DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Optimally Sparse Image Representation by the Easy Path Wavelet Transform

G. Plonka, S. Tenorth, A. Iske

Preprint 15

Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Optimally Sparse Image Representation by the Easy Path Wavelet Transform

G. Plonka, S. Tenorth, A. Iske

Preprint 15

The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

Optimally Sparse Image Representation by the Easy Path Wavelet Transform

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

[1] Department of Mathematics, University of Duisburg-Essen,
Campus Duisburg, 47048 Duisburg, Germany
gerlind.plonka@uni-due.de, stefanie.tenorth@uni-due.de

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

Dedicated to Wolfgang Dahmen on the occasion of his 60th birthday

Abstract

The **Easy Path Wavelet Transform** (EPWT) [19] 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 show that the EPWT leads, for a suitable choice of the pathways, to optimal N-term approximations for piecewise Hölder continuous 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 [12, 13] and directionlets [24] are examples for non-adaptive highly redundant function frames with strong anisotropic directional selectivity.

For piecewise Hölder continuous 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

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

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 [9], Section 7.4). A similar estimate has been achieved by Guo and Labate [12] 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, 7, 8, 10, 11, 14, 15, 16, 18, 19, 20, 21, 22, 23]. 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, 14] (and references therein), being based on ENO reconstructions. We remark that the resulting ENO-EA schemes lead to an optimal N-term approximation, yielding $\|f - f_N\|^2 \le C\, N^{-2}$ 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 many relevant applications to image denoising and image regularization, the space BV containing all functions of bounded variation plays an important role. The space BV seems to be well-adapted to model natural images, since it also allows sharp edges, unlike the bivariate Besov space $B_1^1(L_1)$, see [4]. However, in case of BV spaces, there is no simple data representation, e.g. in terms of wavelet coefficients. For useful relations between Haar decompositions on dyadic rings and the space BV we refer to [5].

In this paper, we prove optimal N-term approximations for a function class being very close to the space BV. For this purpose, we use the locally adaptive *Easy Path Wavelet Transform* (EPWT) which has recently been explored in our previous paper [19]. The EPWT applies a one-dimensional wavelet transform along suitable pathways of data vectors, where local correlations of the given data are essentially exploited. As supported by our numerical experiments in [19], the EPWT leads to an efficient compression method for two-dimensional digital data, especially for image data. But in this paper we focus on the approximation properties of the EPWT, particularly for piecewise smooth images. More precisely, we show that the application of the EPWT leads to an N-term approximation of the form

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

for piecewise Hölder continuous functions of order α (with $0 < \alpha \le 1$) with allowing discontinuities along curves of finite length.

The outline of this paper is as follows. In Section 2, we first introduce the EPWT algorithm, before we apply the resulting approximation method to target functions from a function class to be described. To this end, we recall the basic ideas of the EPWT, where we show that it generates a data-dependent multiresolution analysis and a corresponding adaptive Haar wavelet basis. In order to achieve optimal approximation results, we require specific side conditions for the path vectors that are used in the EPWT algorithm. These side conditions are derived in Subsection 2.3, and further illustrations are given through a numerical example in Subsection 2.4. Finally, Section 3 is devoted to error analysis, where we prove asymptotically optimal N-term error estimates of the form (1.1) for piecewise Hölder continuous functions.

2 EPWT and Adaptive Haar Wavelet Bases

2.1 The EPWT Algorithm

Suppose that $F \in L^2([0,1)^2)$ is a piecewise regular image, being uniformly regular over a finite set of regions $\{\Omega_i\}_{1\le i\le K}$, each of whose boundaries $\partial\Omega_i$ is continuous and of finite length. Moreover, the set $\{\Omega_i\}_{1\le i\le K}$ 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 assumed to be a connected subset of $[0,1]^2$, for $i = 1,\ldots,K$. Furthermore, we assume that F satisfies a Hölder condition in each region Ω_i, $1 \le i \le K$, i.e.,

$$|F(x) - F(x+h)| \le C\|h\|_2^\alpha, \qquad \text{for } x, x+h \in \Omega_i \tag{2.1}$$

for some $\alpha \in (0,1]$ and $C > 0$ which do not depend on i. But F may be discontinuous across the boundaries between adjacent regions.

With assuming that F represents a digital image, the image is given by uniform samples of F over a rectangular grid. For a suitable given integer $J > 1$, let $\{F(2^{-J}n)\}_{n\in I_J}$ be the given samples of F, where $I_J := \{n = (n_1, n_2) : 0 \le n_1 \le 2^J - 1,\, 0 \le n_2 \le 2^J - 1\}$. We regard the piecewise constant function

$$F^{2J}(x) := \sum_{n\in I_J} F(2^{-J}n)\, \chi_{[0,1)^2}(2^J x - n) \qquad \text{for } x \in [0,1)^2$$

as an approximation to F in $L^2([0,1)^2)$. Moreover, by

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

we denote the index set of grid points that are contained in region Ω_i, for $1 \le i \le K$. Obviously,

$$\bigcup_{i=1}^{K} \Gamma_i^J = I_J.$$

Furthermore, for the size $\#\Gamma_i^J$ of Γ_i^J we have $\#\Gamma_i^J \le 2^{2J}$. Consequently, (2.1) yields the error bound

$$|F^{2J}(2^{-J}n) - F^{2J}(2^{-J}m)| \le C\, 2^{-J\alpha}\, \|n - m\|_2^\alpha, \tag{2.2}$$

provided that $n, m \in \Gamma_i^J$ for some $1 \leq i \leq K$.

Remark. Note that the data F^{2J} is obtained from F by interpolation, rather than by its L^2-projection onto the linear space

$$\mathrm{span}\{\chi_{[0,1)^2}(2^J \cdot -n) : n \in I_J\}$$

of piecewise constant functions. We prefer to work with interpolation of F in order to derive the error estimate (2.2) from the Hölder continuity of F in (2.1). $\square$

Now let us briefly recall the EPWT algorithm from our previous work [19]. For the sake of simplicity, we use the (one-dimensional) Haar wavelet basis. To this end, let

$$\phi_{j,k}(t) := 2^{j/2}\,\phi(2^j t - k) \qquad \text{and} \qquad \psi_{j,k}(t) := 2^{j/2}\,\psi(2^j t - k), \tag{2.3}$$

with $\phi(t) := \chi_{[0,1)}(t)$ and $\psi(t) := \chi_{[0,1/2)}(t) - \chi_{[1/2,1)(t)}$, where for any interval $[a, b) \subset \mathbb{R}$, $\chi_{[a,b)}$ denotes its characteristic function.

