Adapted Multiresolution Analysis and Wavelets

Bl. Sendov

Central Laboratory for Parallel Processing
Bulgarian Academy of Sciences
"Acad. G. Bonchev" street, Block 25A, 1113 Sofia, BULGARIA.
E-mail bsendov@argo.bas.bg

Abstract

There are many generalizations of the multiresolution analysis of Mallat [7] and Meyer [8] to accommodate different applications. A typical example is the general setting of Wim Sweldens [12] for the second generation wavelets. See also [4] and other papers cited in [12]. The main purpose of these generalizations is to adapt the instrument for approximation to the objects (which have) to be approximated. Usually the adaptation is addressedto a class of functions.

In this paper we consider adaptation to a particular function. Such an adaptation is useful for compression of signals and images. It is evident, that if a custom-designed instrument is used for approximation of a particular function, one has to provide additional information for the instrument itself, which reduces the compression factor. Here we have an interplay between a better approximation and an increase of the information to be compressed.

The idea to adapt the instrument for approximation to a particular function is not new. For example, the use of wavelet packages and wavelet libraries [8] for signal and image compression is in fact a discrete adaptation to a particular function. The adaptation considered in this paper is not discrete, but in respect to continuous parameters.

Key Words and Phrases : Adaptive wavelets, Adaptive multiresolution analysis.

1 Introduction

In the Approximation theory it is natural to choose the instrument for approximation according to the objects to be approximated. This principle was theoretically developed in the Kolmogorov's theory of widths [6]. Let X be a linear functional space and $A \subset X$ be a set of functions with given properties. For every natural n A. N. Kolmogorov defined the optimal n-dimensional linear subspace $A_n \subset X$ for linear approximation of the elements of A. In other words, if we have to use a linear method for approximation of functions from A, the best instruments are the elements of A_n.

In the classical approximation theory, the objects for approximation are sets A of functions characterized by some properties of smoothness. A proper instrument for approximation has to satisfy the same property of smoothness and to reflect some peculiarities of the functions belonging to A.

The splines and the wavelets, which occupy the minds of the approximation theorists in the last decades, are also useful instruments for approximation, because they reflect some features of the sets of functions, that are important in applications and have to be approximated.

The use of the approximation theory methods in the signal and image processing requires approximation of functions which are far for smooth and even resemble fractal functions. It is not a surprise, that the *fractal approximation theory*, developed by the group around M. F. Barnsley [2] is so helpful in image compression [3].

Is it reasonable to choose a particular instrument for approximation of a particular function ? In fact, the methods in the fractal approximation theory [3] are based on the so called *Iterated Function Systems* (IFS), which are differently chosen for different target functions for approximation. Also, some methods for signal and image compression, based on the so called *wavelet libraries* and *wavelet packets* [8], use in fact instruments for approximation chosen for a particular function.

In this paper we consider a particular approximation instrument for any particular function from a functional space, based on the adaptation of a multiresolution analysis and the respective wavelets.

2 Adaptive multiresolution analysis

Let

$$V_0 \subset V_1 \subset V_2 \subset \dots \quad \text{with} \quad \overline{\bigcup_{j=0}^{\infty} V_j} = L^2(R), \tag{1}$$

be a ladder of nested subspaces of $L^2(R)$. Fix a natural number N and set

$$\Delta_j(N) = [0, (2N-1)2^{-j}]; \quad N = 1, 2, 3, \dots , \quad j = 0, 1, 2, \dots .$$

Let the ladder (1) satisfy two conditions:

1) For every $j = 0, 1, 2, \dots$ there exists a function $\phi_j(x)$ with a finite support $\Delta_j(N)$, such that $\{\phi_j(x - k2^{-j});\ k \in Z\}$ is an orthonormal basis in V_j.

2) For every $j = 0, 1, 2, \dots$ there exists a **filter** $\mathbf{c}(j) = \{c_n(j)\}_{n=0}^{2N-1}$, such that

$$\phi_{j-1}(x) = \sum_{n=0}^{2N-1} c_n(j)\phi_j(x - n2^{-j}); \quad j = 1, 2, 3, \dots . \tag{2}$$

From 1) and 2) there follows, that the necessary conditions for the filter $\mathbf{c}(j) = \{c_n(j)\}_{n=0}^{2N-1}$ in (2) are

$$\sum_{n=0}^{2N-1} c_n^2 = 1, \quad \sum_{n=0}^{2N-k-1} c_n c_{2k+n} = 0; \quad k = 1, 2, 3, \dots, N-1, \tag{3}$$

the so called *orthonormality conditions.*

Definition 1 *Let*

$$\{\mathbf{c}(j)\}_{j=1}^{\infty} = \{(c_0(j), c_1(j), c_2(j), \dots, c_{2N-1}(j))\}_{j=1}^{\infty} \tag{4}$$

be a sequence of filters satisfying the orthonormality conditions (3)*. Call the sequence of filters* (4) **consistent**, *if there exists a sequence of functions* $\phi_j \in L^2(R)$; $j = 0, 1, 2, \dots$ *which satisfy the scaling equations* (2)*.*

It is obvious that a sequence of filters is consistent if all of them are, for example, equal to a Daubechies filter [5]. The necessary and sufficient conditions for the consistency of a sequence of filters is an open problem.

