - m)}) \\
\le\; & C_\nu 2^{J\alpha} \|F\|_{C^\alpha(\Omega_i)},
\end{aligned}$$

where we have used the inequality $|x^k - y^k| \le k|x - y| \max\{|x|, |y|\}^{k-1}$ for $x, y \in \mathbb{R}$, $k \in \mathbb{N}$. The boundedness of $\|F^{2J}\|_{C^\alpha(\Omega_i)}$ by a constant which is independent of J has already been shown in Subsection 2.2. Hence, the assertion follows. □

Let us assume in the sequel that the path function $\tilde{p}$ satisfies the smoothness condition (3.2) in Lemma 3.1. Thus, we obtain for the N-th order modulus of smoothness of $\tilde{f}^{2J}$ the estimate

$$\omega_N(\tilde{f}^{2J}, h)_\infty := \sup_{|\tilde{h}| \le h} \|\Delta_{\tilde{h}}^N \tilde{f}^{2J}\|_\infty \lesssim h^\alpha \|\tilde{f}^{2J}\|_{C^\alpha(T_i)} \lesssim (2^J h)^\alpha \|F^{2J}\|_{C^\alpha(\Omega_i)} \tag{3.4}$$

within the subintervals, where $\tilde{f}^{2J}$ is smooth, i.e., for $N = \lfloor \alpha + 1 \rfloor$ and

$$T_{i,h} := \left\{ t : \, \tilde{p}^{2J}(t + kh) \in \Omega_i, \, k = 0, \ldots, N \right\},$$

see [4]. Next, we consider the L^2-projection $f^{2J} := P_{2J} \tilde{f}^{2J}$ of $\tilde{f}^{2J}$ onto the scaling space

$$V^{2J} := \mathrm{clos}_{L^2[0,1)} \mathrm{span}\{\varphi_{2J,n} : \, n = 0, \ldots, 2^{2J} - 1\},$$

where φ is assumed to be a sufficiently smooth scaling function, see Section 2.2. Then, $f^{2J} = P_{2J} \tilde{f}^{2J} := \sum_{n=0}^{2^{2J}-1} \langle f^{2J}, \tilde{\varphi}_{2J,n} \rangle \, \varphi_{2J,n}$ also satisfies a Hölder smoothness condition of order α. Following along the lines of [4, Theorem 3.3.3], we obtain in the subintervals $T_{i,2^{-2J}}$ the estimate

$$\begin{aligned}
\|\tilde{f}^{2J} - f^{2J}\|_{L^\infty(T_{i,2^{-2J}})} &= \|\tilde{f}^{2J} - P_{2J}\tilde{f}^{2J}\|_{L^\infty(T_{i,2^{-2J}})} \lesssim \omega_N(\tilde{f}^{2J}, 2^{-2J})_\infty \\
&\lesssim (2^{-2J})^\alpha \|\tilde{f}^{2J}\|_{C^\alpha(T_i)} \lesssim (2^{-J})^\alpha \|F^{2J}\|_{C^\alpha(\Omega_i)}.
\end{aligned} \tag{3.5}$$

In particular,

$$|\tilde{f}^{2J}(2^{-2J}\ell) - f^{2J}(2^{-2J}\ell)| = |F^{2J}(\tilde{p}^{2J}(2^{-2J}\ell)) - f^{2J}(2^{-2J}\ell)| \lesssim 2^{-J\alpha}.$$

In the next step, we decompose the function $f^{2J} = \sum_{\ell} c_p^{2J}(\ell)\varphi_{2J,\ell}$ with $c_p^{2J}(\ell) := \langle f^{2J}, \tilde{\varphi}_{2J,\ell}\rangle$ into the low-pass part f^{2J-1} and the high-pass part g^{2J-1}. Applying one level of the one-dimensional wavelet transform to the data set $(c_p^{2J}(\ell))_{\ell=0}^{2^{2J}-1}$, we obtain the decomposition $f^{2J} = f^{2J-1} + g^{2J-1}$ with

$$f^{2J-1} = \sum_{n=0}^{2^{2J-1}-1} c_p^{2J-1}(n)\,\varphi_{2J-1,n} \quad \text{and} \quad g^{2J-1} = \sum_{n=0}^{2^{2J-1}-1} d_p^{2J-1}(n)\,\psi_{2J-1,n},$$

where $c_p^{2J-1}(n) := \langle f^{2J}, \tilde{\varphi}_{2J-1,n}\rangle$ and $d_p^{2J-1}(n) := \langle f^{2J}, \tilde{\psi}_{2J-1,n}\rangle$. From the Hölder smoothness of f^{2J}, we find for $t \in T_i$ the representation

$$f^{2J}(t) = q_\alpha(t - t_0) + R(t - t_0)$$

for $t_0 \in \{2^{-2J}k : k = 0, \ldots, 2^{2J} - 1\} \cap T_i$ and $|t - t_0| \le 2^{-2J}$, where q_α denotes the Taylor polynomial of degree $\lfloor\alpha\rfloor$ of f^{2J} at t_0, and where the remainder R satisfies the bound $|R(t - t_0)| \le c_\varphi(t)2^{-J\alpha}$. Hence, if $\mathrm{supp}(\tilde{\psi}_{2J-1,n}) \in T_i$ for some i, the wavelet coefficients satisfy

$$\begin{aligned} |d_p^{2J-1}(n)| &= |\langle q_\alpha(\cdot - t_0) + R(\cdot - t_0),\, \tilde{\psi}_{2J-1,n}\rangle| = |\langle R(\cdot - t_0),\, \tilde{\psi}_{2J-1,n}\rangle| \\ &\le c_{\varphi,n}\, 2^{-J\alpha}\, \|\tilde{\psi}_{2J-1,n}\|_1 = \tilde{c}_{\varphi,n}\, 2^{(-J+1/2)(\alpha+1)}, \end{aligned}$$

where we have used $\|\tilde{\psi}_{2J-1,n}\|_1 = 2^{-J+1/2}\|\tilde{\psi}\|_1$. Observe that the constant $c_{\varphi,n}$ in this inequality depends on the choice of the wavelet basis but also on the (local) smoothness properties of $\tilde{f}^{2J}$, and hence on the (local) boundedness properties of the derivatives of the chosen path function $\tilde{p}^{2J}$ in Lemma 3.1.