The EPWT is a wavelet transform that works along path vectors through index subsets of I_J. For the characterization of suitable path vectors (in Subsection 2.3), we first need to introduce *neighborhoods* of indices and index sets. For any index $n = (n_1, n_2) \in I_J$, we define its neighborhood by

$$N(n) := \{m = (m_1, m_2) \in I_J \setminus \{n\} : |n_1 - m_1| \leq 1 \text{ and } |n_2 - m_2| \leq 1\}.$$

Hence, an interior index, i.e., an index that does not lie on the boundary of the index domain I_J, has eight neighbors. A boundary index has either five neighbors (when it is not a corner point of the domain) or three neighbors (when it is a corner point).

We consider disjoint dyadic partitions $\{I_\ell^j : \ell = 0, \ldots, 2^j - 1\}$ of I_J for $j = 0, \ldots, 2J$, where $\#I_\ell^j = 2^{2J-j}$, $I_\ell^j \cap I_r^j = \emptyset$ for $\ell \neq r$, and $\bigcup_{\ell=0}^{2^j-1} I_\ell^j = I_J$ for all $j = 0, \ldots, 2J$. These partitions are adaptively chosen during the performance of the EPWT. For fixed j, we say that two subsets I_ℓ^j and I_r^j, with $\ell \neq r$, are *neighbors*,

$$I_r^j \in N(I_\ell^j),$$

iff there exists an index $n \in I_\ell^j$ and an index $m \in I_r^j$ such that $m \in N(n)$.

Now the EPWT algorithm is performed as follows. At the first level, where $j = 2J$, the EPWT is applied along a path vector p^{2J} (of length 2^{2J}) which consists of all indices from I_J in a specific order: the vector p^{2J} is built by concatenating K path vectors p_i^{2J}, $i = 1, \ldots, K$, such that each p_i^{2J} is a *connected vector* containing all grid points from Γ_i^J. We say that a path vector p_i^{2J} of length $\#\Gamma_i^J$ is *connected*, iff for $\ell = 0, \ldots, \lfloor \#\Gamma_i^J/2 - 1 \rfloor$ the two neighboring indices 2ℓ and $2\ell + 1$ are *connected*, i.e., $p_i^{2J}(2\ell + 1) \in N(p_i^{2J}(2\ell))$. Note that by this construction, the complete path vector p^{2J} has only a finite number of at most $K - 1$ *interruptions*, i.e., there are at most $K - 1$ indices 2ℓ, where the condition $p^{2J}(2\ell + 1) \in N(p^{2J}(2\ell))$ is not satisfied.

Note that the components of the path vector p^{2J} lie in I_J, and so p^{2J} contains 2d integer entries, whereas the path vectors p^j, with $j \leq 2J - 1$, contain 1d integer entries. This is in contrast to the notation in [19].

We regard the univariate function

$$\tilde{f}^{2J}(t) := \sum_{\ell=0}^{2^{2J}-1} F^{2J}\left(\tfrac{p^{2J}(\ell)}{2^J}\right)\,\phi(2^{2J}t-\ell) = \sum_{\ell=0}^{2^{2J}-1} f^{2J}(p^{2J}(\ell))\,\phi_{2J,\ell}(t) \qquad \text{for } t\in[0,1)$$

along the path vector $p^{2J} = ((p_1^{2J})^T,\ldots,(p_K^{2J})^T)^T$, with $\phi_{2J,\ell} = 2^J\phi(2^{2J}\cdot -\ell)$ in (2.3), and $f^{2J}(p^{2J}(\ell)) := 2^{-J}F^{2J}(2^{-J}p^{2J}(\ell))$. By using (2.2), the estimate

$$|\tilde{f}^{2J}(2^{-2J+1}\ell) - \tilde{f}^{2J}(2^{-2J}(2\ell+1))| = \left|F^{2J}\left(\tfrac{p^{2J}(2\ell)}{2^J}\right) - F^{2J}\left(\tfrac{p^{2J}(2\ell+1)}{2^J}\right)\right| \le C\,2^{(-J+1/2)\alpha} \tag{2.4}$$

holds for $\ell\in\{0,\ldots,2^{2J-1}-1\}$, as long as $p^{2J}(2\ell)$ and $p^{2J}(2\ell+1)$ are connected and contained in the same index set Γ_i^J for some $1\le i\le K$. Now we apply a one-dimensional Haar wavelet transform to the given data set $\{F^{2J}(2^{-J}n)\}_{n\in I_J} = \{\tilde{f}^{2J}(2^{-2J}\ell)\}_{\ell=0}^{2^{2J}-1}$ along the path vector p^{2J}, and so we obtain the scaling and wavelet coefficients

$$\begin{aligned}
f^{2J-1}(\ell) := \langle \tilde{f}^{2J},\phi_{2J-1,\ell}\rangle &= 2^{-J-1/2}\Big(F^{2J}\left(\tfrac{p^{2J}(2\ell)}{2^J}\right) + F^{2J}\left(\tfrac{p^{2J}(2\ell+1)}{2^J}\right)\Big) &(2.5)\\
&= 2^{-1/2}\Big(f^{2J}(p^{2J}(2\ell)) + f^{2J}(p^{2J}(2\ell+1))\Big),\\
g^{2J-1}(\ell) := \langle \tilde{f}^{2J},\psi_{2J-1,\ell}\rangle &= 2^{-J-1/2}\Big(F^{2J}\left(\tfrac{p^{2J}(2\ell)}{2^J}\right) - F^{2J}\left(\tfrac{p^{2J}(2\ell+1)}{2^J}\right)\Big)\\
&= 2^{-1/2}\Big(f^{2J}(p^{2J}(2\ell)) - f^{2J}(p^{2J}(2\ell+1))\Big),
\end{aligned}$$

for $\ell = 0,\ldots,2^{2J-1}-1$, where we used the identities

$$\begin{aligned}
\langle \phi(2^{2J}\cdot -r),\,\phi_{2J-1,\ell}\rangle &= 2^{-J}\langle \phi_{2J,r},\,2^{-1/2}(\phi_{2J,2\ell}+\phi_{2J,2\ell+1})\rangle\\
&= 2^{-J-1/2}(\delta_{r,2\ell}+\delta_{r,2\ell+1})
\end{aligned}$$

and

$$\langle \phi(2^{2J}\cdot -r),\,\psi_{2J-1,\ell}\rangle = 2^{-J-1/2}(\delta_{r,2\ell}-\delta_{r,2\ell+1})$$

with $\delta_{r,\ell}$ denoting the usual Kronecker symbol. For the wavelet coefficients, (2.4) yields the estimate

$$|g^{2J-1}(\ell)| \le \tfrac{1}{2}\,C\,2^{(-J+1/2)(\alpha+1)},$$

if $p^{2J}(2\ell)$ and $p^{2J}(2\ell+1)$ are connected and contained in the same index set Γ_i^J for some i.

