Hamburger Beiträge
zur Angewandten Mathematik

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

Gerlind Plonka, Stefanie Tenorth, and Armin Iske

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

GERLIND PLONKA[1], STEFANIE TENORTH[1], and ARMIN ISKE[2]

[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) [20] 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. Using polyharmonic spline interpolation, we show in this paper that the EPWT leads, for a suitable choice of the pathways, to optimal N-term approximations for piecewise Hölder smooth functions with singularities along curves.

Key words. sparse data representation, wavelet transform along pathways, image data compression, adaptive wavelet bases, 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 high-dimensional data, especially for signals. In the one-dimensional case, wavelets are particularly efficient to represent piecewise smooth signals with point singularities. In higher dimensions, however, tensor product wavelet bases are no longer optimal for the representation of piecewise smooth functions with discontinuities along curves.

Just very recently, 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 [10, 11] and directionlets [28] 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 $\mathcal{O}(N^{-1})$ [17]. Up to the $(\log_2 N)^3$ factor, this curvelet approximation result is asymptotically optimal (see [7, Section 7.4]). A similar estimate has been achieved by Guo and Labate [10] 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.

In such relevant cases, one should 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, 12, 15, 16, 18, 20, 21, 22, 24, 25, 27]. In [16], 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 [16] 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, 12] (and references therein), being 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 [20], 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 a simple 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 [21], 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 [22], where an efficient edge representation by the EPWT is combined with favorable properties of the biorthogonal tensor product wavelet transform.

In this paper, we show that piecewise Hölder smooth bivariate functions with singularities along smooth curves can optimally be represented by N-term approximations using the EPWT. More precisely, we prove 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 $\alpha > 0$, with allowing discontinuities along smooth curves of finite length.

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 [20, 23]). If, however, smoother wavelet bases are utilized in the EPWT approach, such an interpretation is not obvious. In fact, while Haar wavelets

admit a straight forward 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, leading to a corresponding local approximation order [13]. We will come back to relevant approximation properties of polyharmonic splines in Section 2.

This paper essentially generalizes our results in [23], where we proved optimal N-term approximation of the EPWT for piecewise Hölder continuous functions of order $\alpha \in (0, 1]$. In that paper, the proofs were mainly based on the adaptive multiresolution analysis structure, which is only available for piecewise constant Haar wavelets.

The outline of this paper is as follows. In Section 2, we first introduce the utilized function model 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 the further levels, we require specific side conditions for the path vectors in the EPWT algorithm. These side conditions can be ensured by a suitable path vector construction as proposed in Section 2. Similar conditions for the path vectors have been used already in [23]. 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 $F \in L^2([0, 1)^2)$ is a piecewise smooth bivariate function, being smooth over a finite set of regions $\{\Omega_i\}_{1 \leq i \leq K}$, where each region Ω_i has a sufficiently smooth Lipschitz boundary $\partial \Omega_i$ of finite length. Moreover, the set $\{\Omega_i\}_{1 \leq i \leq K}$ is assumed to be a disjoint partition of $[0, 1)^2$, so that

$$\bigcup_{i=1}^{K} \Omega_i = [0, 1)^2,$$

where each closure $\overline{\Omega}_i$ is a connected subset of $[0, 1]^2$, for $i = 1, \ldots, K$.

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

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

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

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

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

3

that the Hölder space $C^\alpha(\Omega_i)$ of order $\alpha > 0$, being equipped with the norm

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

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

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

see e.g. [4, chapter 3.2]. Now by uniform sampling, the bivariate function F is assumed to be given by its function values taken at a finite rectangular grid. For a suitable integer $J > 1$, let $\{F(2^{-J}n)\}_{n \in I_{2J}}$ be the given samples of F, where

$$I_{2J} := \{n = (n_1, n_2) : 0 \leq n_1 \leq 2^J - 1, \ 0 \leq n_2 \leq 2^J - 1\},$$

and, moreover, let

$$\Gamma_i^{2J} := \left\{ n \in I_{2J} : \frac{n}{2^J} \in \Omega_i \right\} \qquad \text{for } 1 \leq i \leq K$$

be the index sets of grid points that are contained in the regions Ω_i, for $1 \leq i \leq K$. 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 any $1 \leq i \leq K$.