Now let Λ^{2J-1} be the set of all $n \in \{0, \ldots, 2^{2J-1} - 1\}$, where the above estimate for $d^{2J-1}(n)$ is satisfied. Then, the number of the remaining $2^{2J-1} - \#\Lambda^{2J-1}$ wavelet coefficients corresponds to the number of positions, where $\tilde{f}^{2J}$ is discontinuous (e.g. caused by crossing over of the path function to another region) or where derivatives of $\tilde{f}^{2J}$ are not suitably bounded, i.e., the constant $c_{\varphi,n}$ is too large (caused e.g. by tight turns of the path function $\tilde{p}^{2J}$). We assume that this number is bounded by CK, where K is the number of regions in the original image F, and where the constant C does not depend on 2^J.

Now, we consider the low-pass function f^{2J-1} and reconstruct a bivariate function F^{2J-1} as follows. Taking only the path function values $\tilde{p}^{2J}(2^{-2J+1}n)$, we put

$$\tilde{\Gamma}_i^{2J-1} := \{\tilde{p}^{2J}(2^{-2J+1}n) : n = 0, \ldots, 2^{2J-1} - 1,\, \tilde{p}^{2J}(2^{-2J+1}n) \in \Omega_i\}$$

for each $i = 1, \ldots, K$ and $\tilde{\Gamma}^{2J-1} := \cup_{i=1}^K \tilde{\Gamma}_i^{2J-1}$. We compute the polyharmonic spline interpolant

$$F^{2J-1}(x) := \sum_{i=1}^{K} \left(\sum_{y \in \tilde{\Gamma}_i^{2J-1}} c_y^i\, \phi_\alpha\left(\|x - y\|_2\right) + p_m^i(x) \right) \chi_{\Omega_i}(x),$$

satisfying the interpolation conditions

$$F^{2J-1}\left(\tilde{p}^{2J}(2^{-2J+1}n)\right) = f^{2J-1}(2^{-2J+1}n) \qquad \text{for all } n = 0, \dots, 2^{2J-1} - 1.$$

Therefore,

$$\begin{aligned} \left|F^{2J}(\tilde{p}^{2J}(2^{-2J+1}n)) - F^{2J-1}(\tilde{p}^{2J}(2^{-2J+1}n))\right| &= |\tilde{f}^{2J}(2^{-2J+1}n) - f^{2J-1}(2^{-2J+1}n)| \\ &= |\tilde{f}^{2J}(2^{-2J+1}n) - P_{2J-1}\tilde{f}^{2J}(2^{-2J+1}n)| \\ &\lesssim 2^{(-J+1)\alpha} = D^{\alpha}(2^{-J+1/2})^{\alpha}, \end{aligned} \tag{3.6}$$

where $D = \sqrt{2}$, and where the last inequality again follows analogously as in (3.5) since f^{2J-1} is the orthogonal projection of $\tilde{f}^{2J}$ to

$$V^{2J-1} := \mathrm{clos}_{L^2[0,1)}\mathrm{span}\{\varphi_{2J-1,n} : n = 0, \dots 2^{2J-1} - 1\}.$$

The last inequality implies that F^{2J-1} is still a good approximation to F, since the interpolation points have changed only slightly. However, only half of the interpolation points are left, which are irregularly distributed in Ω. Moreover, by (3.1) we have

$$\max_{x\in\Omega_i} \min_{y\in\tilde{\Gamma}_i^{2J-1}} |x - y| \le 2^{-J+1} + \mu 2^{-J} = D\, 2^{-J+1/2}. \tag{3.7}$$

In order to apply the described procedure to the further levels of the EPWT in an analogous manner we need that again $\|F^{2J-1}\|_{W_2^{\alpha+1}(\Omega_i)}$ is bounded independently of J and that

$$\|F^{2J} - F^{2J-1}\|_{L^2(\Omega_i)} \lesssim 2^{(-J+1/2)(\alpha+1)} \qquad \text{for all } i = 1, \dots, K, \tag{3.8}$$

holds. We assume that these conditions are indeed satisfied.

Remark 1. Unfortunately, we are not aware of a rigorous proof of this assertion in the literature for polyharmonic spline approximation (other than interpolation). One needs to show that for a suitably bounded approximation error (as in (3.6)) at the considered quasi-regular point grid satisfying (3.7) the obtained approximant F^{2J-1} has again a bounded norm in the corresponding native space (here being equivalent to $W_2^{\alpha+1}(\Omega_i)$) that is independent of the number of grid points. However, our assumption is strongly supported by the following observations.

Let $X \subset \Omega_i$ be an arbitrary $\mathcal{P}_m$-regular point set of size $c_0 2^{2J-1}$ with fill distance $h_X \sim 2^{-J+1}$ being completely different from the point set Γ_i^{2J} determining the polyharmonic spline F^{2J} (i.e., $X \cap \Gamma_i^{2J} = \emptyset$ and the distance between points of X and Γ_i^{2J} is suitably bounded from below a.e.). Moreover, let $S_{F^{2J},X}$ be the polyharmonic spline interpolant of F^{2J} with $F^{2J}(x) = S_{F^{2J},X}(x)$ for $x \in X$. Then, by Theorem 11.36 in [37], we have

$$\|F^{2J} - S_{F^{2J},X}\|_{\infty} \le C\, 2^{(-J+1/2)\alpha} \|F^{2J}\|_{W_2^{\alpha+1}(\Omega_i)}, \tag{3.9}$$

and further, $\|S_{F^{2J},X}\|_{W_2^{\alpha+1}(\Omega_i)} \lesssim \|F^{2J}\|_{W_2^{\alpha+1}(\Omega_i)} \lesssim \|F\|_{W_2^{\alpha+1}(\Omega_i)}$, see [25], Corollary 4.3. Hence, $\|S_{F^{2J},X}\|_{W_2^{\alpha+1}(\Omega_i)}$ is bounded independently of J. On the other hand, there exists a set $Y \subset \Omega_i$ of points $y \in \Omega_i$, where the distance between F^{2J} and its interpolant satisfies

$$|F^{2J}(y) - S_{F^{2J},X}(y)| > c\, 2^{(-J+1/2)\alpha}, \tag{3.10}$$

with some suitable $c > 0$ being independent of J, since otherwise one could easily derive by an inverse theorem that F^{2J} has even a higher Hölder exponent than α, see e.g. Theorem 4.2 in [36].