The path vector p^{2J} determines a partition of I_J into index sets

$$I_\ell^{2J-1} := \{p^{2J}(2\ell),\,p^{2J}(2\ell+1)\} \qquad \text{for } \ell=0,\ldots,2^{2J-1}-1.$$

Now we consider the "low-pass" image

$$F^{2J-1}(x) := \sum_{\ell=0}^{2^{2J-1}-1} f^{2J-1}(\ell)\,h_{2J-1,\ell}(x) \qquad \text{for } x\in[0,1)^2 \tag{2.6}$$

with the L^2-normalized characteristic functions

$$h_{2J-1,\ell}(x) := 2^{J-1/2}\Big(\chi_{[0,1)^2}(2^Jx - p^{2J}(2\ell)) + \chi_{[0,1)^2}(2^Jx-p^{2J}(2\ell+1))\Big)$$

corresponding to the index sets I_ℓ^{2J-1}, for $\ell = 0, \ldots, 2^{2J-1} - 1$.

As regards the performance of the EPWT at the second level, where $j = 2J - 1$, we first locate a second (almost) connected path vector $p^{2J-1} = (p^{2J-1}(\ell))_{\ell=0}^{2^{2J-1}-1}$ through the index sets I_ℓ^{2J-1}, $\ell = 0, \ldots, 2^{2J-1} - 1$, i.e., the entries of p^{2J-1} are a permutation of the index set $\{0, \ldots, 2^{2J-1} - 1\}$, and we require $I_{p^{2J-1}(2r+1)}^{2J-1} \in N(I_{p^{2J-1}(2r)}^{2J-1})$ for all $r \in \{0, \ldots, 2^{2J-2} - 1\}$ (up to a finite set of indices whose size does not depend on J). Then, we apply the (one level) Haar wavelet transform to

$$\tilde{f}^{2J-1}(t) := \sum_{\ell=0}^{2^{2J-1}-1} f^{2J-1}(p^{2J-1}(\ell))\, \phi_{2J-1,\ell}(t) \qquad \text{for } t \in [0,1)$$

which yields the scaling and wavelet coefficients

$$f^{2J-2}(\ell) := \langle \tilde{f}^{2J-1}, \phi_{2J-2,\ell} \rangle, \quad g^{2J-2}(\ell) := \langle \tilde{f}^{2J-1}, \psi_{2J-2,\ell} \rangle, \quad \ell = 0, \ldots, 2^{2J-2} - 1.$$

Like in the above construction, a corresponding "low pass" image

$$F^{2J-2}(x) := \sum_{\ell=0}^{2^{2J-2}-1} f^{2J-2}(\ell)\, h_{2J-2,\ell}(x) \qquad \text{for } x \in [0,1)^2$$

is obtained, with $h_{2J-2,\ell}$ being the L^2-normalized characteristic functions of the index sets

$$I_\ell^{2J-2} := I_{p^{2J-1}(2\ell)}^{2J-1} \cup I_{p^{2J-1}(2\ell+1)}^{2J-1} \qquad \text{for } \ell = 0, \ldots, 2^{2J-2} - 1,$$

i.e.,

$$h_{2J-2,\ell} := 2^{-1/2}\Big(h_{2J-1,p^{2J-1}(2\ell)} + h_{2J-1,p^{2J-1}(2\ell+1)}\Big).$$

We continue by iteration over the remaining levels $2J - j$, for $j = 2J-2, 2J-3, \ldots, 0$, where at any level $2J - j$ we first locate a path vector $p^{j+1} = (p^{j+1}(\ell))_{\ell=0}^{2^{j+1}-1}$ through the index sets $I_\ell^{j+1} := I_{p^{j+2}(2\ell)}^{j+2} \cup I_{p^{j+2}(2\ell+1)}^{j+2}$, $\ell = 0, \ldots, 2^{j+1} - 1$, before the Haar wavelet transform is applied to

$$\tilde{f}^{j+1}(t) := \sum_{\ell=0}^{2^{j+1}-1} f^{j+1}(p^{j+1}(\ell))\, \phi_{j+1,\ell}(t),$$

yielding

$$f^j(\ell) := \langle \tilde{f}^{j+1}, \phi_{j,\ell} \rangle, \quad \text{and} \quad g^j(\ell) := \langle \tilde{f}^{j+1}, \psi_{j,\ell} \rangle, \qquad \text{for } \ell = 0, \ldots, 2^j - 1.$$

2.2 Adaptive Haar Wavelet Bases

The EPWT algorithm of the previous subsection can be viewed as a tool for adaptive multiresolution analysis. We can further explain this as follows. Let us consider the space V_{2J} of piecewise constant functions

$$V_{2J} = \text{span}\,\{h_{2J,\ell} : \ell = 0, \ldots, 2^{2J} - 1\},$$

where each $h_{2J,\ell}$ is the L^2-normalized characteristic function on one square with edge length 2^{-J}, i.e.,

$$h_{2J,\ell} := 2^J \chi_{[0,1)^2}(2^J \cdot - p^{2J}(\ell)) \qquad \text{for } \ell = 0, \ldots, 2^{2J} - 1.$$

