UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Multiresolution Representation of Cell-Averaged Data

Ami Harten

July 1994

CAM Report 94-21

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

Multiresolution Representation of Cell-Averaged Data.

Ami Harten
School of Mathematical Sciences
Tel-Aviv University
Tel-Aviv, 69978 ISRAEL
and
Department of Mathematics
UCLA

Abstract

In this paper we review recent developments in techniques to represent data in terms of their *local* scale components. These techniques enable us to obtain data compression by eliminating scale-coefficients which are sufficiently small. This capability for data compression can also be used to reduce the cost of many numerical solution algorithms. The purpose of this paper is to promote the use of multiresolution representation schemes which are based on cell-average discretization as the "Standard Method" for data compression.

1. Introduction

Fourier analysis, which provides a way to represent square-integrable functions in terms of their sinusoidal scale-components, has contributed greatly to all fields of science. The main drawback of Fourier analysis is in its globality – a single irregularity in the function dominates the behavior of the scale-coefficients and prevents us from getting immediate information about the behavior of the function elsewhere.

The recent development of the theory of wavelets (see [Me] and [Ma]) was a great step towards local scale decomposition, and has already had great impact on several fields of science. In numerical analysis representation by compactly supported wavelets (see [Da] and [CDF]) is used to reduce the cost of many numerical solution algorithms by either applying it to the numerical solution operator to obtain an approximate sparse form (see [BCR]), or by applying it to the numerical solution itself to obtain an approximate reduced representation in order to solve for less quantities (see [H4-5]). The main drawback of the theory of wavelets is that it attempts to decompose any square integrable function into scale-components which are translates and dilates of a single function. Consequently there are conceptual difficulties in extending wavelets to bounded domains and general geometries. Furthermore, the uniformity of the underlying wavelet approximation makes it impossible to obtain an adaptive (data-dependent) multiresolution *representation* which fits the approximation to the local nature of the data. The only adaptivity which is possible within the theory of wavelets is through redundant "dictionaries."

In a series of works [H1-3] we have studied the question of how to represent discrete data which originates from unstructured grids in bounded domains in terms of scale decomposition. Combining ideas from multigrid methods, numerical solution of conservation laws, hierarchial bases of finite element spaces, subdivision schemes of Computer-Aided Design and of course – the theory of wavelets, we came up with the more general concept of "nested sequence of discretization." We say that a sequence of linear operators $\{D_k\}_{k=0}^{\infty}$ is a nested sequence of discretization if

(i)

$$\mathcal{D}_k : \mathcal{F} \xrightarrow{onto} V^k, \quad \dim V^k = J_k, \tag{1.1a}$$

(ii)

$$\mathcal{D}_k f = 0 \Rightarrow \mathcal{D}_{k-1} f = 0 \tag{1.1b}$$

Here $\mathcal{F}$ is a space of mappings and V^k is a linear space of dimension J_k.

Given any discrete data $v^L = \mathcal{D}_L f$ we show in [H2-3] that it has a multiresolution representation, i.e., a one-to-one correspondence between the given data and its scale-decomposition:

$$v^L \overset{1:1}{\longleftrightarrow} \{d^L, \ldots, d^1, v^0\} =: \hat{v}. \tag{1.2a}$$

The k-th scale-coefficients $d^k = \{d^k_j\}_{j=1}^{J_k - J_{k-1}}$ represent the "difference in information" between $\mathcal{D}_k f$ and $\mathcal{D}_{k-1} f$, and $v^0 = \mathcal{D}_0 f$ is the discretization of f on the coarsest level. Observe that the number of components in $\hat{v}$ is the same as that of v^L because

$$\sum_{k=1}^{L} (J_k - J_{k-1}) + J_0 = J_L. \tag{1.2b}$$

A nested sequence of discretization comes equipped with a decimation operator

$$D_k^{k-1} : V^k \overset{onto}{\longrightarrow} V^{k-1}, \tag{1.3a}$$

which is defined by assigning to $v^k = \mathcal{D}_k f \in V^k$ the value of $\mathcal{D}_{k-1} f \in V^{k-1}$,i.e.