Next we compute a (piecewise) sufficiently smooth approximation to F from its given samples. To this end, we apply polyharmonic spline interpolation to obtain an interpolation to F of the form

$$F^{2J}(x) := \sum_{i=1}^K \left(\sum_{n \in \Gamma_i^{2J}} c_n^i \, \phi_m \left(\left\| x - \frac{n}{2^J} \right\|_2 \right) + p_m^i(x) \right) \chi_{\Omega_i}(x) \qquad \text{for } x \in [0,1)^2 \qquad (2.2)$$

satisfying the interpolation conditions

$$F^{2J}(n/2^J) = F(n/2^J) \qquad \text{for all } n \in I_{2J}. \qquad (2.3)$$

Here, χ_{Ω_i} is the characteristic function of Ω_i, $\phi_m(r) = r^{2m} \log(r)$, for $m := \max(\lfloor \alpha \rfloor, 2)$, is a fixed *polyharmonic spline* kernel, and p_m^i denotes a bivariate polynomial of degree at most m. It is well-known that polyharmonic spline interpolation leads, on given interpolation conditions (2.3), to a *unique* interpolant of the form (2.2). In particular, for the specific choice of a polyharmonic spline kernel ϕ_m, the interpolation scheme achieves to reconstruct polynomials of degree m. Consequently, the *local approximation order* of polyharmonic spline interpolation is $m + 1$, see [13]. Further relevant details on polyharmonic spline interpolation and their approximation properties can be found in [14, Section 3.8].

From the embedding theorem for Besov spaces we obtain $B^\alpha_{\infty,\infty}(\Omega_i) \subset B^\alpha_{2,2}(\Omega_i)$, see e.g. [26] or [4, page 163]. Since $B^\alpha_{2,2}(\Omega_i)$ is equivalent to the Sobolev space $H^\alpha(\Omega_i)$, see [4, 26], this allows us to use the estimate

$$\|F - F^{2^J}\|_{L^2(\Omega)} \le C_F \sum_{i=1}^{K} h^\alpha_{\Omega_i} \|F\|_{B^\alpha_{2,2}(\Omega_i)} \tag{2.4}$$

for the interpolation error in Sobolev spaces, as shown in [19], where the *fill distance*

$$h_{\Omega_i} := \sup_{x \in \Omega_i} \inf_{n \in \Gamma_i^{2^J}} \|x - \frac{n}{2^J}\|_2 \le 2^{-J} \qquad \text{for } 1 \le i \le K$$

measures the density of the interpolation points in Ω_i.

Remark. Note that by the above representation (2.2), we apply polyharmonic spline interpolation separately in each individual region Ω_i. We prefer interpolation (rather than any other projection method), since we essentially require to maintain the local Hölder regularity around each lattice point in I_{2^J}. □

2.2 The EPWT Algorithm

Now let us briefly recall the EPWT algorithm from our previous work [20]. To this end, let $\varphi \in C^\beta$ with $\beta \ge \alpha$ 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 $V_0 := \text{clos}_{L^2} \text{span} \{\varphi(\cdot - k) : k \in \mathbb{Z}\}$. 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 $j, k \in \mathbb{Z}$, we use the notation

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

likewise for $\tilde\varphi$ and $\tilde\psi$. The functions $\varphi, \tilde\varphi$ 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 any polynomial p_m of degree less than or equal $m = \max(\lfloor \alpha \rfloor, 2)$, and so,

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

5

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 any 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

$$c^{j-1}(n) = \langle f^j, \tilde{\varphi}_{j-1,n}\rangle \quad \text{and} \quad d^{j-1}(n) = \langle f^j, \tilde{\psi}_{j-1,n}\rangle. \tag{2.5}$$

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 index subsets of I_{2J}. For the characterization of suitable path vectors we first need to introduce *neighborhoods* of indices. For any index $n = (n_1, n_2) \in I_{2J}$, we define its neighborhood by

$$N(n) := \{m = (m_1, m_2) \in I_{2J} \setminus \{n\} : \|n - m\|_2 \leq \sqrt{2}\},$$

where $\|n - m\|_2^2 = (n_1 - m_1)^2 + (n_2 - m_2)^2$. Hence, an interior index, i.e., an index that does not lie on the boundary of the index domain I_{2J}, has eight neighbors.