Definition 2 *Call a ladder* (1) *of nested subspaces of* $L^2(R)$*, which satisfies the conditions* 1) *and* 2)*, an* **adaptive multiresolution analysis** (AMRA) *of* **rank** N *and the functions* ϕ_j **adaptive scaling functions**, *or* **adaptive father wavelets**.

If the filters $\mathbf{c}(j) = \mathbf{c}$, in the *scaling equations* (2), do not depend on j and are, for example, equal to the filters of the well known Daubechies wavelets ${}_N\phi(x)$, see [5], then we get $\phi_j(x) = {}_N\phi(2^j x)$; $j = 0, 1, 2, \dots$. In this case we have a multiresolution analysis of Mallat – Meyer – Daubechies. The simplest case in such a setting, $N = 1$, $c_0(j) = c_1(j) = 1/\sqrt{2}$, is the Haar multiresolution analysis.

The motivation for Definition 2 is the following. Let W_{j-1} be the orthogonal compliment of V_{j-1} in V_j, i. e.,

$$V_j = V_{j-1} \oplus W_{j-1}, \quad V_{j-1} \perp W_{j-1}$$

and

$$\psi_{j-1}(x) = \sum_{n=0}^{2N-1} (-1)^n c_{2N-n-1}(j) \phi_j(x - n2^{-j}), \tag{5}$$

then $\{\psi_{j-1}(x - k2^{-j+1});\ k \in Z\}$ is an orthonormal basis in W_{j-1}.

Call the functions (5) **adaptive wavelets** (adaptive mother wavelets).

From the definition

$$L^2(R) = V_0 \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots .$$

If f_A denotes the projection of a function $f \in L^2(R)$ on the subset $A \subset L^2(R)$, then

$$\|f\|_2^2 = \|f_{V_0}\|_2^2 + \|f_{W_0}\|_2^2 + \|f_{W_1}\|_2^2 + \|f_{W_2}\|_2^2 + \cdots .$$

The basic idea of a MRA is that in the space $V_j = V_{j-1} \oplus W_{j-1}$, most of the "energy" ($\|f_{V_j}\|_2^2 = \|f_{V_{j-1}}\|_2^2 + \|f_{W_{j-1}}\|_2^2$) stays in the "cruder" subspace V_{j-1} and a smaller part of it remains in the "detail" subspace W_{j-1}. It is natural to use the freedom to choose the filter $\mathbf{c}(j)$ to maximize $\|f_{V_{j-1}}\|_2^2$ or, which is the same, to minimize $\|f_{W_{j-}}\|_2^2$.

Fix a function $f \in L^2(R)$ and let the filters $\mathbf{c}(j)$ be chosen for $j = s+1, s+2, s+3, \dots$. Define

$$f_k(j-1) = \int_{-\infty}^{\infty} f(t)\phi_{j-1}(x - k2^{-j+1})\, dt =$$

$$\sum_{n=0}^{2N-1} c_n(j) \int_{-\infty}^{\infty} f(t)\phi_j(x-(2k+n)2^{-j})\, dt = \sum_{n=0}^{2N-1} c_n(j) f_{2k+n}(j)$$

and

$$F_{p,q}(j) = \sum_{k\in Z} f_{2k+p}(j) f_{2k+q}(j).$$

It is obvious that

$$F_{p,q}(j) = F_{q,p}(j) \quad \text{and} \quad F_{p+2,q+2}(j) = F_{p,q}(j).$$

Then

$$\|f_{V_{j-1}}\|_2^2 = \sum_{k\in Z} f_k(j-1)^2 = \sum_{k\in Z} \left(\sum_{n=0}^{2N-1} c_n(j) f_{2k+n}(j) \right)^2,$$

or

$$\|f_{V_{j-1}}\|_2^2 = \sum_{m,n=0}^{2N-1} F_{m,n}(j) c_n(j) c_m(j).$$

To find the optimal filter $\mathbf{c}(j)$ is equivalent to find the point of supremum of a symmetric quadratic form under additional conditions. More precisely, to solve the following

Problem 1 *Let U_N be the subset of the $2N$-dimensional Euclidean space E^{2N}, which consists of the points $\mathbf{u} = (u_0, u_1, u_2, \dots, u_{2N-1})$ satisfying the* **orthonormality** *conditions*

$$\sum_{n=0}^{2N-1} u_n^2 = 1, \qquad \sum_{n=0}^{2N-k-1} u_n u_{2k+n} = 0; \quad k = 1, 2, 3, \dots, N-1.$$

Find a point $\mathbf{u}^*$ for which

$$\sum_{m,n=0}^{2N-1} F_{m,n} u_n^* u_m^* = \sup_{\mathbf{u}\in U_N} \left\{ \sum_{m,n=0}^{2N-1} F_{m,n} u_n u_m \right\},$$

where $\|F_{m,n}\|$ is a $2N \times 2N$ symmetric matrix.