$$D_k^{k-1}(\mathcal{D}_k f) =: \mathcal{D}_{k-1} f. \tag{1.3b}$$

To see that the decimation operator is well defined by the above observe that (1.1a) implies that

$$V^k = \mathcal{D}_k(\mathcal{F}), \quad V^{k-1} = \mathcal{D}_{k-1}(\mathcal{F})$$

and that if $v^k = \mathcal{D}_k f_1 = \mathcal{D}_k f_2$, then it follows from the nestedness (1.1b) that $\mathcal{D}_{k-1} f_1 = \mathcal{D}_{k-1} f_2$ because

$$0 = \mathcal{D}_k f_1 - \mathcal{D}_k f_2 = \mathcal{D}_k(f_1 - f_2) \Longrightarrow 0 = \mathcal{D}_{k-1}(f_1 - f_2) = \mathcal{D}_{k-1} f_1 - \mathcal{D}_{k-1} f_2;$$

thus the definition $D_k^{k-1} v^k = \mathcal{D}_{k-1} f$ is independent of the particular f . We refer to the assignment of $f \in \mathcal{F}$ to $v^k \in V^k$ such that $v^k = \mathcal{D}_k f$ as reconstruction, and denote it by $\mathcal{R}_k$

$$\mathcal{R}_k : V^k \to \mathcal{F}, \quad \mathcal{D}_k \mathcal{R}_k = I_k, \tag{1.4}$$

where I_k is the identity operator in V^k ; note that $\mathcal{R}_k$ is a right-inverse of $\mathcal{D}_k$. We would like to point out that the reconstruction need not be a linear operator, i.e. it may depend on the data v^k.

Given any $v^L \in V^L$ we evaluate $\{v^k\}_{k=0}^{L-1}$ by repeated decimation

$$v^{k-1} = D_k^{k-1} v^k, \quad k = L, \ldots, 1. \tag{1.5}$$

Starting from v^{k-1} in (1.5) we approximae v^k by

$$v^k \approx \mathcal{D}_k(\mathcal{R}_{k-1} v^{k-1}),$$

and denote

$$P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1} : V^{k-1} \to V^k. \tag{1.6a}$$

P_{k-1}^k , the prediction operator, is a right-inverse of the decimation D_k^{k-1}

$$D_k^{k-1} P_{k-1}^k = I_{k-1}; \tag{1.6b}$$

To see that take $f = \mathcal{R}_{k-1} v^{k-1}$ in (1.3b) and use (1.4). We observe that the prediction error e^k

$$e^k = v^k - P_{k-1}^k v^{k-1} = (I_k - P_{k-1}^k D_k^{k-1}) v^k \tag{1.7a}$$

satisfies the relation

$$D_k^{k-1} e^k = D_k^{k-1} v^k - (D_k^{k-1} P_{k-1}^k) v^{k-1} = v^{k-1} - v^{k-1} = 0$$

and therefore it is in the null space of the decimation operator

$$e^k \in \mathcal{N}(D_k^{k-1}) = \{v | \quad v \in V^k,\ D_k^{k-1} v = 0\}. \tag{1.7b}$$

It follows from (1.3a) that

$$\dim \mathcal{N}(D_k^{k-1}) = J_k - J_{k-1} \tag{1.8a}$$

and therefore the prediction error e^k, which is described in terms of J_k components in V^k, can be represented in $\mathcal{N}(D_k^{k-1})$ by $(J_k - J_{k-1})$ coefficients $d^k = \{d_j^k\}_{j=1}^{J_k - J_{k-1}}$,which we consider as a representation of the k-th scale. To be specific, let $\{\mu_j^k\}_{j=1}^{J_k - J_{k-1}}$ be any basis of $\mathcal{N}(D_k^{k-1})$,

$$\mathcal{N}(D_k^{k-1}) = span\{\mu_j^k\}_{j=1}^{J_k - J_{k-1}}, \tag{1.8b}$$

and let d^k denote the coordinates of e^k in this basis

$$e^k = \sum_{j=1}^{J_k - J_{k-1}} d_j^k \mu_j^k =: E_k d^k, \quad d^k =: G_k e^k. \tag{1.9a}$$