Obviously, the function set $\{h_{2J,\ell} : \ell = 0, \ldots, 2^{2J} - 1\}$ forms an orthonormal basis of V_{2J}, and the function F^{2J}, as defined in the last subsection, can be written as

$$F^{2J}(x) = \sum_{\ell=0}^{2^{2J}-1} 2^{-J} F(2^{-J} p^{2J}(\ell))\, h_{2J,\ell}(x) = \sum_{\ell=0}^{2^{2J}-1} f^{2J}(p^{2J}(\ell))\, h_{2J,\ell}(x) \quad \text{for } x \in [0,1)^2$$

with $f^{2J}(p^{2J}(\ell)) := 2^{-J}\, F(2^{-J} p^{2J}(\ell))$. By applying the first level of the EPWT with Haar filters along the path vector $(p^{2J}(\ell))_{\ell=0}^{2^{2J}-1}$, we determine the coarser function spaces

$$\begin{aligned} V_{2J-1}(F) &:= \operatorname{span}\left\{ h_{2J-1,\ell} := 2^{-1/2}\left(h_{2J,2\ell} + h_{2J,2\ell+1}\right) \text{ for } \ell = 0, \ldots, 2^{2J-1} - 1 \right\}, \\ W_{2J-1}(F) &:= \operatorname{span}\left\{ \Psi_{2J-1,\ell} := 2^{-1/2}\left(h_{2J,2\ell} - h_{2J,2\ell+1}\right) \text{ for } \ell = 0, \ldots, 2^{2J-1} - 1 \right\}. \end{aligned}$$

Note that the space $V_{2J-1}(F)$ (resp. $W_{2J-1}(F)$) is generated by piecewise constant functions whose support usually consists of two neighboring squares of edge length 2^{-J}. We have $V_{2J-1}(F) \subset V_{2J}$ and $W_{2J-1}(F) \subset V_{2J}$. Moreover,

$$V_{2J-1}(F) + W_{2J-1}(F) = V_{2J}, \qquad V_{2J-1}(F) \perp W_{2J-1}(F).$$

Hence, the first step of the EPWT yields an orthonormal decomposition of F^{2J} into $F^{2J-1} \in V_{2J-1}(F)$ and $G^{2J-1} \in W_{2J-1}(F)$, with F^{2J-1} in (2.6), and

$$G^{2J-1}(x) = \sum_{\ell=0}^{2^{2J-1}-1} g^{2J-1}(\ell)\, \Psi_{2J-1,\ell}(x) \qquad \text{for } x \in [0,1)^2.$$

Proceeding further along these lines, we obtain for $j = 2J-1, 2J-2, \ldots, 0$ the adaptive scaling and wavelet spaces from the path vectors $(p^{j+1}(\ell))_{\ell=0}^{2^{j+1}-1}$,

$$\begin{aligned} V_j(F) &:= \operatorname{span}\left\{ h_{j,\ell} := 2^{-1/2}\left(h_{j+1,p^{j+1}(2\ell)} + h_{j+1,p^{j+1}(2\ell+1)}\right) \text{ for } \ell = 0, \ldots, 2^j - 1 \right\}, \\ W_j(F) &:= \operatorname{span}\left\{ \Psi_{j,\ell} := 2^{-1/2}\left(h_{j+1,p^{j+1}(2\ell)} - h_{j+1,p^{j+1}(2\ell+1)}\right) \text{ for } \ell = 0, \ldots, 2^j - 1 \right\}. \end{aligned}$$

The support of the orthogonal basis functions $h_{j,\ell} \in V_j(F)$ (resp. $\Psi_{j,\ell} \in W_j(F)$) usually consists of connected areas generated by 2^{2J-j} squares with edge length 2^{-J}. Obviously, we have $V_j(F) + W_j(F) = V_{j+1}(F)$, $V_j(F) \perp W_j(F)$, and the function set

$$\{h_{0,0}\} \cup \{\Psi_{j,\ell} : j = 0, \ldots, 2J-1,\, \ell = 0, \ldots, 2^j - 1\}$$

forms an orthonormal basis of V_{2J}.

The application of $2J$ levels to a function $F^{2J} \in V_{2J}$ by using the EPWT algorithm finally yields a unique orthonormal decomposition

$$F^{2J} = F^0 + \sum_{j=0}^{2J-1} G^j$$

with $F^0 = f^0(0)h_{0,0} = f^0(0)\,\chi_{[0,1)^2}$, where

$$f^0(0) = \int_{[0,1)^2} F^{2J}(x)\,dx = 2^{-2J}\sum_{\ell=0}^{2^{2J}-1} F^{2J}(p^{2J}(\ell))$$

and

$$G^j(x) := \sum_{\ell=0}^{2^j-1} g^j(\ell)\,\Psi_{j,\ell}(x) \qquad \text{for } x \in [0,1)^2 \qquad \text{for } j = 0,\ldots,2J-1.$$

2.3 Conditions for the Path Vectors

In this subsection, we fix two specific side conditions for the path vectors. The side conditions are required for our error analysis in the following Section 3. The two side conditions are termed (a) *region condition* and (b) *diameter condition*, as stated below.

To introduce the two conditions, let $J > 0$ be an arbitrary fixed integer. Suppose that the path vectors $(p^j(\ell))_{\ell=0}^{2^j-1}$, $j = 1,\ldots,2J$, are iteratively chosen at the levels of the EPWT algorithm. We assume that their corresponding index sets I_ℓ^j, $\ell = 0,\ldots,2^j-1$,

$$\begin{aligned} I_\ell^{2J} &:= \{p^{2J}(\ell)\} && \text{for } \ell = 0,\ldots,2^{2J}-1, \\ I_\ell^j &:= I^{j+1}_{p^{j+1}(2\ell)} \cup I^{j+1}_{p^{j+1}(2\ell+1)} && \text{for } \ell = 0,\ldots,2^j-1,\ 1 \le j < 2J, \end{aligned}$$

are satisfying the following two conditions for every $J > 0$.

(a) Region condition. After the performance of the $(2J-j)$th level of the EPWT, $j \in \{2J-1,\ldots,0\}$, there are at most $C_1 K$ index sets I_ℓ^j, whose components are not completely contained in one region index set Γ_i^J for some $i \in \{1,\ldots,K\}$, and the constant C_1 does not depend on J or j.

(b) Diameter condition. After the performance of the $(2J-j)$th level of the EPWT, $j \in \{2J-1,\ldots,0\}$, almost all index sets I_ℓ^j, $\ell \in \{0,\ldots,2^j-1\}$, whose components are completely contained in one region index set Γ_i^J, for some $i \in \{1,\ldots K\}$, possess an almost optimal "diameter", i.e., there is a constant $D < \infty$, being independent of J and j, such that

$$\operatorname{diam} I_\ell^j = \max_{k_1,k_2 \in I_\ell^j} \|k_1 - k_2\|_2 \le D\,2^{J-j/2}. \tag{2.7}$$

Moreover, the number of index sets with $I_\ell^j \subset \Gamma_i^J$, for some i, that do not satisfy this optimal diameter condition, is bounded by a constant $C_2 < \infty$ which does not depend on J or j.