Problem 1 has a solution, as U_N is a compact. But the numerical solution of this problem is not trivial even for $N = 2$.

Observe, that the optimal filter for the j level may be calculated if we know this filter for all higher levels. In the applications of a MRA we use finite number of nested subsets V_j, say s. Then we may choose a filter of a "good" wavelet, say a Daubechies wavelet, and take the filters for the levels $j \geq s$ equal to this filter. Then for $j = s-1, s-2, \ldots, 1, 0$ calculate the optimal filter. This will transform the MRA to a AMRA, partially adapted to the particular function f.

The $2N$-dimensional vector of the filter of the Daubechies wavelets is defined from N orthonormality conditions (2) and N vanishing moments

$$\int_{-\infty}^{\infty} x^k \psi(x)\, dx = \sum_{n=0}^{2N-1} (-1)^n n^k c_{2N-n-1} = 0; \quad k = 0, 1, 2, \ldots, N. \tag{6}$$

The orthonormality conditions (3) are necessary for the mechanism of the MRA. The conditions (6) are not obligatory for the functioning of the MRA. They ensure that all algebraic polynomials of degree $< N$ stay in V_0, so to say, they adapt the MRA to the algebraic polynomials. Our idea is to adapt a given MRA to a particular function, using the freedom to choose the consecutive filters.

If we use adapted wavelets to expand a function with $n = 2^s$ wavelet coefficients, it is necessary to add also the information for the adapted filters, which is of the volume of $O(\log n) = O(s)$.

An AMRA of rank N is defined by the respective sequence $\mathbf{C} = \{\mathbf{c}(j)\}_{j=1}^{\infty}$ of filters.

Definition 3 *An AMRA of rank N and with sequence of filters $\mathbf{C}$ is optimal for a function f if all the filters of this AMRA are optimal for f.*

The set $\mathcal{A}(\mathbf{C})$ of the functions for which the AMRA with sequence of filters $\mathbf{C}$ is optimal is not empty. The scaling function ϕ_0 defined by (2) is such a function. In the general case, the characterization of $\mathcal{A}(\mathbf{C})$ is not clear.

It is easy to see that the Haar multiresolution analysis in $L^2[0,1)$ is optimal for a function f if

$$f(1-x) = f(x) \quad \text{for} \quad x \in [0,1]. \tag{7}$$

Really, if a function satisfies the symmetry condition (7), then $F_{0,0}(j) = F_{1,1}(j)$; $j = 1, 2, 3, \ldots$ in Problem 1, hence the maximum of the quadratic form is achieved for $u_0 = u_1 = 1/\sqrt{2}$.

3 Adaptive multiresolution analysis of rank 1

In this section we consider the simplest case of a multiresolution analysis of rank 1 in $L^2[0,1)$. This is a generalization of the Haar multiresolution analysis.

In this case the consistency problem is easy and the scaling equation is solved explicitly, if the dyadic representation of the real numbers is used.

Every $x \in [0,1)$ is uniquely represented in the form

$$x = x_1 2^{-1} + x_2 2^{-2} + x_3 2^{-3} + \cdots \quad ,$$

where $x_i = 0, 1$, if the finite representation for a dyadic rational x is taken.

Let a sequence

$$\{\mathbf{c}(j)\}_{j=1}^{\infty} = \{(c_0(j), c_1(j))\}_{j=1}^{\infty}; \quad c_0(j)^2 + c_1(j)^2 = 1; \quad j = 1, 2, 3, \ldots \tag{8}$$

of filters of rank 1 be given. Let

$$\lambda_j(0) = \sqrt{2} c_0(j), \quad \lambda_j(1) = \sqrt{2} c_1(j); \quad j = 1, 2, 3, \ldots \tag{9}$$

and define the sequence $\delta = \{\delta_j\}_{j=1}^{\infty}$ in such a way that

$$\lambda_j(0) = \frac{1+\delta_j}{\sqrt{1+\delta_j^2}}, \quad \lambda_j(1) = \frac{1-\delta_j}{\sqrt{1+\delta_j^2}}; \quad j = 1, 2, 3, \ldots \quad . \tag{10}$$

A sequence of filters of rank 1 is defined uniquely by the corresponding δ-sequence.

Definition 4 *Let $\delta = \{\delta_j\}_{j=1}^{\infty}$ be a sequence of real numbers and $\lambda_j(0)$, $\lambda_j(1)$ are defined by* (10). *A function ϕ of the form*

$$\phi(x) = \varphi(\delta; x) = \prod_{k=1}^{\infty} \lambda_k(x_k); \quad x \in [0,1) \tag{11}$$

is called a **starting function** *with sequence δ, see* [9], [10].

From a theorem of Paley and Zigmund [1, p. 53] it follows that a starting function (11) is Lebesgue integrable if

$$\sum_{i=1}^{\infty} \delta_i^2 < \infty. \tag{12}$$

A starting function may have an arbitrary fractal dimension between 1 and 2, see [11].