If we choose the constant c in a way such that Y contains a quasi-uniformly distributed point set Y^{2J-1} with $\#Y^{2J-1} \sim 2^{2J-1}$ and with fill distance $\sim 2^{-J+1}$, then our argument works as follows. Consider now a polyharmonic spline interpolant $S_{Y^{2J-1}}$ of $S_{F^{2J},X}$ with $S_{Y^{2J-1}}(y) = S_{F^{2J},X}(y)$ for $y \in Y^{2J-1}$. Then we conclude as in Subsection 2.2 that

$$\|S_{Y^{2J-1}}\|_{W_2^{\alpha+1}(\Omega_i)} \lesssim \|S_{F^{2J},X}\|_{W_2^{\alpha+1}(\Omega_i)} \lesssim \|F^{2J}\|_{W_2^{\alpha+1}(\Omega_i)}$$

and

$$\|S_{Y^{2J-1}} - F^{2J}\|_{L^2(\Omega_i)} \le \|S_{F^{2J},Y} - S_{F^{2J},X}\|_{L^2(\Omega_i)} + \|S_{F^{2J},X} - F^{2J}\|_{L^2(\Omega_i)} \lesssim 2^{(-J+1/2)(\alpha+1)},$$

while for $y \in Y^{2J-1}$ we have

$$|F^{2J}(y) - S_{Y^{2J-1}}(y)| > c\, 2^{(-J+1/2)\alpha}$$

by (3.10). If we can choose the point set X such that $Y^{2J-1} = \tilde{\Gamma}^{2J-1}$, then the needed assumptions on F^{2J-1} follow.

We finally show that one is able to choose the constant c in (3.10) independently of J such that Y contains a point set Y^{2J-1} with the desired properties, i.e., with $\#Y^{2J-1} \sim 2^{2J-1}$ and with fill distance $\sim 2^{-J+1}$. This can be seen by a geometric argument. Using the inverse error estimate in Theorem 5.1 in [25] (with $f = F^{2J} - S_{F^{2J},X}$, $\mu = \alpha + 1$, and $q_X = 2^{(-J+1/2)(\alpha+1)}$), we see that

$$\|F^{2J} - S_{F^{2J},X}\|_{L^2(\Omega_i)} \sim 2^{(-J+1/2)(\alpha+1)}.$$

Considering now a triangulation $\mathcal{T}$ of Ω_i with the vertex set X, let us consider in each triangle $T \in \mathcal{T}$ the point x_T, where $c_T := \max_{x \in T} |F^{2J}(x) - S_{F^{2J},X}(x)||F^{2J}(x_T) - S_{F^{2J},X}(x_T)|$. With $\min_{x_1,x_2 \in X} |x_1 - x_2| \sim 2^{-J+1/2}$, $h_X \sim 2^{-J+1}$, and recalling that $F^{2J}(x) - S_{F^{2J},X}(x) = 0$ for $x \in X$, it follows by linear approximation that $\int_T |F^{2J}(x) - S_{F^{2J},X}(x)|^2 dx \sim (c_T\, 2^{(-2J+1)})^2$, and hence that

$$\|F^{2J} - S_{F^{2J},X}\|_{L^2(\Omega_i)}^2 \sim 2^{-4J+2} \sum_{T \in \mathcal{T}} c_T^2,$$

i.e., $\sum_{T \in \mathcal{T}} c_T^2 \sim 2^{(-2J+1)\alpha}\, 2^{2J-1}$. Since $\#\mathcal{T} = 2^{2J-1}$, it follows from (3.9) that c_T needs to be of size $\sim 2^{(-J+1/2)\alpha}$ for almost all triangles T.

Remark 2. Compared with the original EPWT algorithm in [26], the condition (3.1) is an important relaxation. Here the scaling and wavelet coefficients are computed from the function $f^{2J}(t)$ that approximates $F^{2J}(\tilde{p}^{2J}(t))$, but since $\tilde{p}^{2J}(2^{-2J}\ell)$, $\ell = 0, \dots, 2^{2J} - 1$, do not interpolate but only approximate the original grid points I_{2J}, we work with new function values $F^{2J}(\tilde{p}^{2J}(2^{-2J}\ell))$ of the polyharmonic spline function F^{2J} instead of the given values $F(y)$, $y \in I_{2J}$. The original EPWT relates to the special case $\mu = 0$, where we have interpolation in (3.1).

3.2 Conditions for the Path Vectors

Before we proceed with the error estimates for the further levels of the EPWT algorithm, we want to fix specific side conditions for the path functions that are required for our error analysis, and that have been implicitly used already in the estimates at the highest level in Section 3.1. The side conditions are termed (a) *region condition*, (b) *path smoothness condition*, and (c) *diameter condition*, as stated below.

The region condition and the diameter condition have been similarly stated already in [29] to prove the N-term approximation estimates. The region condition ensures that at each level of the EPWT, the path function "jumps" only C_1K times from one region Ω_i to another region or inside a region. The diameter condition ensures that the remaining points in $\tilde{\Gamma}^j$ are quasi-uniformly distributed, such that there is a constant D, not depending on J or j, satisfying

$$\max_{x\in\Omega_i}\min_{y\in\tilde{\Gamma}^j}\|x-y\|_2 \le (1+\sqrt{2})D2^{-j/2}. \tag{3.11}$$

Further, at each level j of the EPWT, we assume that we can find a smooth function $\tilde{p}^j : [0,1] \to \Omega$ with uniformly bounded derivatives that approximates the values in $\tilde{\Gamma}^j$. This leads us to the vector p^j, whose components from $\tilde{\Gamma}^j$ are suitably ordered, and we have $|\tilde{p}^j(2^{-j}\ell) - p^j(\ell)| \le \mu 2^{-j}$ for $\ell = 0,\dots,2^j-1$. Further, we assume that $\tilde{p}^j(t) \in \Omega_i$, for $t \in [\ell/2^j, (\ell+1)/2^j]$, if $\tilde{p}^j(2^{-2J}\ell)$ and $\tilde{p}^j(2^{-2J}(\ell+1))$ are in Ω_i.