Here G_k denotes the operator which assigns to $e^k \in \mathcal{N}(D_k^{k-1})$ its coordinates d^k in the basis $\{\mu_j^k\}_{j=1}^{J_k - J_{k-1}}$; observe that $E_k G_k$ is the identity operator in $\mathcal{N}(D_k^{k-1})$, i.e.

$$E_k G_k e^k = e^k \quad \text{for any } e^k \in \mathcal{N}(D_k^{k-1}). \tag{1.9b}$$

Using the above definition of scale-coefficients we obtain the multiresolution representation (MR) in (1.2a). The direct MR transform $\hat{v} = M \cdot v^L$ is given by the algorithm

$$\begin{cases} DO\ k = L, \ldots, 1 \\ \\ v^{k-1} = D_k^{k-1} v^k \\ \\ d^k = G_k (I_k - P_{k-1}^k D_k^{k-1}) v^k =: G_k^D v^k \end{cases} \tag{1.10}$$

The inverse MR transform $v^L = M^{-1} \cdot \hat{v}$ is given by

$$\begin{cases} DO\ k = 1, \ldots, L \\ \\ v^k = P_{k-1}^k v^{k-1} + E_k d^k. \end{cases} \tag{1.11}$$

In order to apply this multiresolution representation to real-life problems for purposes of analysis and data compression, we have to make sure that the direct MR transform and its inverse are stable with respect to perturbations. In [H3] we present stability analysis for MR schemes and derive a sufficient condition which seems to be "close" to necessary; this condition also implies existence of a multiresolution basis for mappings in $\mathcal{F}$. In appendix A we review some elements of this analysis and relate them to the particular examples of the present paper.

We remark that in multigrid terminology D_k^{k-1} is "restriction" and P_{k-1}^k is "prolongation." In signal processing D_k^{k-1} plays the role of "low-pass filter" while G_k^D, which is defined in (1.10), plays the role of "high-pass filter." This framework is a generalization of the theory of wavelets in the sense that under conditions of uniformity its natural result is wavelets.

The selection of a nested sequence of discretization in (1.1a) constitutes a setting for the multiresolution representation (1.1b), i.e. it determines the operators D_k^{k-1}, G_k and E_k. Once this is done, any choice of corresponding reconstruction operators $\{\mathcal{R}_k\}_{k=0}^L$ defines a MR scheme for discrete data v^L in V^L by taking $P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$ in the transforms (1.10)-(1.11). This opens up a tremendous number of possibilities for the design of MR schemes, where the primary consideration is the selection of an appropriate discretization. In this paper we claim that if we had to choose one of the many discretizations as a setting for the "standard data compression" scheme, then the best choice is that of cell-average discretization.

The considerations in selecting a particular discretization are : (1)Simplicity and flexibility; (2)Information contents; (3)Suitability to applications.

In section 2 we show how to construct a nested sequence of cell-average discretization in any compact set in any number of space-dimensions. In section 3 we consider the one-dimensional case and show that any known method of interpolation gives rise to a corresponding technique for reconstruction from cell-averages. Thus the only competition to cell-average discretization with respect to simplicity and flexibility comes from discretization by point values (see [AH]) ; however, cell-average discretization carries more information than point-value discretization.

In section 6 we demonstrate the difference in information contents between these two discretizations by describing the subcell reconstruction technique by which a discontinuous piecewise-polynomial function can be recovered exactly from its cell-averages; this is not true for point-value discretization. In [H1] we consider discretizations by weighted-averages

and conclude that there are no major differences in regions of smoothness, but there are differences in the information about irregularities of the sampled function, which can be recovered from higher moments of the data. Thus "hat-averages" contain more information than cell-averages, and one can exactly recover discontinuities as well as distributions in piecewise-polynomial functions from their "hat-averages" (see [ADH]). In this respect "spectral discretization", e.g. taking the first J_k Fourier coefficients, has the most information. However, cell-average discretization clearly wins over higher-moment discretizations in terms of simplicity and flexibility.