For notational simplicity, we denote by Λ^j the set of all indices $\ell \in \{0,\ldots,2^j-1\}$ for which the index set I_ℓ^j is contained in Γ_i^J, for some i, and which satisfies the diameter condition (2.7). By assumptions (a) and (b), at each level $2J-j$, the number of indices ℓ being not contained in Λ^j is bounded by $C_1K + C_2$.

Remark. Since each index set I_ℓ^j has by definition 2^{2J-j} elements, the diameter condition can for instance be satisfied for even j, if the indices in I_ℓ^j are arranged in a square of length $2^{J-j/2}$. In this case, condition (2.7) holds with $D = \sqrt{2}$. □

2.4 Example

For the purpose of illustration, we consider one specific example for an image of size 16×16. This numerical example demonstrates the efficacy of the EPWT algorithm. Furthermore, it helps explain the region condition and the diameter condition of the previous subsection. To this end, we consider a piecewise Hölder continuous function F and its interpolation F^{2J} for $J = 4$, see Figure 1(a). The image F^{2J} consists of three regions, and we assume that condition (2.2) (with a suitable constant C) is satisfied in each of these regions.

We aim to apply the EPWT algorithm such that the region condition (a) and the diameter condition (b) are satisfied with small constants C_1, D, and C_2. In a first step, we determine a path p^8 of length 256 that is built by concatenating three connected pathways p_1^8 (background), p_2^8 (approximation of circle) and p_3^8 (approximation of sloping bar), see Figure 1(b). The first indices of the pathways p_1^8, p_2^8 and p_3^8 are indicated by small circles. In this example, the vectors p_ν^8, for $\nu = 1, 2, 3$, are even completely connected, i.e., we have $p_\nu^8(\ell + 1) \in N(p_\nu^8(\ell))$, for $\nu = 1, 2, 3$ and for $\ell = 0, \ldots, \#\Gamma_\nu^4 - 2$. The path vector p^8 determines a low-pass image F^7, where we have also indicated the index sets I_ℓ^7, $\ell = 0, \ldots, 127$ that are determined by p^8, see Figure 1(c).

There is only one index set, which does not satisfy the region condition, see the last index $(0, 15)$ of p_1^8 in the right upper corner and the first index $(2, 4)$ of p_2^8 in Figure 1(c). Hence, we have $C_1 = 1/3$. Since there are index sets I_ℓ^7 of the form $\{m,\, n\}$ with $\|m - n\|_2 = \sqrt{2}$, the diameter condition (2.7) is satisfied with $D = 1$, and we have $C_2 = 0$.

At the second level of the EPWT, a suitable path vector p^7 determines a low-pass image F^6, see Figure 1(d), with indicated index sets of size 4. There are only two index sets, which do not satisfy the region condition, i.e., $C_1 = 2/3$. Furthermore, there are two index sets at the bottom of the image, which satisfy the diameter condition (2.7) only with the constant $D = 3/2$, wheras all other index sets satisfy (2.7) with $D \leq \sqrt{5}/2$.

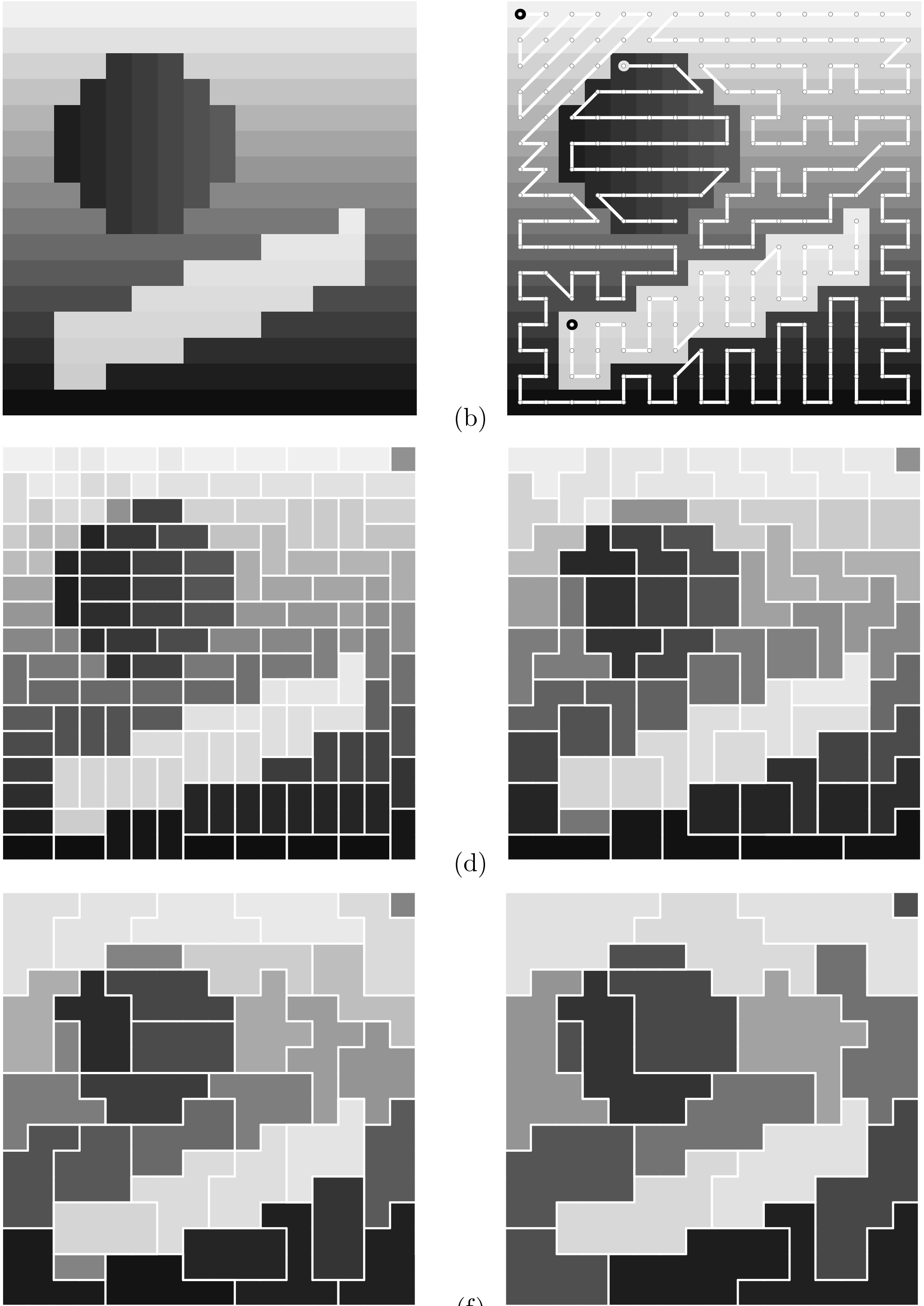
At the third level of the EPWT, a path vector p^6 determines the low-pass image F^5, see Figure 1(e), where we have also indicated the index sets of size 8. An appropriate concatenation of those two index sets of size 4 that did not meet the region condition in F^6 now leads to only one exception of the region condition, i.e., $C_1 = 1/3$. We observe that all other index sets in Figure 1(e) satisfy the diameter condition (2.7) with $D = \sqrt{18}/\sqrt{8} = 3/2$.