Set the notations

$$\phi_j(x) = \begin{cases} 2^{j/2} \prod_{i=j+1}^{\infty} \lambda_i(x_i) & \text{for} \quad x \in [0, 2^{-j}) \\ 0 & \text{for} \quad x \notin [0, 2^{-j}) \end{cases} ; \quad j = 0, 1, 2, \ldots \quad . \tag{13}$$

Lemma 1 *The functions* (13) *satisfy the scaling equations* (2).

Proof. From (13) we have for $x \in [0, 2^{-j})$

$$\phi_{j-1}(x) = 2^{(j-1)/2} \prod_{i=j}^{\infty} \lambda_i(x_i) =$$

$$\frac{1}{\sqrt{2}}\lambda_j(0)\chi_{[0,2^{-j-1})}(x)2^{j/2} \prod_{i=j+1}^{\infty} \lambda_i(x_i) + \frac{1}{\sqrt{2}}\lambda_j(1)\chi_{[2^{-j-1},2^{-j})}(x)2^{j/2} \prod_{i=j+1}^{\infty} \lambda_i(x_i) =$$

$$c_0(j)\phi_j(x) + c_1(j)\phi_j(x - 2^{-j}),$$

where $\chi_\Delta(x)$ is the characteristic function of the interval Δ.

Lemma 2 *The functions*

$$\phi_j(x),\ \phi_j(x - 2^{-j}),\ \phi_j(x - 2 \cdot 2^{-j}), \ldots,\ \phi_j(x - (2^j - 1)2^{-j})$$

are orthonormal in $L^2[0,1)$.

Proof. The orthogonality follows from the separation of the supports of these functions. According to the normality, from (13) we have

$$\int_0^1 \varphi_j(t)^2\, dt = \frac{1}{2}\lambda_{j+1}(0)^2 \int_0^{2^{-j-1}} \varphi_{j+1}(t)^2\, dt + \frac{1}{2}\lambda_{j+1}(1)^2 \int_{2^{-j-1}}^{2^{-j}} \varphi_{j+1}(t)^2\, dt$$

$$= \int_0^1 \varphi_{j+1}(t)^2\, dt.$$

On the other hand the product $\prod_{j=1}^{\infty} \lambda_j(0)$ is convergent, hence

$$\lim_{j\to\infty} \prod_{k=j}^{\infty} \lambda_k(0) = 1$$

and

$$\lim_{j\to\infty} \int_0^1 \varphi_j(t)^2\, dt = 2^j \lim_{j\to\infty} \int_0^{2^{-j}} \prod_{k=j+1}^{\infty} \lambda_k(t_j)\, dt = 1. \quad \square$$

From Lemma 1 and Lemma 2 follows

Theorem 1 *A sequence of filters of rank 1* (8) *is consistent if its* δ*-sequence satisfies the condition* (12).

Corollary 1 *Every consistent filter of rank 1 is asymptotically equal to the Haar filter.*

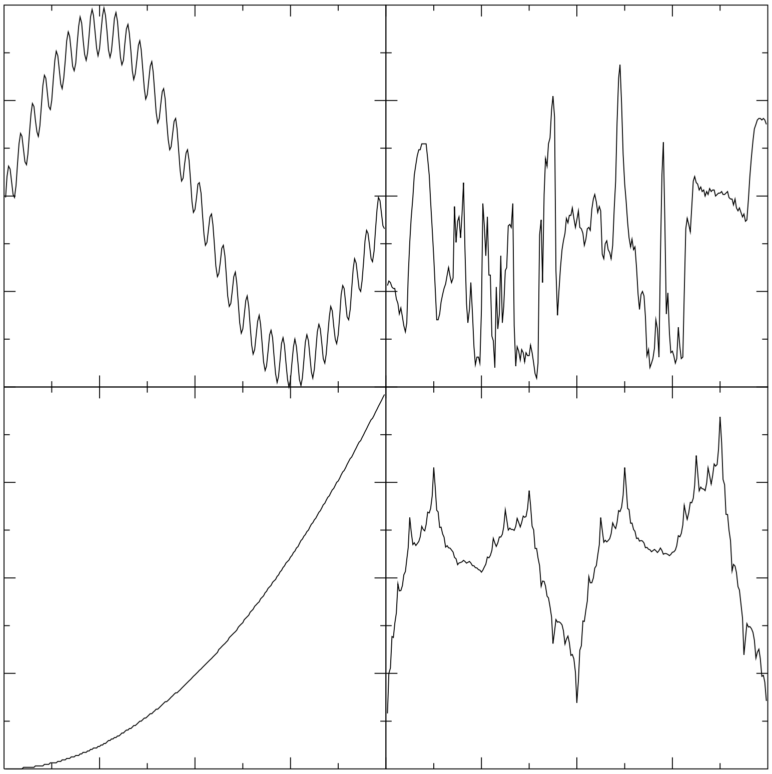


Figure 1: Test functions $\frac{f_2|f_4}{f_1|f_3}$.