In section 4 we describe the relation between the biorthogonal wavelets of [CDF] and our formulation. We show that biorthogonal wavelets can be thought of as the "uniform constant-coefficient" case of our framework which corresponds to a choice of some weighted-average discretization in $\mathbb{R}$. We study the case which corresponds to cell-average discretization and show that the corresponding biorthogonal wavelets are the hierarchial form of piecewise-polynomial reconstruction. The "competition" between Daubechies' orthonormal wavelets [Da] to the biorthogonal wavelets of [CDF] is generally settled in favor of the latter, because in many applications the orthogonality doesn't buy you much, and it costs quite a bit.

In section 5 we present stability analysis of data compression algorithms, which is based on the results of [H3] and summarized in appendix A. We also present our ideas on error-control algorithms which enable us to prescribe an upper bound on the compression error, and provides a meaningful and stable way to apply adaptive techniques. Because of the simplicity of cell-average discretization we can actually prescribe a strategy for truncation of scale-coefficients and prove our contention about error control.

In section 7 we present some numerical experiments and discuss some aspects of suitability of the discretization to various applications.

2. MR schemes for cell-average discretization.

Consider absolutely integrable functions $f \in \mathcal{F}$

$$f : \Omega \subset \mathbb{R}^m \longrightarrow \mathbb{R}, \qquad \mathcal{F} = L^1(\Omega) \tag{2.1a}$$

where Ω is a compact set, and let $C^k = \{c_i^k\}_{i=1}^{J_k}$ be a set of cells such that

$$\Omega^k =: \overline{\cup_{i=1}^{J_k} c_i^k} \subseteq \Omega, \qquad c_i^k \cap c_j^k = \emptyset \ \text{ for } \ i \neq j, \tag{2.1b}$$

We define the cell-average discretization by

$$(\mathcal{D}_k f)_i = \frac{1}{|c_i^k|} \int_{c_i^k} f(x)dx, \qquad |c_i^k| = \int_{c_i^k} dx, \tag{2.1c}$$

and consider a refinement sequence $\{C^k\}_{k=0}^{L}$ in which C^k is formed from C^{k-1} by dividing each cell c_i^{k-1} into, say q, disjoint cells $\{c_{i_\ell}^k\}_{\ell=1}^{q}$,

$$\overline{\cup_{\ell=1}^{q} c_{i_\ell}^k} = \overline{c_i^{k-1}}. \tag{2.1d}$$

Alternatively we can consider (2.1d) to be a coarsening procedure in which we agglomerate every q cells of C^k into a larger cell of C^{k-1}; the only reason that we take here a fixed q is to simplify the notations. It follows from the additivity of the integral that

$$(\mathcal{D}_{k-1} f)_i = \frac{1}{|c_i^{k-1}|} \sum_{\ell=1}^{q} \int_{c_{i_\ell}^k} f(x)dx = \frac{1}{|c_i^{k-1}|} \sum_{\ell=1}^{q} |c_{i_\ell}^k| (\mathcal{D}_k f)_{i_\ell}. \tag{2.2a}$$

This shows that condition (1.1b) is satisfied, and provides the definition of the decimation operator in (1.3)

$$v_i^{k-1} = (D_k^{k-1} v^k)_i =: \frac{1}{|c_i^{k-1}|} \sum_{\ell=1}^{q} |c_{i_\ell}^k| v_{i_\ell}^k \tag{2.2b}$$

Let e^k denote the prediction error in (1.7a), then

$$D_k^{k-1} e^k = 0 \Rightarrow \sum_{\ell=1}^{q} |c_{i_\ell}^k| e_{i_\ell}^k = 0. \tag{2.3}$$

This relation shows that we can define the scale coefficients d^k by taking $(q-1)$ properly chosen linear combinations of the q prediction errors $\{e_{i_\ell}^k\}_{\ell=1}^{q}$ in each cell c_i^{k-1} . These linear combinations should be chosen so that together with (2.3) they constitute an invertible system of q linear equations for the prediction errors $\{e_{i_\ell}^k\}_{\ell=1}^{q}$ in the cell c_i^{k-1} . (see e.g. [HY] for such combinations in representation of matrices).