At the fourth level of the EPWT, we obtain the low-pass image F^4 in Figure 1(f) with only one index set violating the region condition. Five index sets satisfy the diameter condition with $D = \sqrt{45}/4 \approx 1.67705$. The remaining index sets satisfy (2.7) with $D \leq \sqrt{40}/4 \approx 1.58114$. In this case, (level-independent) constants are $C_1 = 2/3$, $D = \sqrt{45}/4$, and $C_2 = 0$.

3 Approximation Properties of the EPWT Algorithm

Recall that for given integer $J > 0$, the function F^{2J} is assumed to be the piecewise constant approximation of the image F satisfying (2.2). In this section, we shall prove the optimal N-term approximation to F by a suitably chosen EPWT, where the path vectors are required to satisfy the region condition and the diameter condition of Subsection 2.3.

Let us first prove suitable estimates for the scaling and the wavelet coefficients.



(a) (b) (c) (d) (e) (f)

Figure 1: Application of the EPWT algorithm to a 16×16 image. (a) interpolated image F^8, (b) connected path p^8 for F^8, (c) low-pass image F^7 with index sets of size 2, (d) low-pass image F^6 with index sets of size 4, (e) low-pass image F^5 with index sets of size 8, (f) low-pass image F^4 with index sets of size 16.

Theorem 3.1 *Let F^{2J} be an image in V^{2J} satisfying (2.2) for each index set Γ_i^J, $i = 1,\ldots,K$, as determined in Subsection 2.1. For the application of the EPWT to F^{2J} (according to Section 2), we assume that the path vectors $(p^{j+1}(\ell))_{\ell=0}^{2^{j+1}-1}$, $j = 2J-1,\ldots,0$, satisfy the region condition (a) and the diameter condition (b) of Subsection 2.3. Let $f^{2J}(p^{2J}(\ell)) := 2^{-J}F^{2J}(2^{-J}p^{2J}(\ell))$, $\ell = 0,\ldots,2^{2J}-1$, and let $f^j(\ell) = \langle \tilde{f}^{j+1}, \phi_{j,\ell}\rangle$, $j = 2J-1, 2J-2,\ldots,0$, $\ell = 0,\ldots,2^j-1$ be the scaling coefficients that are obtained by the EPWT. Then, for all $j = 2J,\ldots,0$ and $\ell \in \Lambda^{j-1}$, the estimate*

$$|f^j(p^j(2\ell)) - f^j(p^j(2\ell+1))| \le 2^{\alpha/2}\, C\, D^\alpha\, 2^{-j(\alpha+1)/2} \tag{3.1}$$

holds, where $D > 1$ is the constant of the diameter condition (2.7), and where C and α are the Hölder constant and the Hölder exponent in (2.2). Furthermore, for all $\ell \in \{0,\ldots,2^{j-1}-1\} \setminus \Lambda^{j-1}$, we find the estimate

$$|f^j(p^j(2\ell)) - f^j(p^j(2\ell+1))| \le C'\, 2^{-j/2} \tag{3.2}$$

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

Proof. For $j = 2J$, estimate (3.1) follows directly from (2.4). By using the representation $\phi_{2J-1,\ell} = 2^{-1/2}(\phi_{2J,2\ell} + \phi_{2J,2\ell+1})$ we find

$$f^{2J-1}(\ell) = 2^{-J-1/2}\left(F^{2J}\left(\tfrac{p^{2J}(2\ell)}{2^J}\right) + F^{2J}\left(\tfrac{p^{2J}(2\ell+1)}{2^J}\right)\right) = 2^{-J-1/2} \sum_{n \in I_\ell^{2J-1}} F^{2J}\left(\tfrac{n}{2^J}\right),$$

see (2.5). Hence,

$$\begin{aligned}
&|f^{2J-1}(p^{2J-1}(2\ell)) - f^{2J-1}(p^{2J-1}(2\ell+1))| \\
&= 2^{-J-1/2}\left| \sum_{n \in I^{2J-1}_{p^{2J-1}(2\ell)}} F^{2J}\left(\tfrac{n}{2^J}\right) - \sum_{n \in I^{2J-1}_{p^{2J-1}(2\ell+1)}} F^{2J}\left(\tfrac{n}{2^J}\right)\right| \\
&\le 2^{-J-1/2}\, 2C\left(\tfrac{2D}{2^J}\right)^\alpha = 2^{\alpha/2}\, C\, D^\alpha\, 2^{-(J-1/2)(\alpha+1)},
\end{aligned}$$

follows by $I_\ell^{2J-2} = I^{2J-1}_{p^{2J-1}(2\ell)} \cup I^{2J-1}_{p^{2J-1}(2\ell+1)}$ and (2.2), where we note that the sets $I^{2J-1}_{p^{2J-1}(2\ell)}$ and $I^{2J-1}_{p^{2J-1}(2\ell+1)}$ contain only two indices and, moreover, the diameter condition

$$\operatorname{diam} I_\ell^{2J-2} \le 2D$$

holds. Likewise, for general $j \in \{1,\ldots,2J-1\}$ we observe that $f^j(\ell)$, as obtained by the application of the $(2J-j)$th level of the EPWT algorithm, can be viewed as a weighted average of function values $F^{2J}(2^{-J}n)$ with $n \in I_\ell^j$, i.e.,