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 index set I_{2J}. This path vector is a suitable permutation of all indices in I_{2J}, which can be determined by using the following strategy from [20]. Recall that $I_{2J} = \cup_{i=1}^{K}\Gamma_i^{2J}$, where Γ_i^{2J} corresponds to lattice points in Ω_i. Start with one index $p^{2J}(0)$ in Γ_1^{2J}. Now, for a given n-th component $p^{2J}(n)$ being contained in the index set Γ_i^{2J}, for some $i \in \{1, \ldots, K\}$, we choose the next component $p^{2J}(n+1)$ of the path vector p^{2J}, such that

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

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

In situations where $(N(p^{2J}(n)) \cap \Gamma_i^{2J}) \setminus \{p^{2J}(0), \ldots p^{2J}(n)\}$ is an empty set, the path is *interrupted*, and we need to start a *new pathway* by choosing the next index $p^{2J}(n+1)$ from $\Gamma_i^{2J} \setminus \{p^{2J}(0), \ldots, p^{2J}(n)\}$. If, however, this set is also empty, we choose $p^{2J}(n+1)$ from the set of remaining indices $I_{2J} \setminus \{p^{2J}(0), \ldots, p^{2J}(n)\}$. For a more detailed description of the path vector construction we refer to [20].

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, see [23]. The so obtained vector p^{2J} is composed of *connected* pathways, i.e., each pair of consecutive components in these pathways are neighboring.

Now, we consider the data vector

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

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 (2.5). 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, where only the wavelet coefficients corresponding to an interruption (from one region to another) may possess significant amplitudes.

The path vector p^{2J} determines a new subset of indices

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

where $\Gamma_i^{2J-1} := \{p^{2J}(2\ell) : p^{2J}(2\ell) \in \Gamma_i^{2J}\}$.

As regards the next level of the EPWT, where $j = 2J - 1$, 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 indices in Γ^{2J-1}. Similar as before, we require that $p^{2J-1}(n)$ and $p^{2J-1}(n+1)$ are neighbors lying in the same index 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

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

Again, the number of path interruptions can be bounded by $C_1 K$, where C_1 does not depend on J. Then we apply one level of the one-dimensional wavelet transform to the permuted data vector $(c^{2J-1}(p^{2J-1}(\ell)))_{\ell=0}^{2^{2J-1}-1}$ to 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 $j + 1$ for $j = 2J - 3, \ldots, 0$, where at any level $j + 1$ we first construct a path $p^{j+1} = (p^{j+1}(\ell))_{\ell=0}^{2^{j+1}-1}$ through the index set

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

with applying similar strategies as described above. Here, $p^{j+1}(n)$ and $p^{j+1}(n+1)$ are called neighbors, iff
$$\left\|p^{j+1}(n) - p^{j+1}(n+1)\right\|_2 \leq D 2^{J-(j+1)/2},$$

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 $(c^{j+1}(p^{j+1}(\ell))_{\ell=0}^{2^{j+1}-1}$, yielding c^j and d^j.
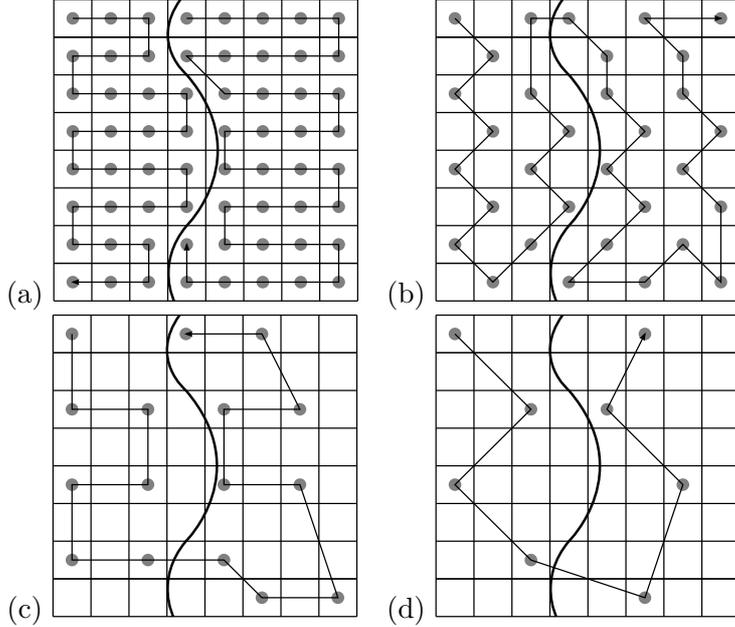
Figure 1: Path construction. (a) level six, (b) level five, (c) level four, (d) level three.

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 $[0,1)^2$ be divided into only two regions, Ω_1 and Ω_2. The function F is assumed to be Hölder smooth 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 in the underlying index set $I_6 = \Gamma_1^6 \cup \Gamma_2^6$, such that each pair of consecutive components in the path are neighbors. We first pick all indices 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 index set $\Gamma^5 = \Gamma_1^5 \cup \Gamma_2^5$ (containing only each second index 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 index sets Γ^4 and Γ^3 along with their corresponding path vectors.