Using agglomeration as a coarsening technique may result in cells which are general polygons. In the following we describe a piecewise-polynomial reconstruction technique

which is suitable for this purpose (see [HC]): Let us denote by $\mathcal{S}_i^k$ a stencil of $s(r)$ *cells* in C^k which includes c_i^k, i.e.

$$\mathcal{S}_i^k = \{c_{i_m}^k\}_{m=1}^{s(r)}, \qquad c_i^k \in \mathcal{S}_i^k; \tag{2.4a}$$

here $s(r)$ is the number of coefficients in a polynomial of degree $(r-1)$ in $\mathbb{R}^m$. Let $p_i^k(x; \mathcal{D}_k f)$ denote the unique polynomial of degree $(r-1)$ which attains the averages $(\mathcal{D}_k f)_{i_m}$ in $\mathcal{S}_i^k$, i.e. the one which satisfies the following system of $s(r)$ linear equations for its $s(r)$ coefficients:

$$\frac{1}{|c_{i_m}^k|}\int_{c_{i_m}^k} p_i^k(x; \mathcal{D}_k f)dx = (\mathcal{D}_k f)_{i_m}, \qquad m = 1, \ldots, s(r), \tag{2.4b}$$

and define

$$(\mathcal{R}_k \mathcal{D}_k f)(x) = p_i^k(x; \mathcal{D}_k f) \quad \text{for} \quad x \in c_i^k. \tag{2.4c}$$

Clearly (2.4) defines a reconstruction of $\mathcal{D}_k f$ which is exact for polynomial functions of degree less or equal $(r-1)$, and thus is r-th order accurate.

Note that for $r = 1$ in (2.4) we have $s(r) = 1$ and we get the piecewise-constant reconstruction

$$(\mathcal{R}_k \mathcal{D}_k f)(x) = \sum_i (\mathcal{D}_k f)_i \chi_{c_i^k}(x), \tag{2.5a}$$

where $\chi_C(x)$ denotes the characteristic function of the set C,

$$\chi_C(x) = \begin{cases} 1 & x \in C \\ 0 & \text{otherwise} \end{cases}; \tag{2.5b}$$

In [HC] we present an hierarchial algorithm for the selection of a "centered" stencil, which is applicable even to completely unstructured meshes C^k in $\mathbb{R}^m$. In this context the "centered" stencil is defined as the one which minimizes the reconstruction error for the one-higher degree polynomials (i.e. degree r). This algorithm is of "crystal growth" type: starting with the cell c_i^k we begin to add successively, one cell at a time, to the cluster of cells that we have at the beginning of each step. The cell which is being added is selected from the set of all side-neighbors of the existing cluster by the requirment that it will minimize the reconstruction error of suitably chosen monomials.

In [HC] we also present an adaptive "crystal growth" algorithm which is designed to assign a stencil $\mathcal{S}_i^k$ from the smooth part of $f(x)$, if available, to all cells c_i^k which are themselves in the smooth part of $f(x)$. This way a Gibbs-like phenomenon is avoided, and the resulting approximation is r-th order accurate everywhere, except at cells which contain a discontinuity. This is accomplished by selecting the cell from the set of side-neighbors which minimizes the derivatives of the so-defined reconstruction. We refer to this adaptive technique as Essentially Non-Oscillatory (ENO) reconstuction, and we shall describe a one-dimensional version of it in section 6 of this paper.

We refer the reader to [Ab] for details of special ENO reconstruction techniques for triangulated meshes.

3. MR schemes in [0,1].

In this section we take in (2.1) $\Omega = [0,1]$ and let X^L be an arbitrary partition of $[0,1]$

$$X^L = \{x_i^L\}_{i=0}^{J_L},\ x_0^L = 0,\ x_{J_L}^L = 1,\ J_L = 2^L J_0 \tag{3.1a}$$

where the sequence above is strictly increasing and J_0 is some integer. We define the grids $X^k = \{x_i^k\}_{i=0}^{J_k},\ k = L-1,\ldots,1$ by the coarsening

$$x_i^{k-1} = x_{2i}^k, \quad i = 0,\ldots,J_{k-1} =: J_k/2, \tag{3.1b}$$

in which we delete from X^k all the points with odd indices (The only reason that we remove *every* other point is to simplify the notations). We consider the covering of $[0,1]$ by cells