3.1 Numerical examples

In this section we illustrate the process of adaptation using a Haar multiresolution analysis on $L^2[0,1)$. Let

$$h_j(x) = \begin{cases} 2^{j/2} & \text{for} \quad x \in [0, 2^{-j}), \\ 0 & \text{for} \quad x \notin [0, 2^{-j}); \end{cases} \qquad j = 0, 1, 2, \dots \quad ,$$

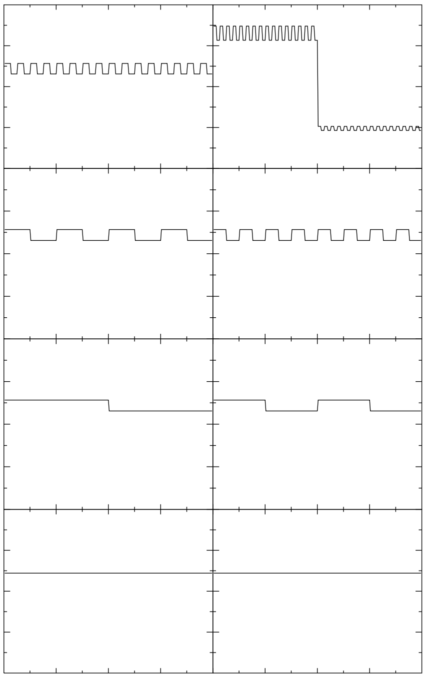
then V_j is spanned over the orthonormal system $\{h_j(x - k2^{-j}\}_{k=0}^{2^j-1}$.

For the illustrations we take a function f from V_s, that means a step function with 2^j regular steps. Then we optimize consecutively the filters for $j = s-1, s-2, \dots, 1$. In our illustrations $s = 8$ and the graphs of the four test functions considered, are depicted in Fig. 1. The test functions are:

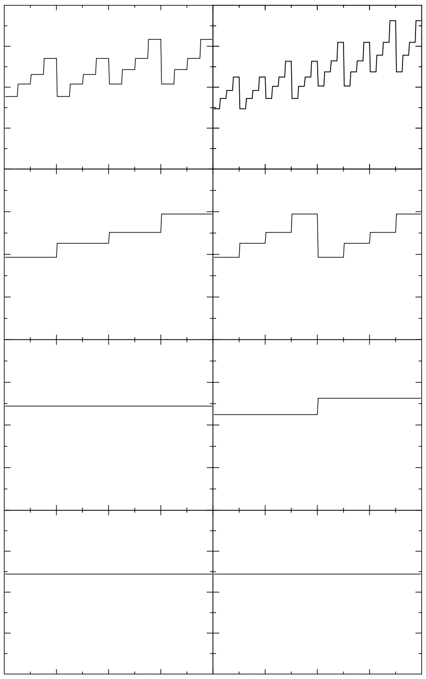
$f_1(x) = x^2, \quad f_2(x) = \sin x + \frac{1}{8}\sin 32x,$

f_3 is a fractal function and f_4 is a row from the photograph of Lenna.

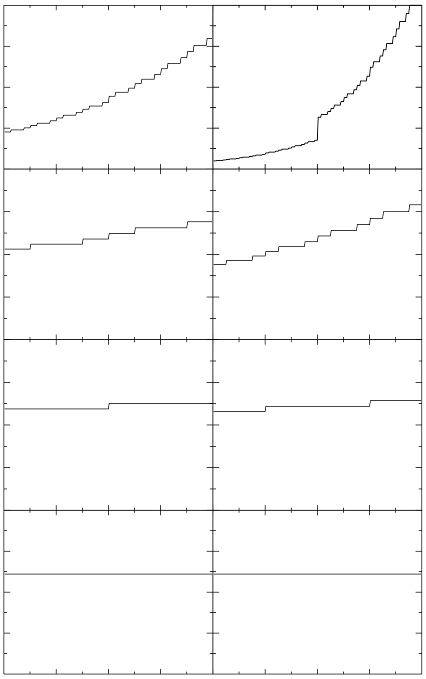
On Fig. 2 the consecutive scaling functions of the AMRA of rank 1, adapted to the respective test functions are depicted. The supports of the consecutive scaling functions are normalized. The scaling functions for the highest indeks, on the bottom, are almost constants, i.e., they are as the Haar scaling function.



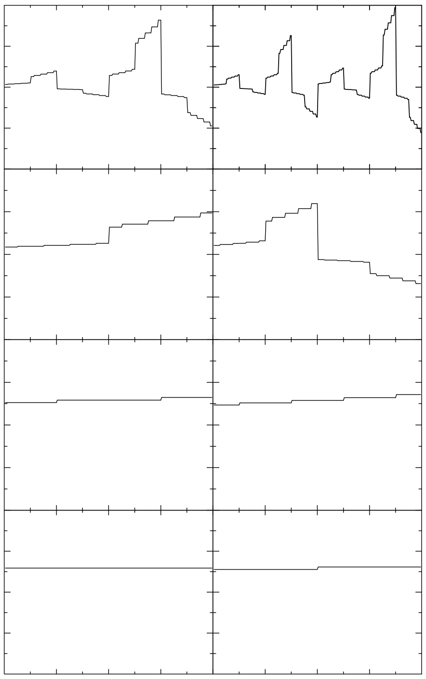
Test function f_2



Test function f_4



Test function f_1



Test function f_3

Figure 2: Rank 1 filter, one parameter adaptation

Function Entropy	Rank 1 Haar	Rank 1 adapted	Rank 2 Daubechies	Rank 2 adapt. 1 par.	Rank 2 adapt. 2 par.
f_1	0.936993	0.239505	1.388035	0.860206	0.203622
f_2	0.914136	0.878660	1.419075	0.556471	0.481989
f_3	4.367164	3.845035	2.926231	2.909769	0.963871
f_4	2.449377	2.361259	2.444133	2.283421	1.996557