In this example, we have

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

(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 $D = \sqrt{10}/2 \approx 1.5811$. This simple example also illustrates that the path construction leads at each level to index sets Γ_i^j with quasi-uniformly distributed indices. Further details concerning the required quasi-uniformity of the index distributions are explained in Section 3.2.

8

Remark 1. Note that the components of the path vector p^j lie in Γ^{2J} with containing 2d integer entries. This is in contrast to the notation in [20]. Further, unlike in [20], we do no longer consider index sets but define a neighborhood of pixels by the Euclidean distance between corresponding indices. The constant D should be chosen rather small, e.g. $D \in (\sqrt{2}, 2)$, as in the previous example. We remark that the choice of a sufficiently small constant D leads to a sequence of paths, where the distribution of the indices in Γ^j is quasi-uniform at each level of the EPWT, see Section 3.2.

Remark 2. 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.

To show this fact also theoretically, we have proven in [23] that for piecewise Hölder continuous functions of order $\alpha \in (0, 1]$, we obtain optimal convergence rates for N-term approximations when using univariate Haar wavelets. However, our numerical experiments show that the EPWT is much more effective when smoother wavelet bases are used, e.g. the Daubechies wavelets with two vanishing moments or the $7 - 9$ biorthogonal transform.

Therefore, the goal of this paper is to show that a higher Hölder smoothness of the target function within the different regions leads to optimal N-term approximations, when using the EPWT in combination with smooth wavelet bases. In this sense, the results of this paper can be viewed as a generalization of our previous results in [23].

Remark 3. In contrast to the above procedure, where we have used the data vectors $\left(c^j(\ell) \right)_{\ell=0}^{2^j-1} = \left(F^j(2^{-J}p^j(\ell) \right)_{\ell=0}^{2^j-1}$, we will consider slightly different data vectors $c_p^j(\ell)$ in the theoretical estimates of Sections 3 and 4, where we will be using L^2-projection operators determined by the dual scaling and wavelet functions, $\tilde{\varphi}$ and $\tilde{\psi}$.

3 Decay of Wavelet Coefficients using 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.

As already explained in the previous section, we consider applying polyharmonic spline interpolation, from given image values $F(2^{-J}n)$, separately in the individual domains Ω_i, $i = 1, \ldots, K$. We assume that F is Hölder smooth of order α on each Ω_i, and that the polyharmonic spline interpolant F^{2J} in (2.2), being Hölder smooth of order at least α, reconstructs bivariate polynomials of degree $m = \max(\lfloor \alpha \rfloor, 2)$, thus yielding *local* approximation order $m + 1$.

At the $2J$-th level of the EPWT, we define a path p^{2J} through all indices of I_{2J} such that consecutive components of p^{2J} are neighboring indices lying in the same index set Γ_i. The path p^{2J} can be constructed in a way such that only $C_1 K$ "interruptions" occur. Next, we consider a one-dimensional function $f^{2J}(t) = \sum_{k=0}^{2^{2J}-1} c_p^{2J}(k) \varphi_{2J,k}(t)$, $t \in [0, 1)$ that suitably approximates a smooth one-dimensional and scaled restriction of F resp. F^{2J}

"along the path p^{2J}" with

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

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

Significant wavelet coefficients will only occur at a finite number of locations on the interval $[0, 1)$ that correspond to interruptions of the path. However, the number of such interruptions does not depend on J but only 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 levels of the EPWT, we consider the index sets $\Gamma^j = \cup_{i=1}^K \Gamma_i^j$. By construction, we have the inclusion $\Gamma^j \subset \Gamma^{j+1}$, for $j = 1, \ldots, 2J - 1$, and, moreover, $\#\Gamma^j = 2^j$. To obtain sufficiently accurate polyharmonic spline interpolations F^j to F with $F^j(2^{-J}n) = F(2^{-J}n)$ for $n \in \Gamma^j$, two specific conditions on the index sets Γ^j and the path vectors p^j need to be satisfied. We can briefly explain these two conditions, termed *region 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 should prefer to traverse the indices belonging to one region set Γ_i^j, before "jumping" to another region. In this way, the region condition ensures that we can optimally exploit the smoothness of the function F along the path. Secondly, the diameter condition requires a quasi-uniform distribution of remaining pixels in each Γ_i^j, and so the diameter condition leads to a sufficiently accurate polyharmonic spline interpolation F^j at each level of the EPWT.