Let us introduce the three conditions more explicitly.

(a) Region condition. At each level j of the EPWT, the path function $\tilde{p}^j$ is chosen, such that it contains only at most C_1K discontinuities caused by crossing over from one region Ω_i to another region or by jumping within one region Ω_i.

(b) Path smoothness condition. The path function $\tilde{p}^j$ satisfies in each subinterval T_i the smoothness condition

$$\|\tilde{p}^j\|_{C^\beta(T_i)} \le C_2\, 2^{j\beta/2}$$

for all $\beta \le \max\{\alpha, 1\}$ and with a constant C_2 being independent of j, i.e., $\tilde{p}^j \in C_\alpha(T_i)$ for $\alpha > 1$ and $\tilde{p}^j \in C_1(T_i)$ for $\alpha \le 1$, $i = 1,\dots,K$, with bounded derivatives.

(c) Diameter condition. At each level of the EPWT, we require for almost all values $2^{-j}\ell$, $\ell = 0,\dots,2^j-2$, the condition

$$\|\tilde{p}^j(2^{-j}\ell) - \tilde{p}^j(2^{-j}(\ell+1))\|_2 \le D\, 2^{-j/2}, \tag{3.12}$$

where D is independent of J and j, and where the number of values of $\tilde{p}^j(2^{-j}\ell)$ which do not satisfy the diameter condition, is bounded by a constant C_3 not depending on j. Hence, at each level j, consecutive components $\tilde{p}^j(2^{-j}\ell)$ in the path function should be spatial neighbors.

The conditions (a) and (c) can be enforced by the proposed path construction in Subsection 2.2, see (2.12) for the region condition and (2.13) for the diameter condition. Particularly, the diameter condition ensures a quasi-uniform distribution of the points in $\tilde{\Gamma}^j$. This can be seen inductively as follows. For $j = 2J$, the assertion (3.11) is obvious,

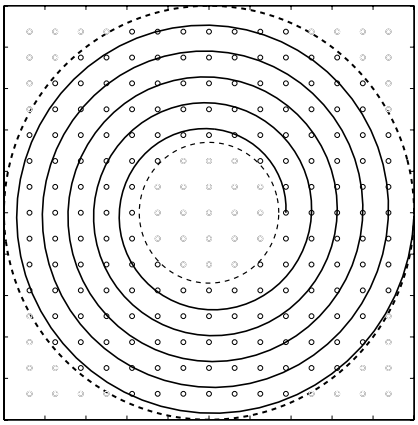

Figure 2: **Archimedean spiral.** A not simply connected domain Ω. The boundary of Ω is displayed with dashed lines. The smooth path $\tilde{p}^{2J}$ (here for $J = 3$ and displayed with a solid line) is a planar spiral curve, an Archimedean spiral. The set Γ^{2J} is illustrated by uniformly distributed points.

see also (3.7). Generally, assuming (3.11) and (3.12) at level $j+1$, it follows that

$$\begin{aligned}\max_{x\in\Omega_i}\min_{y\in\tilde{\Gamma}^j}\|x-y\|_2 &\le \max_{x\in\Omega_i}\min_{y\in\tilde{\Gamma}^{j+1}}\|x-y\|_2+\max_{y\in\tilde{\Gamma}^{j+1}}\min_{z\in\tilde{\Gamma}^j}\|y-z\|_2\\ &\le (1+\sqrt{2})D2^{-(j+1)/2}+D2^{-(j+1)/2}=(1+\sqrt{2})D\,2^{-j/2}.\end{aligned}$$

Unfortunately, the given path smoothness condition is rather restrictive. We consider discussing an example, where the conditions for the path $\tilde{p}^j$ are satisfied for $\alpha \in (1,2)$.

Example. We consider the domain Ω being the difference set between two concentric disks with radii $r_1 < r_2$, a circular ring. In this case, we let the smooth path $\tilde{p}^{2J}$ (at level J and for J large enough) be a planar spiral curve. More precisely, we take an Archimedean spiral with parametric form $(a\phi\cos\phi,\, a\phi\sin\phi)$ with $a = 2^{-J}/2\pi$ for $\phi \in [\phi_1,\phi_2] = [2\pi\lceil 2^J r_1\rceil, 2\pi\lfloor 2^J r_2\rfloor]$, see Figure 2. Then each straight line through the origin intersects the successive turnings of the spiral in points with a constant separation distance $2\pi a = 2^{-J}$. We can assume that we have to approximate $\mathcal{O}(2^{2J}(r_2^2-r_1^2)\pi)$ quasi-uniformly distributed points of Γ^{2J} in the circular ring while the length of the spiral is

$$\frac{a}{2}\left[\phi\sqrt{\phi^2+1}+\ln(\phi\sqrt{\phi^2+1})\right]_{\phi_1}^{\phi_2}=\mathcal{O}(2^J(r_2^2-r_1^2)).$$

Hence, we find a suitable parameterization for $\tilde{p}^{2J}$, such that each point in Γ^{2J} is approximated sufficiently well by $\tilde{p}^{2J}(2^{-J}\ell)$, i.e., the path $\tilde{p}^{2J}$ meets the conditions (b) and (c) for $\alpha \in (1,2)$. Figure 2 shows an example for a set Γ^{2J} of uniformly distributed points which can be approximated sufficiently well by points on the Archimedean spiral, e.g., by their Euclidean projections onto the curve $\tilde{p}^{2J}$.