Table 1: Entropy of the wavelet coefficients

We calculate the wavelet coefficients of the function f in Haar wavelets and in optimized wavelets. To compare the quality of these wavelet coefficients for compression, we use the

3.2 Entropy Criterion

Let $\{\phi_i\}_{i=0}^m$ be an orthonormal basis in V_s and

$$a_i(\phi; f) = \int_0^1 f(t)\phi_i(t)\ dt, \quad i = 0, 1, 2, \ldots, m$$

and

$$\sum_{i=0}^m a_i(\phi; f)^2 = 1.$$

The entropy criterion, for the adaptation of an orthonormal basis to the function f is to minimize the entropy

$$\epsilon^2(f; \{\phi\}) = -\sum_{i=0}^m |a_i(\phi; f)|^2 \log_2 |a_i(\phi; f)|^2.$$

This criterion is the basis for adaptation of wavelet packets [8].

In Table 1 the entropy of the wavelet coefficients of the test functions for Haar wavelets and the adapted wavelets is given. The advantage of the adapted wavelets is evident.

4 Adaptive multiresolution analysis of rank 2

In this section we consider a multiresolution analysis of rank 2 in $L^2(R)$. This is a generalization of the Daubechies multiresolution analysis with the scaling functions ${}_2\phi(x)$, see [5] and filter

$$c_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad c_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad c_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad c_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}. \tag{14}$$

For $N = 2$ the orthogonality conditions for every filter are two and every filter depends on two parameters. The general form of the filters is the following

$$c_0 = \cos\alpha\cos\beta, \quad c_1 = \cos\alpha\sin\beta, \quad c_2 = \sin\alpha\sin\beta, \quad c_3 = -\sin\alpha\cos\beta,$$

where $\alpha, \beta \in [0, \pi]$ are two parameters.

4.1 Optimization with two parameters

The optimization problem is to find the maximum of a function $\Phi(\alpha, \beta)$ of two variables in the square $[0, \pi] \times [0, \pi]$, or to solve the system

$$\frac{\partial\Phi}{\partial\alpha} = (F_{1,1} - F_{0,0})\sin 2\alpha\cos 2\beta \;-\; 2F_{0,1}\sin 2\alpha\sin 2\beta$$

$$+\,(F_{0,2} - F_{1,3})\cos 2\alpha\sin 2\beta \;-(F_{1,2} + F_{0,3})\cos 2\alpha\cos 2\beta \;+\; (F_{1,2} - F_{0,3})\cos 2\alpha = 0,$$

$$\frac{\partial\Phi}{\partial\beta} = (F_{1,1} - F_{0,0})\cos 2\alpha\sin 2\beta \;-\; 2F_{0,1}\cos 2\alpha\cos 2\beta$$

$$+\,(F_{0,2} - F_{1,3})\sin 2\alpha\cos 2\beta \;+\; (F_{1,2} + F_{0,3})\sin 2\alpha\sin 2\beta = 0.$$

On Fig. 3 the consecutive scaling functions of the AMRA of rank 2, adapted by two parameters to the respective test functions are depicted.

4.2 Optimization with one parameter

If we add the condition for zero moment vanishing to the orthonormality conditions, i. e.,

$$c_0 + c_2 = c_1 + c_3,$$

for every filter, then every filter will depend on one parameter. In this case every filter will have the form

$$c_0 = \frac{1}{2\sqrt{2}}(1 + \sin\alpha + \cos\alpha), \quad c_1 = \frac{1}{2\sqrt{2}}(1 + \sin\alpha - \cos\alpha),$$

$$c_2 = \frac{1}{2\sqrt{2}}(1 - \sin\alpha - \cos\alpha), \quad c_3 = \frac{1}{2\sqrt{2}}(1 - \sin\alpha + \cos\alpha),$$