We remark that the two conditions can be satisfied by using the strategies for the path construction as proposed in Subsection 2.2. This then allows 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 further levels of the EPWT and the corresponding estimates are then derived in a similar manner.

We consider a sufficiently smooth parametric curve $\tilde{p}^{2J}(t)$, $t \in [0, 1)$, through the plane interpolating the path p^{2J}, i.e., with $\tilde{p}^{2J}(\ell/2^{2J}) = 2^{-J}p^{2J}(\ell)$, for $\ell = 0, \ldots, 2^{2J} - 1$, and with the convention that $\tilde{p}^{2J}(t) \in \Omega_i$, for $t \in [\ell/2^{2J}, (\ell+1)/2^{2J}]$, if $2^{-J}p^{2J}(\ell)$ and $2^{-J}p^{2J}(\ell+1)$ 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 $\tilde{p}^{2J}$,

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

For each interruption in the path vector p^{2J}, originated within one region Ω_i or by a jump from one region to another, there are indices ℓ and $\ell+1$, where $p^{2J}(\ell)$ and $p^{2J}(\ell+1)$ are not neighbors or where there is a discontinuity between $F^{2J}(2^{-J}p^{2J}(\ell))$ and $F^{2J}(2^{-J}p^{2J}(\ell+1))$. These interruptions correspond to small subintervals of $[0, 1)$ of length 2^{-2J}, where the univariate function $\tilde{f}^{2J}$ may also have discontinuities. More precisely, an interruption

10

between the path components $p^{2J}(\ell)$ and $p^{2J}(\ell+1)$ generates such a "jump interval" at $[\ell/2^{2J}, (\ell+1)/2^{2J}]$. For simplicity, we assume that we only have path interruptions from one region to another. Note that this is only a mild restriction, since we may otherwise decide to further subdivide a region into several smaller subregions according to the regularity of F along the paths. Finally, we remark that for any convex region Ω_i we can show that there is at least one path without any interruptions, see [23].

Recall that the trace theorem for Hölder resp. Besov spaces (see [26]) 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 $\{t \in [0,1) : \tilde{p}^{2J}(t) \in \Omega_i\}$ with assuming that the corresponding path vector p^{2J} has no interruptions. In particular, we obtain for the N-th order modulus of smoothness the estimate

$$\omega_N(\tilde{f}^{2J}, h)_\infty := \sup_{|\tilde{h}| \le h} \|\Delta^N_{\tilde{h}} \tilde{f}^{2J}\|_\infty \lesssim (2^J h)^\alpha \|\tilde{f}^{2J}\|_{B^\alpha_{\infty,\infty}} \tag{3.1}$$

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

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

see [4]. Observe that the factor $2^{J\alpha}$ in (3.1) is due to the scaling of $\tilde{f}^{2J}$ on $[0,1]$ while the length of the complete curve $\tilde{p}^{2J}$ is $c\, 2^J$, where the constant c does not depend on J.

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

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

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. Along the lines of [4, Theorem 3.3.3], we now have in the subintervals $T_{i,2^{-2J}}$

$$\begin{aligned}
\|\tilde{f}^{2J} - f^{2J}\|_{L^\infty(T_{i,2^{-2J}})} &= \|\tilde{f}^{2J} - P_{2J}\tilde{f}^{2J}\|_{L^\infty(T_{i,2^{-2J}})} \\
&\lesssim \omega_N(\tilde{f}^{2J}, 2^{-2J})_\infty \lesssim (2^{-J})^\alpha \|\tilde{f}^{2J}\|_{B^\alpha_{\infty,\infty}}.
\end{aligned} \tag{3.2}$$

In particular,

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

In the next step, we decompose the function $f^{2J} = \sum_\ell c^{2J}_p(\ell) \varphi_{2J,\ell}$ with $c^{2J}_p(\ell) := \langle f^{2J}, \tilde{\varphi}_{2J,\ell} \rangle$ 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 $(c^{2J}_p(\ell))_{\ell=0}^{2^{2J}-1}$, we obtain the decomposition $f^{2J} = f^{2J-1} + g^{2J-1}$ with

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

where $c^{2J-1}_p(n) := \langle f^{2J}, \tilde{\varphi}_{2J-1,n} \rangle$ and $d^{2J-1}_p(n) := \langle f^{2J}, \tilde{\psi}_{2J-1,n} \rangle$. From the Hölder smoothness of f^{2J} in $T_i := \{t \in [0,1) : \tilde{p}^{2J}(t) \in \Omega_i\}$, we find for $t \in T_i$ the representation

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

11

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 $|R(t - t_0)| \le c_\varphi 2^{-J\alpha}$. Hence, if $\mathrm{supp}(\tilde{\psi}_{2J-1,n}) \in T_i$ for some i, the wavelet coefficients satisfy

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

where we have used $\|\tilde{\psi}_{2J-1,n}\|_1 = 2^{-J+1/2}\|\tilde{\psi}\|_1$.