$$f^j(\ell) = 2^{-J}\, \tfrac{1}{2^{J-j/2}} \sum_{n \in I_\ell^j} F^{2J}\left(\tfrac{n}{2^J}\right),$$

where $\# I_\ell^j = 2^{2J-j}$. Hence, by using the diameter property for $\ell \in \Lambda^{j-1}$, we obtain

$$\begin{aligned}
|f^j(p^j(2\ell)) - f^j(p^j(2\ell+1))| &= 2^{-2J+j/2}\left| \sum_{n \in I^j_{p^j(2\ell)}} F^{2J}\left(\tfrac{n}{2^J}\right) - \sum_{n \in I^j_{p^j(2\ell+1)}} F^{2J}\left(\tfrac{n}{2^J}\right)\right| \\
&\le 2^{-2J+j/2}\, 2^{2J-j}\, C\left(\frac{D\, 2^{J-(j-1)/2}}{2^J}\right)^\alpha \\
&= 2^{\alpha/2}\, C\, D^\alpha\, 2^{-j(\alpha+1)/2}.
\end{aligned}$$

Finally, since F (resp. F^{2J}) is bounded, we obtain

$$|f^j(p^j(2\ell)) - f^j(p^j(2\ell+1))| \leq C'\, 2^{-j/2}$$

for all indices $\ell \in \{0, \ldots, 2^{j-1}-1\} \setminus \Lambda^{j-1}$, with some constant C' being independent from j. Note that the last estimate follows from the previous one by letting $\alpha = 0$. □

We are now in a position to estimate the wavelet coefficients obtained by the EPWT.

Theorem 3.2 *For $j = 2J-1, \ldots, 0$, let $g^j(\ell) = \langle \tilde{f}^{j+1}, \psi_{j,\ell}\rangle$, $\ell = 0, \ldots, 2^j - 1$, denote the wavelet coefficients that are obtained by applying the EPWT algorithm to $F^{2J} \in V_{2J}$ (according to Section 2), where we assume that F^{2J} satisfies (2.2). 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 2.3. Then, for all $j = 2J-1, \ldots, 0$ and $\ell \in \Lambda^j$, the estimate*

$$|g^j(\ell)| \leq \tfrac{1}{2}\, C\, D^\alpha\, 2^{-j(\alpha+1)/2} \tag{3.3}$$

holds, where $D > 1$ is the constant of the diameter condition (2.7), and where C and α are the Hölder constant and the Hölder exponent in (2.2). Furthermore, for all $\ell \in \{0, \ldots, 2^j - 1\} \setminus \Lambda^j$, we find the estimate

$$|g^j(\ell))| \leq \tfrac{1}{2}\, C'\, 2^{-j/2} \tag{3.4}$$

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

Proof. The proof follows from Theorem 3.1, with observing that the one-dimensional Haar wavelet satisfies $\psi_{j,\ell} = 2^{-1/2}\,(\phi_{j+1,2\ell} - \phi_{j+1,2\ell+1})$, and by using $\langle \phi_{j+1,r}, \phi_{j+1,\ell}\rangle = \delta_{r,\ell}$. By (3.1), we obtain

$$\begin{aligned}
|g^j(\ell)| &= |\langle \tilde{f}^{j+1}, \psi_{j,\ell}\rangle| \\
&= \Big| \sum_{r=0}^{2^{j+1}-1} f^{j+1}(p^{j+1}(r))\, \langle \phi_{j+1,r},\, 2^{-1/2}\,(\phi_{j+1,2\ell} - \phi_{j+1,2\ell+1})\rangle \Big| \\
&= 2^{-1/2}\, |f^{j+1}(p^{j+1}(2\ell)) - f^{j+1}(p^{j+1}(2\ell+1))| \\
&\leq 2^{-1/2}\, 2^{\alpha/2}\, C\, D^\alpha\, 2^{-(j+1)(\alpha+1)/2} = \tfrac{1}{2}\, C\, D^\alpha\, 2^{-j(\alpha+1)/2}.
\end{aligned}$$

Likewise, for all $\ell \in \{0, \ldots, 2^j - 1\} \setminus \Lambda^j$

$$|g^j(\ell)| \leq 2^{-1/2}\, C'\, 2^{-(j+1)/2} = C'\, 2^{-(j+2)/2}$$

follows from (3.2). □

Observe that the complete image F^{2J} is now covered by the vector of wavelet coefficients (as generated by the EPWT)

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

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

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

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

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

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

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

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.

Using the orthogonal decomposition of F^{2J} of Subsection 2.2, the L^2-error can be represented as

$$\epsilon_N = \|F^{2J} - F_N^{2J}\|_2^2 = \sum_{(j,\ell)\notin S_N^{2J}} |g^j(\ell)|^2, \tag{3.5}$$

where F_N^{2J} is the approximation of F^{2J} that is reconstructed from the N wavelet coefficients $g^j(\ell)$, $(j, \ell) \in S_N^{2J}$ with largest absolute value.

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

Theorem 3.3 *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*

$$\epsilon_N = \|F^{2J} - F_N^{2J}\|_2^2 \leq \tilde{C}\, N^{-\alpha} \tag{3.6}$$

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

Proof. We organize the proof into two parts.

1. Let the sequence of all wavelet coefficients $g^j(\ell)$, $j = 0, \ldots, 2J-1$, $\ell = 0, \ldots, 2^j - 1$, and $g^{-1} = g^{-1}(0)$ be sorted in decreasing order, such that we obtain the new sequence $(g_\mu)_{\mu=0}^{2^{2J}-1}$ with $|g_\mu| \geq |g_{\mu+1}|$ for $\mu = 0, \ldots, 2^{2J} - 2$.