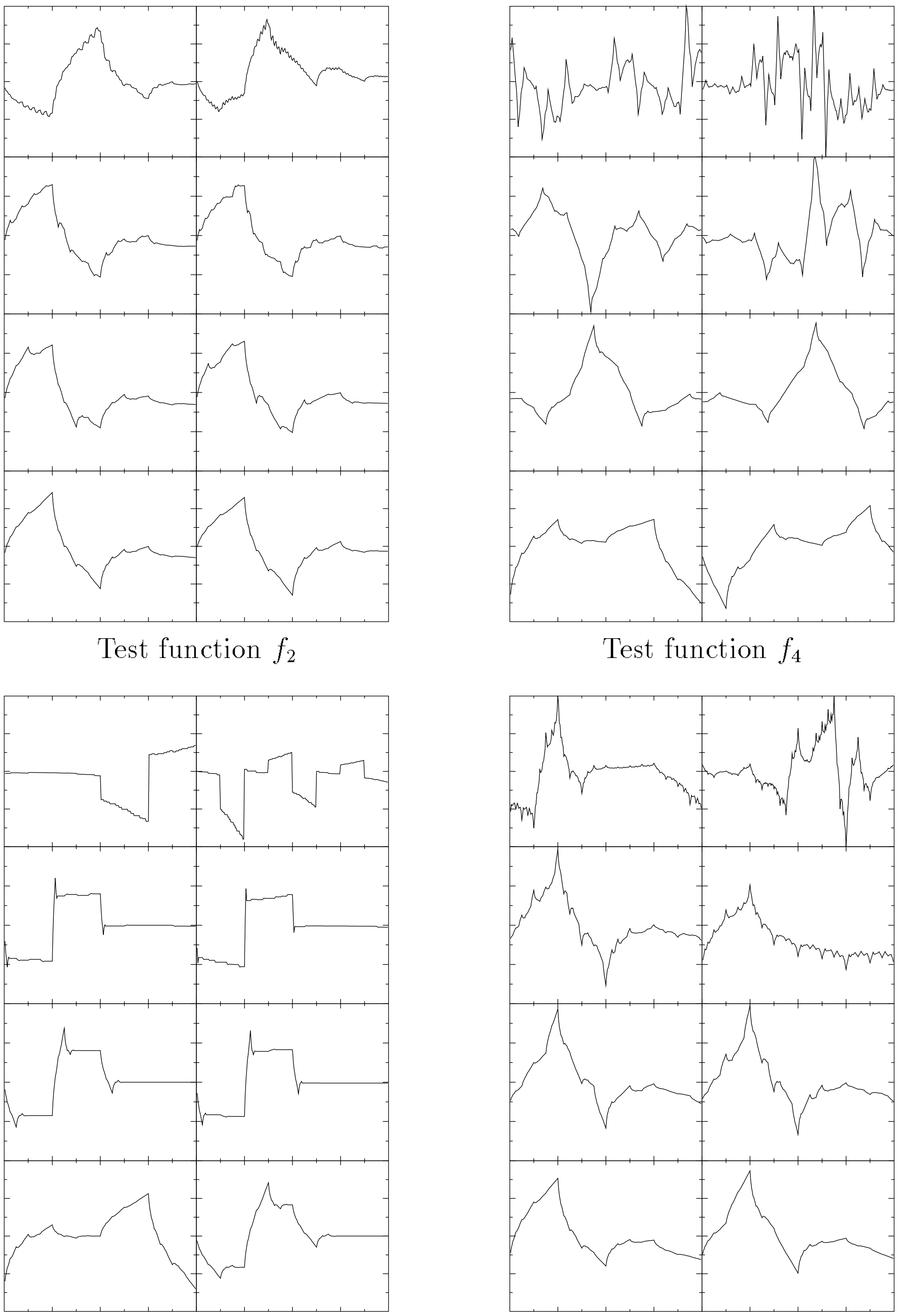
where $\alpha \in [0, \pi]$.

For $\alpha = 2\pi/3$ we get the Daubechies filter (14).

In this case the optimization problem is to find the maximum of a function $\Psi(\alpha)$ of one variable in the interval $[0, \pi]$ or to solve the equation

$$\frac{\partial\Psi}{\partial\alpha} = (F_{1,1} - F_{0,0} + F_{0,2} - F_{1,3})\cos 2\alpha \;+\; (F_{0,3} + F_{1,2} - 2F_{0,1})\sin 2\alpha$$

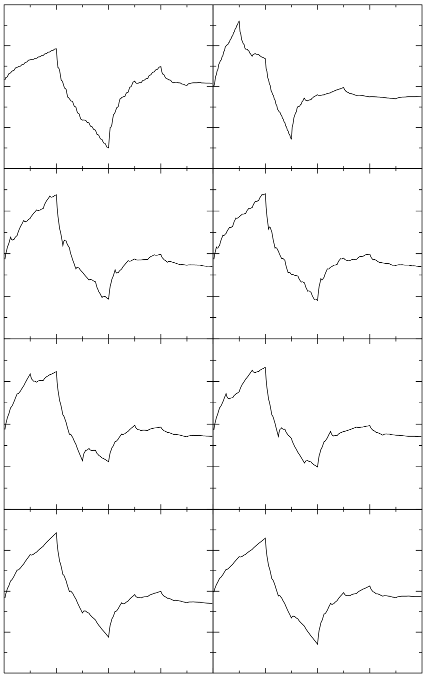
$$+\,(F_{0,3} - F_{1,2})\sin 2\alpha = 0.$$