Remark 1. To understand the strength of the path smoothness condition (b), we follow the arguments of one of Jean-Marie Mirebeau and show that the path vector p^j can only be taken straight along a line if $\mu < 1/4$. In particular, if the path function $\tilde{p}^{2J}$ is forced to interpolate the grid points of the form $(2^{-J}n_1, 2^{-J}n_2)$, $n_1, n_2 \in \mathbb{Z}$ (regular grid), as in the situation of the EPWT in [26], the path smoothness condition cannot be satisfied. Indeed, with (3.1) we find the estimate

$$\begin{aligned}&\|p^{2J}(\ell+1) - 2p^{2J}(\ell) + p^{2J}(\ell-1)\|_\infty \\ &\leq \|\tilde{p}^{2J}(2^{-2J}(\ell+1)) - 2\tilde{p}^{2J}(2^{-2J}\ell) + \tilde{p}^{2J}(2^{-2J}(\ell-1))\|_\infty + 2^{-J}\cdot 4\mu.\end{aligned}$$

But the smoothness condition (3.2) implies that

$$\begin{aligned}&\|\tilde{p}^{2J}(2^{-2J}(\ell+1)) - 2\tilde{p}^{2J}(2^{-2J}\ell) + \tilde{p}^{2J}(2^{-2J}(\ell-1))\|_\infty \\ = \;&\|(\tilde{p}^{2J}(2^{-2J}(\ell+1)) - \tilde{p}^{2J}(2^{-2J}\ell)) - (\tilde{p}^{2J}(2^{-2J}\ell) + \tilde{p}^{2J}(2^{-2J}(\ell-1)))\|_\infty \\ = \;&\|\int_{I_0} (\tilde{p}^{2J})'(t) - (\tilde{p}^{2J})'(t-2^{-2J})\,dt\|_\infty \\ \leq \;&2^{-2J}\,2^{-2J(\alpha-1)}|\tilde{p}^{2J}|_{C^\alpha(I)} = 2^{-2J\alpha}|\tilde{p}^{2J}|_{C^\alpha(I)} \leq C_2 2^{-J\alpha},\end{aligned}$$

where $I_0 := [2^{-2J}\ell, 2^{-2J}(\ell+1)] \subset I$, and where I denotes an interval where $\tilde{p}^{2J}$ is C^α smooth. Hence,

$$\|p^{2J}(\ell+1) - 2p^{2J}(\ell) + p^{2J}(\ell-1)\|_\infty \leq 2^{-J}(C_2 2^{J(1-\alpha)} + 4\mu).$$

For $\alpha > 1$ the term $C2^{J(1-\alpha)}$ tends to zero, as $J \to \infty$. On the other hand, if the points $p^{2J}(\ell+1), p^{2J}(\ell), p^{2J}(\ell-1)$ (being samples from the regular grid $(2^{-J}n_1, 2^{-J}n_2)$, $n_1, n_2 \in \mathbb{Z}$) do not lie on the same line, then the norm on the left hand side is at least 2^{-J}. Hence, for $\mu < 1/4$ the path conditions are incompatible.

Remark 2. The conditions (a) and (c) can slightly be relaxed the sense that the constants C_1 and C_3 may be allowed to depend polynomially on j (but not exponentially). In fact, considering the proof of Theorem 4.1 (resp. Theorem 3.3 in [29]), one needs to ensure that a weighted sum of all wavelet coefficients that are only satisfying the estimate (3.15) is still bounded.

3.3 The Further Levels of the EPWT

Let us now explain the further levels of the EPWT. These are performed following along the lines of the $2J$-th level. We start with the polyharmonic spline interpolant

$$F^j(x) := \sum_{i=1}^{K} \left(\sum_{y \in \tilde{\Gamma}_i^j} c_y^i\, \phi_\alpha\left(\|x-y\|_2\right) + p_m^i(x) \right) \chi_{\Omega_i}(x)$$

that satisfies the interpolation conditions $F^j(\tilde{p}^{j+1}(2^{-j}n)) = f^j(2^{-j}n)$ for all $n = 0, \ldots, 2^j - 1$. We fix a suitable path function $\tilde{p}^j$ approximating the set $\tilde{\Gamma}^j = \{\tilde{p}^{j+1}(2^{-j}n) : n = 0, \ldots, 2^j - 1\}$ with corresponding data values $\{F^j\left(\tilde{p}^{j+1}(2^{-j}n)\right) : n = 0, \ldots, 2^j - 1\}$, such that the three path conditions in Section 3.2 are satisfied. Then we determine the one-dimensional restriction of F^j along $\tilde{p}^j$, $\tilde{f}^j(t) := F^j\left(\tilde{p}^j(t)\right)$, $t \in [0,1]$. The piecewise Hölder smoothness of $\tilde{f}^j(t)$ ensures that

$$\omega_N(\tilde{f}^j, h)_\infty \lesssim (2^j h)^\alpha \qquad \text{for} \quad T_{i,h} := \{t : \tilde{p}^j(t+kh) \in \Omega_i,\, k = 0, \ldots, N\}.$$

Considering the L^2-projection $P_j\tilde{f}^j := \sum_{n=0}^{2^j-1} c_p^j(n)\,\varphi_{j,n}$ with $c_p^j(n) := \langle \tilde{f}^j,\, \tilde{\varphi}_{j,n}\rangle$ of $\tilde{f}^j$ onto the scaling space

$$V^j := \mathrm{clos}_{L^2[0,1)}\mathrm{span}\{\varphi_{j,n} :\, n = 0,\dots,2^j-1\},$$

where φ is assumed to be a smooth scaling function as in Subsection 3.1, we have

$$\|\tilde{f}^j - P_j\tilde{f}^j\|_{L^\infty(T_{i,2^{-j}})} \lesssim w_N(\tilde{f}^j, 2^{-J-j/2}) \lesssim 2^{-j\alpha/2},$$

where we have applied the diameter condition (3.12). We use the decomposition

$$P_j\tilde{f}^j = \sum_\ell c_p^j(\ell)\,\varphi_{j,\ell} = f^{j-1} + g^{j-1} = \sum_n c_p^{j-1}(n)\,\varphi_{j-1,n} + \sum_n d_p^{j-1}(n)\,\psi_{j-1,n}.$$