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 wavelet coefficients $2^{2J-1} - \#\Lambda^{2J-1}$ corresponds to the number of discontinuities of $\tilde{f}^{2J}$ and is hence bounded by CK, where K is the number of regions in the original image F, and where the constant C does not depend on 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 components of p^{2J} with even indices, we put

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

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

$$
F^{2J-1}(x) := \sum_{i=1}^K \left(\sum_{y \in \Gamma_i^{2J-1}} c_y^i \, \phi_m \left(\left\| x - \frac{y}{2^J} \right\|_2 \right) + p_m^i(x) \right) \chi_{\Omega_i}(x),
$$

satisfying the interpolation conditions

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

Therefore,

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

where $D = \sqrt{2}$, and where the last inequality again follows analogously as in (3.2) since f^{2J-1} is the orthogonal projection of $\tilde{f}^{2J}$ to

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

The last inequality implies that F^{2J-1} is still a good approximation for F, since the interpolation points have changed only slightly. However, only half of the interpolation points are left, which are irregularly distributed in $[0,1)^2$. Moreover, we have

$$
\max_{x \in \Omega_i} \min_{y \in \Gamma_i^{2J-1}} |x - 2^{-J}y| \le 2^{-J+1} = D \, 2^{-J+1/2} \quad \text{and} \quad \min_{y_1, y_2 \in \Gamma_i^{2J-1}} |y_1 - y_2| \ge 2^{-J}.
$$

Together with (2.4) we observe

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

12

3.2 Conditions for the Path Vectors

Before we proceed with the error estimates for the further levels of the EPWT algorithm, we need to fix two specific side conditions for the path vectors that are required for our error analysis. The two side conditions are termed (a) *region condition* and (b) *diameter condition*, as stated below.

Similar conditions have already been applied to the N-term estimates in [23]. The region condition ensures that at each level of the EPWT, the path vector should be chosen in a way such that all indices belonging to one region Ω_i, $i \in \{1, \dots, K\}$, should be taken first before jumping to another region, and that one should have only a small number of index pairs $p^j(\ell)$, $p^j(\ell + 1)$ belonging to different regions Γ_i^j and Γ_k^j. The diameter condition ensures that the remaining indices in Γ^j are quasi-uniformly distributed, such that there is a constant D, not depending on J or j, satisfying

$$\max_{x \in \Omega_i} \min_{y \in \Gamma^j} \|x - 2^{-J}y\|_2 \leq D2^{-j/2}.$$

Let us introduce the two conditions more explicitly.

(a) **Region condition.** At each level j of the EPWT, the path p^j is chosen, such that it contains only at most $C_1 K$ interruptions caused by pairs of components, $p^j(\ell)$ and $p^j(\ell + 1)$, which are no neighbors or are not belonging to the same index set Γ_i^j.

(b) **Diameter condition.** At each level of the EPWT, we require for almost all components of the path the bound

$$\|p^j(\ell) - p^j(\ell + 1)\|_2 \leq D\, 2^{J-j/2}, \tag{3.3}$$

where D is independent of J and j, and where the number of components of p^j which are not satisfying the diameter condition, is bounded by a constant C_2 not depending on J or j.

The two conditions, (a) and (b), are already ensured by the proposed path construction in Subsection 2.2. Particularly, the diameter condition ensures a quasi-uniform distribution of the indices in Γ^j.

3.3 The Further Levels of the EPWT

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

$$F^{j+1}(x) := \sum_{i=1}^{K} \left(\sum_{y \in \Gamma_i^{j+1}} c_y^i\, \phi_m \left(\left\| x - \frac{y}{2^J} \right\|_2 \right) + p_m^i(x) \right) \chi_{\Omega_i}(x)$$

satisfying the interpolation conditions

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

13

We first fix a suitable path vector p^{j+1} passing through the set

$$\Gamma^{j+1} = \{p^{j+2}(2n) : n = 0, \ldots, 2^{j+1} - 1\}$$

with corresponding data values $\{F^{j+1}\left(\frac{p^{j+2}(2n)}{2^J}\right) : n = 0, \ldots, 2^{j+1} - 1\}$, such that the region condition and the diameter condition in Section 3.2 are satisfied.