We first show that $\sum_{\mu=0}^{2^{2J}-1} |g_\mu|^p$ with $\frac{1}{2} < \frac{1}{p} < \frac{\alpha+1}{2}$ is bounded independently from the choice of the integer $J > 1$. For that purpose, we use the estimates in Theorem 3.2, where we distinguish between *type I wavelet coefficients* $g^j(\ell)$ satisfying the estimate (3.3) and *type II wavelet coefficients* satisfying only the estimate (3.4). From the region condition (a) and the diameter condition (b) on the path vectors p^j, it follows that there are at most $C_1K + C_2$ wavelet coefficients of type II in each level j and the sum of these type II coefficients is bounded by

$$\begin{aligned}
\sum_{g_\mu \text{ of type II}} |g_\mu|^p &\leq |g^{-1}(0)|^p + (C_1K + C_2) \sum_{j=0}^{2J-1} (\frac{1}{2}C' 2^{-j/2})^p \\
&= |g^{-1}(0)|^p + (C_1K + C_2)(C'/2)^p \sum_{j=0}^{2J-1} 2^{-jp/2} \\
&\leq |g^{-1}(0)|^p + (C_1K + C_2)(C'/2)^p (1 - 2^{-p/2})^{-1}
\end{aligned}$$

for all $p > 0$, independently from J.

For the type I coefficients we obtain the bound

$$\begin{aligned}\sum_{g_\mu \text{ of type I}} |g_\mu|^p &\le \sum_{j=0}^{2J-1} 2^j \, |(C\, D^\alpha/2)\, 2^{-j(\alpha+1)/2}|^p \\ &= (C\, D^\alpha/2)^p \sum_{j=0}^{2J-1} 2^{-j(p(\alpha+1)/2-1)}\end{aligned}$$

by using (3.3). This expansion is finite (independent from J), if $p(\alpha+1)/2 > 1$, i.e., if $\frac{1}{p} < \frac{\alpha+1}{2}$.

2. We now estimate the error ϵ_N as follows. For $k \le 2^{2J-1}$, we consider the partial sums

$$S_k^J := \sum_{\mu=k}^{2k-1} |g_\mu|^p.$$

Since $\sum_{\mu=0}^{2^{2J}-1} |g_\mu|^p < \infty$ for arbitrary $J > 1$, it follows that S_k^J is bounded (independently from J). By $|g_{2k}| \le |g_m|$ for $k \le m \le 2k-1$, we now find

$$S_k^J \ge k\, |g_{2k}|^p \quad \text{and} \quad S_k^J \ge k\, |g_{2k-1}|^p$$

for $k < 2^{2J-1}$ and so $|g_{2k}| \le |g_{2k-1}| \le \left(\frac{S_k^J}{k}\right)^{1/p}$, i.e.,

$$|g_k| \le \left(\frac{S_{\lceil k/2\rceil}^J}{k/2}\right)^{1/p} = \left(\frac{2}{k}\, S_{\lceil k/2\rceil}^J\right)^{1/p}.$$

For arbitrary $J > 1$ we obtain

$$\begin{aligned}\sum_{\mu=N+1}^{2^{2J}-1} |g_\mu|^2 &\le \left(\sup_{\mu>N} |S_{\lceil \mu/2\rceil}^J|^{2/p}\right) \sum_{\mu=N+1}^{\infty} \left(\frac{2}{\mu}\right)^{2/p} \\ &= 2^{2/p}\left(\sup_{\mu>N} |S_{\lceil \mu/2\rceil}^J|^{2/p}\right) \sum_{\mu=N+1}^{\infty} \mu^{-2/p}.\end{aligned}$$

Finally, since $\sup_{\mu>N} |S_{\lceil \mu/2\rceil}^J|^{2/p}$ is bounded for each $J > 1$, the assertion of the theorem follows for $p \to 2/(\alpha+1)$ by using

$$\sum_{\mu=N+1}^{\infty} \mu^{-2/p} \le \int_N^\infty x^{-2/p}\, dx = \frac{1}{(2/p-1)} N^{1-2/p}.$$

$\square$

Let us conclude by stating the following corollary.

Corollary 3.4 *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 $J \ge J(\epsilon)$ the N-term estimate*

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

holds, where $\tilde{C}$ is the constant in (3.6).

Proof. For given $J > 0$ and $n = (n_1, n_2) \in I_J$ let

$$A_n^J := [2^{-J}n_1, 2^{-J}(n_1+1)) \times [2^{-J}n_2, 2^{-J}(n_2+1)).$$

Then,

$$\|F - F^{2J}\|_{L^2}^2 = \sum_{n \in I_J} \int_{A_n^J} |F(x) - F(2^{-J}n)|^2 \, dx.$$

We consider two types of indices $n \in I_J$. If $A_n^J \cap \partial\Omega_i = \emptyset$ for all $i \in \{1, \ldots, K\}$, we say that n belongs to the subset I_J' of I_J. Otherwise, it belongs to I_J''. With assuming finite length for $\cup_{i=1}^K \partial\Omega_i$, there exists a constant L being independent of J, such that $\#I_J'' \leq L\, 2^J$ for all J. For the indices in I_J' we can use the Hölder condition (2.1) and find

$$\sum_{n \in I_J'} \int_{A_n^J} |F(x) - F(2^{-J}n)|^2 \, dx \leq \sum_{n \in I_J'} 2^{-2J} (C 2^{(-J+1/2)\alpha})^2 \leq C^2\, 2^{(-2J+1)\alpha}$$

with using $\#I_J' \leq \#I_J = 2^{2J}$. Since the image F is bounded, i.e., $|F(x)| < C'$ for some $C' > 0$, we also have

$$\sum_{n \in I_J''} \int_{A_n^J} |F(x) - F(2^{-J}n)|^2 \, dx < \sum_{n \in I_J''} 2^{-2J} (2C')^2 < 2^{-J+2}\, L\, C'.$$

Hence, $\|F - F^{2J}\|_{L^2}^2 < C^2 2^{(-2J+1)\alpha} + 2^{-J+2} L C'$ holds. Moreover, for any $\epsilon > 0$ we can find an integer $J(\epsilon)$, such that $\|F - F^{2J}\|_{L^2}^2 < \epsilon$ holds for all $J \geq J(\epsilon)$. This in combination with (3.6) concludes our proof. $\square$

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, W. Dahmen, I. Daubechies, and R.A. DeVore, Harmonic analysis of the space BV, *Rev. Mat. Iberoam.* **19**(1) (2003), 235–263.

[5] A. Cohen, R.A. DeVore, P. Petrushev, and H. Xu, Nonlinear approximation and the space $BV(\mathbb{R}^2)$, *Amer. J. Math.* **121** (1999), 587–628.

[6] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.

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

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

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

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

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

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

[13] 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.

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

[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] 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.

[20] G. Plonka and D. Roşca, Easy path wavelet transform on triangulations of the sphere, preprint, 2008.

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

[22] 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.

[23] 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.

[24] 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.

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.