Then the Hölder smoothness of $P_j\tilde{f}^j$ in the intervals T_i yields the Taylor expansion

$$P_j\tilde{f}^j(t) = q_\alpha(t-t_0) + R(t-t_0) \quad \text{ with } |R(t-t_0)| \le c_\varphi(t)\,D^\alpha\,2^{-j\alpha/2}$$

for $t, t_0 \in T_i$, where D is the constant in the diameter condition (3.12), and this gives the estimate for the wavelet coefficients corresponding to the region Ω_i,

$$\begin{aligned} |d_p^{j-1}(n)| &= |\langle R(t-t_0), \tilde{\psi}_{j-1,n}\rangle| \le c_{\varphi,n}\,D^\alpha\,2^{-j\alpha/2}\,2^{-(j-1)/2}\,\|\tilde{\psi}\|_1 \\ &\le c_{\varphi,n}\,D^\alpha\,2^{-(j-1)(\alpha+1)/2}, \end{aligned} \tag{3.13}$$

where $c_{\varphi,n}$ depends on the local smoothness of $P_j\tilde{f}^j$ and hence on the derivative bounds of $\tilde{p}_j$ in (3.2). Again, let Λ^{j-1} be the set of indices n from $\{0,\dots,2^{j-1}-1\}$, where $d_p^{j-1}(n)$ satisfies the above estimate. Then the number of wavelet coefficients $2^{j-1} - \#\Lambda^{j-1}$ which are not satisfying this estimate (since $\mathrm{supp}\ \tilde{\psi}_{j-1,n} \not\subset T_i$ for some i) is bounded by a constant independent of J and j.

Finally, we obtain the polyharmonic spline interpolant

$$F^{j-1}(x) := \sum_{i=1}^K \left(\sum_{y\in\tilde{\Gamma}_i^{j-1}} c_y^i\,\phi_\alpha\left(\|x-y\|_2\right) + p_m^i(x) \right) \chi_{\Omega_i}(x),$$

where $\tilde{\Gamma}_i^{j-1} := \{\tilde{p}^j(2^{-j+1}n) :\, n = 0,\dots,2^{j-1}-1,\, \tilde{p}^j(2^{-j+1}n) \in \Omega_i\}$, $\tilde{\Gamma}^{j-1} := \cup_{i=1}^K \tilde{\Gamma}_i^{j-1}$, through the interpolation conditions $F^{j-1}(\tilde{p}^j(2^{-j+1}n)) = f^{j-1}(2^{-j+1}n)$ for all $n = 0,\dots,2^{j-1}-1$. Hence, we obtain the estimate

$$\left|F^j\left(\tilde{p}^j(2^{-j+1}n)\right) - F^{j-1}\left(\tilde{p}^j(2^{-j+1}n)\right)\right| = |\tilde{f}^j(2^{-j+1}n) - P_{j-1}\tilde{f}^j(2^{-j+1}n)| \lesssim 2^{-j\alpha/2}.$$

Let us summarize the above findings on the decay of the EPWT wavelet coefficients in the following theorem.

Theorem 3.2 *For $j = 2J-1,\dots,0$, let $d_p^j(\ell) = \langle f^{j+1},\, \tilde{\psi}_{j,\ell}\rangle$, $\ell = 0,\dots,2^j-1$, denote the wavelet coefficients that are obtained by applying the EPWT algorithm to F (according to the above definitions in Section 3), where we assume that $F \in L^2(\Omega)$ is piecewise Hölder smooth of order α as prescribed in Subsection 2.1. Further assume that the path functions $\tilde{p}^{j+1}$, $j = 2J-1,\dots,0$, in the EPWT algorithm satisfy the region condition (a), the path*

smoothness condition (b), and the diameter condition (c) of Subsection 3.2. Then, for all $j = 2J-1, \dots, 0$ *and* $\ell \in \Lambda^j$*, the estimate*

$$|d_p^j(\ell)| \le C\, D^\alpha\, 2^{-j(\alpha+1)/2} \tag{3.14}$$

holds, where $D > 1$ *is the constant of the diameter condition (3.12),* α *is the Hölder exponent of* F*, and* C *depends on the utilized wavelet basis and on the Hölder constant in (2.1). Furthermore, for all* $\ell \in \{0, \dots, 2^j - 1\} \setminus \Lambda^j$*, we obtain the estimate*

$$|d_p^j(\ell)| \le C'\, 2^{-j/2} \tag{3.15}$$

with some constant C' *being independent of* J *and* j*.*

Proof. The proof of (3.14) follows directly from (3.13). Likewise, for all $\ell \in \{0, \dots, 2^j - 1\} \setminus \Lambda^j$, i.e., for point sets that do not satisfy the diameter, the path smoothness, or the region condition, we observe at least

$$|d_p^j(\ell)| \le C'\, 2^{-j/2} = C'\, 2^{-j/2}$$

since we can assume that F^j is bounded, and hence the above estimate (3.14) holds for $\alpha = 0$. Thus (3.15) follows. □

4 Optimal N-term Approximation obtained from the EPWT

Consider now the vector of all EPWT wavelet coefficients

$$\mathbf{d}_p = ((d_p^{2J-1})^T, \dots, d_p^0, d_p^{-1})^T$$

with $d_p^j = (d_p^j(\ell))_{\ell=0}^{2^j-1}$ for $j = 0, \dots, 2J-1$, and with the mean value

$$d_p^{-1} = d_p^{-1}(0) := f^0(0) = 2^{-2J} \sum_{y \in I_J} F^{2J}(y),$$

together with the side information on the path functions $\tilde{p}^{2J}, \dots, \tilde{p}^1$ at each iteration step. With this information the reconstructed image F_{rec}^{2J} is uniquely recovered, where F_{rec}^{2J} is the polyharmonic spline interpolation satisfying

$$F_{rec}^{2J}\left(\tilde{p}^{2J}(2^{-2J}n)\right) = f^{2J}(2^{-2J}n), \quad n = 0, \dots, 2^{2J} - 1.$$

Indeed, reconsidering the $(j+1)$-th level of the EPWT procedure, we observe that the scaling coefficients $c_p^j(n) = \langle \tilde{f}^{j+1}, \tilde{\varphi}_{j,n} \rangle$ and the wavelet coefficients $d_p^j = \langle \tilde{f}^{j+1}, \tilde{\psi}_{j,n} \rangle$ determine f^j and g^j, and hence $P_{j+1}\tilde{f}^{j+1}$ uniquely. Further, the polyharmonic spline interpolation F^{j+1} is entirely determined by $\tilde{f}^{j+1}(2^{-(j+1)}n)$ and the side information about the path function $\tilde{p}^{j+1}(2^{-(j+1)}n)$.