Next, we consider a sufficiently smooth parametric curve $\tilde{p}^{j+1}(t)$, $t \in [0, 1)$, through the plane interpolating p^{j+1}, i.e., with $\tilde{p}^{j+1}(2^{-j-1}\ell) = 2^{-J} p^{j+1}(\ell)$ for $\ell = 0, \ldots, 2^{j+1} - 1$, such that $\tilde{p}^{j+1}(t) \in \Omega_i$ for $t \in [2^{-(j+1)}\ell, 2^{-(j+1)}(\ell + 1)]$, if $2^{-J} p^{j+1}(\ell)$ and $2^{-J} p^{j+1}(\ell + 1)$ are in the same region Ω_i.

Then we determine the one-dimensional restriction of F^{j+1} along $\tilde{p}^{j+1}$,

$$\tilde{f}^{j+1}(t) := F^{j+1}\left(\tilde{p}^{j+1}(t)\right) \qquad t \in [0, 1).$$

As in the above discussion concerning the $2J$-th level, the piecewise Hölder smoothness of $\tilde{f}^{j+1}(t)$ ensures that

$$\omega_N(\tilde{f}^{j+1}, h)_\infty \le \tilde{c}\,(2^J h)^\alpha \qquad \text{for} \qquad T_{i,h} := \{t : \tilde{p}^{j+1}(t + kh) \in \Omega_i,\ k = 0, \ldots, N\}.$$

Considering the L^2-projection

$$f^{j+1} = P_{j+1}\tilde{f}^{j+1} := \sum_{n=0}^{2^{j+1}-1} c_p^{j+1}(n)\,\varphi_{j+1,n}$$

with $c_p^{j+1}(n) := \langle \tilde{f}^{j+1}, \tilde{\varphi}_{j+1,n} \rangle$ of $\tilde{f}^{j+1}$ onto the scaling space

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

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

$$
\begin{aligned}
\|\tilde{f}^{j+1} - f^{j+1}\|_{L^\infty(T_{i,2^{-(j+1)}})} &= \|\tilde{f}^{j+1} - P_{j+1}\tilde{f}^{j+1}\|_{L^\infty(T_{i,2^{-(j+1)}})} \\
&\lesssim w_N(\tilde{f}^{j+1}, 2^{-J-(j+1)/2}) \lesssim 2^{-(j+1)\alpha/2},
\end{aligned}
$$

where we have used the diameter condition (3.3), so that

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

We use the decomposition

$$
\begin{aligned}
f^{j+1} &= \sum_\ell c_p^{j+1}(\ell)\,\varphi_{j+1,\ell} \\
&= f^j + g^j = \sum_n c_p^j(n)\,\varphi_{j,n} + \sum_n d_p^j(n)\,\psi_{j,n}.
\end{aligned}
$$

Then the Hölder smoothness of f^{j+1} in the intervals T_i yields the Taylor expansion

$$f^{j+1}(t) = q_\alpha(t - t_0) + R(t - t_0) \quad \text{with } |R(t - t_0)| \le c_\varphi\,D^\alpha\,2^{-(j+1)\alpha/2}$$

14

for $t, t_0 \in T_i$, and this gives the following estimate for the wavelet coefficients corresponding to the region Ω_i:

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

Again, let Λ^j be the set of indices n from $\{0, \ldots, 2^j - 1\}$, where $d_p^j(n)$ satisfies the above estimate. Then the number of wavelet coefficients $2^j - \#\Lambda^j$ which are not satisfying this estimate (since supp $\tilde{\psi}_{j,n} \not\subset T_i$ for some i) is bounded by a constant independent of J and j.