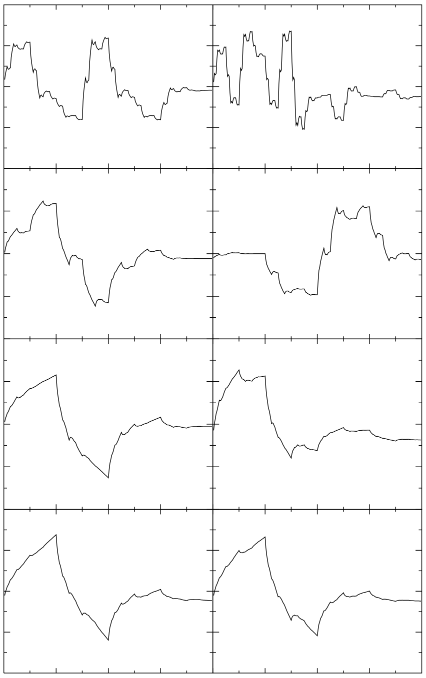
Test function f_2 Test function f_4

Test function f_1 Test function f_3

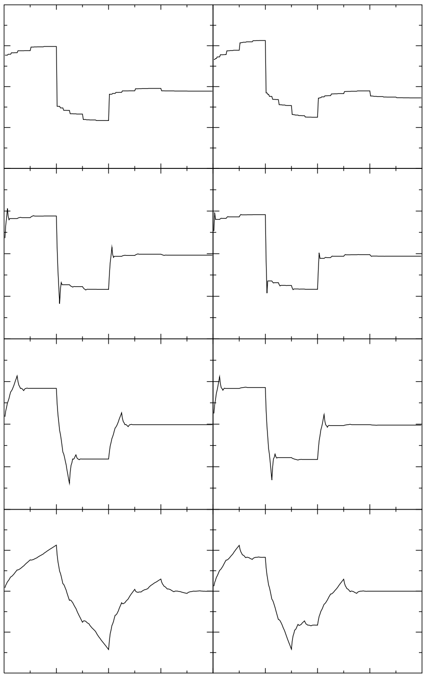
Figure 3: Rank 2, two parameters adaptation



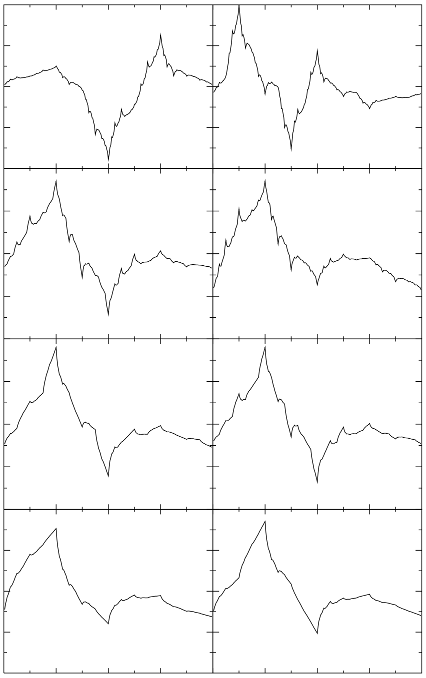
Test function f_2



Test function f_4



Test function f_1



Test function f_3

Figure 4: Rank 2, one parameter adaptation

On Fig. 4 the consecutive scaling functions of the AMRA of rank 2, adapted by one parameter to the respective test functions are depicted. On Table 1 we see the respective entropy of the wavelet coefficients.

Acknowledgment. I would like to thank my college Dr. Pencho Marinov for the various computer programs for calculation and visualization of the above results.

References

[1] G. Alexits (1961), *Convergence problems of orthogonal systems*, Publ. House of the Hungarian Acad. of Sci., Budapest.

[2] Barnsley, M. F. (1988), *Fractals Everywhere*, Boston.

[3] Barnsley, M. F. and L. P. Hurd (1993), *Fractal Image Compression*, AK Peters, Ltd. Wellesley, Massachusetts.

[4] J. M. Carnicer, W. Dahmen, and J. M. Peña (1996): *Local Decompositions of Refinable Spaces*, J. of Appl. and Comput. Harmon. Anal., **3**, 127 - 153.

[5] I. Daubechies (1992): *Ten Lectures on Wavelets* , SIAM, Philadelphia, PA.

[6] A. N. Kolmogorov (1936), *Über die beste annaherung von functionen einer functionsklasse*, Math. Ann., **37**, 107 - 111.

[7] S. Mallat (1989): Trans. Amer. Math. Soc., **315**, 69 - 88.

[8] Y. Meyer (1993), *Wavelets*, SIAM, Philadelphia.

[9] Bl. Sendov (1999), *Adaptive Multiresolution Analysis on the Dyadic Topological Group.* J. of Approx. Theory, **96**, 258 - 280.

[10] Bl. Sendov (1999): *Walsh-similar functions.* East J. on Approx., **5**(1): 1- 65.

[11] Bl. Sendov (1999): *Adaptive Approximation with Walsh-similar Functions*, Proceed. of IDoMAT 98: Approximation Theory, (to appear).

[12] W. Sweldens (1997): *The Lifting Scheme: a Construction of Second Generation Wavelets*, Siam J. Math. Anal., **29** (2), 512 - 546.