By the choice of the wavelet basis, it further follows for $n = 0, \dots, 2^{j+1} - 1$ that

$$\begin{aligned} |F^{j+1}(\tilde{p}^{j+1}(2^{-(j+1)}n)) - F_{rec}^{j+1}(\tilde{p}^{j+1}(2^{-(j+1)}n))| &= |\tilde{f}^{j+1}(2^{-(j+1)}n) - P_{j+1}\tilde{f}^{j+1}(2^{-(j+1)}n)| \\ &\lesssim 2^{-(j+1)\alpha/2}, \end{aligned}$$

where F_{rec}^{j+1} is uniquely determined by $P_{j+1}\tilde{f}^{j+1}(2^{-(j+1)}n)$, $n = 0, \ldots, 2^{j+1}-1$ and the side information about the path $\tilde{p}^{j+1}$.

In order to find a sparse approximation of the digital image F resp. F^{2J}, we apply a shrinkage procedure to the EPWT wavelet coefficients $d_p^j(\ell)$, using the hard threshold function

$$s_\sigma(x) = \begin{cases} x & |x| \geq \sigma, \\ 0 & |x| < \sigma, \end{cases} \qquad \text{for some } \sigma > 0.$$

We now study the error of a sparse representation using only the N wavelet coefficients with largest absolute value for an approximative reconstruction of F^{2J}. For convenience, let S_N^{2J} be the set of indices (j,ℓ) of the N wavelet coefficients with largest absolute value.

Moreover, let $F_{N,rec}^{2J}$ denote the polyharmonic spline interpolation determined by the reconstructed function $f_N^{2J} = \sum_n c_{p,N}^{2J}(n)\varphi_{2J,n}$ using only the N wavelet coefficients with largest absolute value, satisfying the interpolation conditions

$$F_{N,rec}^{2J}\left(\tilde{p}^{2J}(2^{-2J}n)\right) = f_N^{2J}(2^{-2J}n) \qquad \text{for} \quad n = 0, \ldots, 2^{2J}-1.$$

While the wavelet basis used above is not orthonormal but stable, we can still estimate the distance of F^{2J} and $F_{N,rec}^{2J}$ by

$$\epsilon_N = \|F^{2J} - F_{N,rec}^{2J}\|_{L^2(\Omega)}^2 \lesssim \sum_{(j,\ell)\notin S_N^{2J}} |d_p^j(\ell)|^2.$$

This estimate is a direct consequence of Theorem 3.2. Indeed, at each level of the EPWT, we observe that

$$\begin{aligned} \|F^{j+1} - F_{rec}^j\|_{L^2(\Omega)} &\leq \|F^{j+1} - F^j\|_{L^2(\Omega)} + \|F^j - F_{rec}^j\|_{L^2(\Omega)} \\ &\lesssim \left(\sum_{n=0}^{2^j-1} |d_p^j(n)|^2\right)^{1/2} \lesssim 2^{-j\alpha/2}, \end{aligned}$$

where the number of wavelet coefficients satisfying (3.14) is $2^j - C_1K + C_3$, and where the constants C_1 and C_3 do not depend on j, see Section 3.2.

Now we obtain the main result of this paper, showing the optimal N-term approximation of the EPWT algorithm.

Theorem 4.1 *Let F_N^{2J} be the N-term approximation of F^{2J} as constructed above, and let the assumptions of Theorem 3.2 be satisfied. Then the estimate*

$$\epsilon_N = \|F^{2J} - F_N^{2J}\|_2^2 \leq \tilde{C}\, N^{-\alpha} \tag{4.1}$$

holds for all $J \in \mathbb{N}$, where the constant $\tilde{C} < \infty$ does not depend on J.

Proof. The proof can be carried out by following along the lines of the proof of Theorem 3.3 in [29]. □

Let us finally conclude by stating the following corollary.

Corollary 4.2 *Let $F \in L^2(\Omega)$ be piecewise Hölder continuous (as assumed in Subsection 2.1). Then, for each $\epsilon > 0$ there exists an integer $J(\epsilon)$, such that for all $J \geq J(\epsilon)$ the N-term estimate*

$$\|F - F_N^{2J}\|_{L^2}^2 < \tilde{C} N^{-\alpha} + \epsilon$$

holds, where $\tilde{C}$ is the constant in (4.1).

Proof. The proof follows directly from Theorem 4.1 and (2.10). □

Remark. Observe that the above "construction" using polyharmonic spline interpolations F^j and smooth path functions $\tilde{f}^j$ is only done for analytical purposes. For a pragmatic implementation of the EPWT we refer to our previous works [26, 28, 33], where different versions of the EPWT are tested, relying on various strategies for path construction. In particular, the relaxed EPWT in [26] already covers the path conditions heuristically, where in the numerical implementation "path snakes" are preferred approximating the path smoothness condition.

Acknowledgment

We thank the referees for detailed and helpful comments to improve the paper. We especially thank Jean-Marie Mirebeau, who pointed out to us the strong dependence between the smoothness properties of the restricted functions $\tilde{f}^j$ and the boundedness of derivatives of the path functions $\tilde{p}^j$. This issue had not been considered rigorously in the first version of the paper. A closer look at this problem has led to the path smoothness condition as developed in the present paper that needs to be satisfied in order to obtain best N-term approximation results. Further advices by Jean-Marie led us to the idea to relax the original interpolation condition for the parametric path curve $\tilde{p}^{2J}$ and to replace it by an approximation condition with $\mu > 0$ in (3.1). This relaxation is indeed crucial as shown in the remark at the end of Section 3.2. Finally, we thank him for suggesting the construction of Example 1 (cf. Figure 2) in Section 3.2.

Furthermore, the first author thanks Amos Ron for his useful advice regarding possible path constructions during an early stage of the paper and Robert Schaback for inspiring discussions concerning suitable error estimates for polyharmonic splines interpolation.

This work is supported by the priority program SPP 1324 of the Deutsche Forschungsgemeinschaft (DFG), projects PL 170/13-2 and IS 58/1-2.

References

[1] F. Arandiga, A. Cohen, R. Donat, N. Dyn, and B. Matei, Approximation of piecewise smooth functions and images by edge-adapted (ENO-EA) nonlinear multiresolution techniques, *Appl. Comput. Harmon. Anal.* **24** (2008), 225–250.