Finally, we obtain the polyharmonic spline interpolant

$$F^j(x) := \sum_{i=1}^{K} \left(\sum_{y \in \Gamma_i^j} c_y^i \, \phi_m \left(\left\| x - \frac{y}{2^J} \right\|_2 \right) + p_m^i(x) \right) \chi_{\Omega_i}(x),$$

where $\Gamma_i^j := \{ p^{j+1}(2n) : n = 0, \ldots, 2^j - 1, \ 2^{-J} \, p^{j+1}(2n) \in \Omega_i \}$, $\Gamma^j := \cup_{i=1}^K \Gamma_i^j$, through the interpolation conditions

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

Hence, we obtain the estimate

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

Particularly,

$$\left| F^{2J} \left(\frac{p^{j+1}(2n)}{2^J} \right) - F^j \left(\frac{p^{j+1}(2n)}{2^J} \right) \right| \le \sum_{\nu=j}^{2J-1} \left| F^{\nu+1} \left(\frac{p^{j+1}(2n)}{2^J} \right) - F^\nu \left(\frac{p^{j+1}(2n)}{2^J} \right) \right|$$

$$\lesssim \sum_{\nu=j}^{2J-1} 2^{-(\nu+1)\alpha/2} \le \frac{2^{-(j+1)\alpha/2}}{1 - 2^{-\alpha/2}}. \tag{3.4}$$

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

Theorem 3.1 *For $j = 2J - 1, \ldots, 0$, let $d_p^j(\ell) = \langle f^{j+1}, \tilde{\psi}_{j,\ell} \rangle$, $\ell = 0, \ldots, 2^j - 1$, denote the wavelet coefficients that are obtained by applying the EPWT algorithm to F (according to Subsection 2.2), where we assume that $F \in L^2([0,1)^2)$ is piecewise Hölder smooth of order α as prescribed in Subsection 2.1. Further assume that the path vectors $(p^{j+1}(\ell))_{\ell=0}^{2^{j+1}-1}$, $j = 2J - 1, \ldots, 0$, in the EPWT algorithm satisfy the region condition (a) and the diameter condition (b) of Subsection 3.2. Then, for all $j = 2J - 1, \ldots, 0$ and $\ell \in \Lambda^j$, the estimate*

$$|d_p^j(\ell)| \le C \, D^\alpha \, 2^{-j(\alpha+1)/2} \tag{3.5}$$

holds, where $D > 1$ is the constant of the diameter condition (3.3), 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 $\ell \in \{0, \ldots, 2^j - 1\} \setminus \Lambda^j$, we obtain the estimate

$$|d_p^j(\ell)| \le C' \, 2^{-j/2} \tag{3.6}$$

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

Proof. The proof of (3.5) follows directly from the above considerations. Likewise, for all $\ell \in \{0, \ldots, 2^j - 1\} \setminus \Lambda^j$, i.e., for index sets that do not satisfy the diameter or the region condition, we observe at least

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

since we can assume that F^j is bounded, and hence the above estimate (3.5) holds for $\alpha = 0$. Hence (3.6) follows. $\qquad \square$

4 N-term Approximation with the EPWT

Consider now the vector of all EPWT wavelet coefficients

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

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

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

together with the side information on the path vectors in each iteration step

$$\mathbf{p} = ((p^{2J})^T, \, \ldots, (p^1)^T)^T \in \mathbb{R}^{2(2^{2J}-1)}.$$

With this information the image F_{rec}^{2J} is uniquely recovered, where F_{rec}^{2J} is the polyharmonic spline interpolation satisfying

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

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 f^{j+1} uniquely. Further, the polyharmonic spline interpolation F^{j+1} is entirely determined by the function values $\tilde{f}^{j+1}(2^{-(j+1)}n)$.

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

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

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

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

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 3.1 and (3.4). Indeed, at each level of the EPWT, we observe that

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

where the number of wavelet coefficients satisfying (3.5) is $2^j - C_1 K + C_2$, and where the constants C_1 and C_2 do not depend on J or 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 3.1 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 [23]. $\square$

Let us finally conclude by stating the following corollary.

Corollary 4.2 *Let $F \in L^2([0,1)^2)$ be piecewise Hölder continuous (as assumed in Subsection 2.1). Then, for any $\epsilon > 0$ 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 (2.4). $\square$

Acknowledgment

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

[11] K. Guo, W.-Q. Lim, D. Labate, G. Weiss, and E. Wilson, Wavelets with composite dilations, *Electr. res. Announc. of AMS* **10** (2004), 78–87.

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

[13] 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.

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

[15] 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.

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

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

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

[19] 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.

[20] 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.

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

[22] G. Plonka, S. Tenorth, D. Roşca. A hybrid method for image approximation using the easy path wavelet transform. *IEEE Trans. Image Process.*, 2011, to appear, DOI 10.1109/TIP.2010.2061861.

[23] G. Plonka, S. Tenorth, and A. Iske, Optimally sparse image representation by the easy path wavelet transform. Preprint, 2010.

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

[25] 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.

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

[27] 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.

[28] 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.