[2] E.J. Candès and D.L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise singularities, *Comm. Pure Appl. Math.* **57** (2004), 219–266.

[3] E.J. Candès, L. Demanet, D.L. Donoho, and L. Ying, Fast discrete curvelet transforms, *Multiscale Model. Simul.* **5** (2006), 861–899.

[4] A. Cohen, *Numerical Analysis of Wavelet Methods*. Studies in Mathematics and its Applications **32**, Elsevier, Amsterdam, 2003.

[5] S. Dekel and D. Leviatan, Adaptive multivariate approximation using binary space partitions and geometric wavelets, *SIAM J. Numer. Anal.* **43** (2006), 707–732.

[6] L. Demaret, N. Dyn, and A. Iske, Image compression by linear splines over adaptive triangulations, *Signal Processing* **86** (2006), 1604–1616.

[7] R.A. DeVore, Nonlinear approximation, *Acta Numerica*, 1998, 51–150.

[8] M.N. Do and M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, *IEEE Trans. Image Process.* **14** (2005), 2091–2106.

[9] D.L. Donoho, Wedgelets: Nearly minimax estimation of edges, *Ann. Stat.* **27** (1999), 859–897.

[10] J. Duchon, Splines minimizing rotation-invariant semi-norms in Sobolev spaces. *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller (eds.), Springer, Berlin, 1977, 85–100.

[11] A. Fraysse and S. Jaffard, The Sobolev embeddings are usually sharp, Abstr. Appl. Anal. **2005**(4) (2005), 437–448.

[12] P. Grohs and G. Kutyniok, Parabolic Molecules, Technical University Berlin, preprint 2012.

[13] K. Guo and D. Labate, Optimally sparse multidimensional representation using shearlets, *SIAM J. Math. Anal.* **39** (2007), 298–318.

[14] K. Guo, W.-Q. Lim, D. Labate, G. Weiss, and E. Wilson, Wavelets with composite dilations, *Electron. Res. Announc. Amer. Math. Soc.* **10** (2004), 78–87.

[15] A. Harten, Multiresolution representation of data: general framework, *SIAM J. Numer. Anal.* **33** (1996), 1205–1256.

[16] D. Heinen and G. Plonka, Wavelet shrinkage on paths for denoising of scattered data, *Result. Math.* **62**(3) (2012), 337–354.

[17] A. Iske, On the approximation order and numerical stability of local Lagrange interpolation by polyharmonic splines, *Modern Developments in Multivariate Approximation*, W. Haußmann, K. Jetter, M. Reimer, J. Stöckler (eds.), ISNM 145, Birkhäuser, Basel, 2003, 153–165.

[18] A. Iske, *Multiresolution Methods in Scattered Data Modelling*. Lecture Notes in Computational Science and Engineering **37**, Springer, Berlin, 2004.

[19] L. Jaques and J.-P. Antoine, Multiselective pyramidal decomposition of images: wavelets with adaptive angular selectivity, *Int. J. Wavelets Multiresolut. Inf. Process.* **5** (2007), 785–814.

[20] G. Kutyniok, J. Lemvig, and W.-Q Lim, Optimally Sparse Approximations of 3D Functions by Compactly Supported Shearlet Frames, *SIAM J. Math. Anal.* **44** (2012), 2962–3017.

[21] E. Le Pennec and S. Mallat, Bandelet image approximation and compression, *Multiscale Model. Simul.* **4** (2005), 992–1039.

[22] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1999.

[23] S. Mallat, Geometrical grouplets, *Appl. Comput. Harmon. Anal.* **26** (2009), 161–180.

[24] C.A. Micchelli, Interpolation of scattered data: distance matrices and conditionally positive definite functions. Constr. Approx. **2** (1986), 11–22.

[25] F.J. Narcowich, J.D. Ward, and H. Wendland, Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions, Constr. Approx. **24** (2006), 175–186.

[26] G. Plonka, The easy path wavelet transform: a new adaptive wavelet transform for sparse representation of two-dimensional data, *Multiscale Modelling Simul.* **7** (2009), 1474–1496.

[27] G. Plonka and D. Roşca, Easy path wavelet transform on triangulations of the sphere, *Mathematical Geosciences* **42**(7) (2010), 839–855.

[28] G. Plonka, S. Tenorth, D. Roşca. A hybrid method for image approximation using the easy path wavelet transform. *IEEE Trans. Image Process.* **20**(2) (2011), 372–381.

[29] G. Plonka, S. Tenorth, and A. Iske, Optimally sparse image representation by the easy path wavelet transform. Int. J. Wavelets Multiresolut. Inf. Process. **10**(1) (2012), 1250007 (20 pages).

[30] D.D. Po and M.N. Do, Directional multiscale modeling of images using the contourlet transform, *IEEE Trans. Image Process.* **15** (2006), 1610–1620.

[31] R. Shukla, P.L. Dragotti, M.N. Do, and M. Vetterli, Rate-distortion optimized tree structured compression algorithms for piecewise smooth images, *IEEE Trans. Image Process.* **14** (2005), 343–359.

[32] W. Sickel and H. Triebel, Hölder inequalities and sharp embeddings in function spaces of B^s_{pq} and F^s_{pq} type, Z. Anal. Anwendungen **14**(1), 105–140 (1995).

[33] S. Tenorth, Adaptive Waveletmethoden zur Approximation von Bildern, PhD thesis, University of Göttingen, 2011.

[34] H. Triebel, *Theory of Function Spaces*. Birkhäuser, Basel 1983.

[35] M.B. Wakin, J.K. Romberg, H. Choi, and R.G. Baraniuk, Wavelet-domain approximation and compression of piecewise smooth images, *IEEE Trans. Image Process.* **15** (2006), 1071–108.

[36] J.P. Ward, L^p Bernstein inequalities and inverse theorems for RBF approximation on $\mathbb{R}^d$, J. Approx. Theory **164**(12) (2012), 1577–1593.

[37] H. Wendland, *Scattered Data Approximation*, Cambridge University Press, 2005.

[38] V. Velisavljević, B. Beferull-Lozano, M. Vetterli, and P.L. Dragotti, Directionlets: anisotropic multidirectional representation with separable filtering, *IEEE Trans. Image Process.* **15**(7) (2006), 1916–1933.