$$C^k = \{c_i^k\}_{i=1}^{J_k}, \quad c_i^k = (x_{i-1}^k, x_i^k), \tag{3.2a}$$

where $\{x_i^k\}$ are the gridpoints of X^k in (3.1); observe that

$$\overline{c_i^{k-1}} = \overline{c_{2i-1}^k \cup c_{2i}^k}. \tag{3.2b}$$

Let $v^k = \mathcal{D}_k f$ denote the cell-averages in (2.1c); v^{k-1} is obtained from v^k by the decimation (2.2b), i.e.

$$v_i^{k-1} = \frac{1}{|c_i^{k-1}|}(|c_{2i-1}^k| v_{2i-1}^k + |c_{2i}^k| v_{2i}^k),\ i = 1,\ldots,J_{k-1}. \tag{3.3}$$

From (1.7b) we get that the prediction error e^k satisfies the relation

$$|c_{2i-1}^k| e_{2i-1}^k + |c_{2i}^k| e_{2i}^k = 0 \quad for \quad i = 1,\ldots,J_{k-1}. \tag{3.4a}$$

Therefore, if we define $d^k = G_k e^k$ in (1.9a) by

$$d_j^k = \frac{1}{|c_j^{k-1}|}(|c_{2j-1}^k| e_{2j-1}^k - |c_{2j}^k| e_{2j}^k) \quad for \quad j = 1,\ldots,J_{k-1} \tag{3.4b}$$

we recover the prediction error by $e^k = E_k d^k$ which is defined as follows:

$$\begin{cases} e_{2i-1}^k = d_i^k \cdot \frac{|c_i^{k-1}|}{2|c_{2i-1}^k|} \\ \\ e_{2i}^k = -d_i^k \cdot \frac{|c_i^{k-1}|}{2|c_{2i}^k|} \quad . \end{cases} \tag{3.4c}$$

The direct MR transform $\hat{v} = M \cdot v^L$ is given in this case by the algorithm

$$\begin{cases} DO\ k = L, \ldots, 1 \\ v_i^{k-1} = \frac{1}{|c_i^{k-1}|}(|c_{2i-1}^k| v_{2i-1}^k + |c_{2i}^k| v_{2i}^k), \quad i = 1, \ldots, J_{k-1} \\ d_j^k = \frac{2|c_{2j-1}^k|}{|c_j^{k-1}|}[v_{2j-1}^k - (P_{k-1}^k v^k - 1)_{2j-1}], \quad j = 1, \ldots, J_{k-1} \end{cases} \tag{3.5}$$

The inverse MR transform $v^L = M^{-1} \cdot \hat{v}$ is given by

$$\begin{cases} DO\ k = 1, \ldots, L \\ DO\ i = 1, \ldots, J_{k-1} \\ v_{2i-1}^k = (P_{k-1}^k v^k - 1)_{2i-1} + \frac{|c_i^{k-1}|}{2|c_{2i-1}^k|} d_i^k \\ v_{2i}^k = (|c_i^{k-1}| v_i^{k-1} - |c_{2i-1}^k| v_{2i-1}^k)/|c_{2i}^k| \end{cases} \tag{3.6}$$

Observe that the last statement of (3.6) is obtained from (3.3), and it is the "inverse" of the first step in algorithm (3.5).

In [HEOC] we showed that any interpolation method gives rise to a corresponding method for reconstruction from cell-averages by the following "reconstruction via primitive function" technique: Given cell-averages $v^k = \mathcal{D}_k f$ we calculate the point values of the "primitive function"

$$F_i^k = F(x_i^k), \quad F(x) = \int_0^x f(y)dy$$

by

$$F_0^k = 0, \quad F_i^k = \sum_{j=1}^{i} |c_j^k| v_j^k, \quad 1 \le i \le J_k, \tag{3.7a}$$

and define

$$(\mathcal{R}_k v^k)(x) = \frac{d}{dx} \mathcal{I}_k(x; F^k), \tag{3.7b}$$

where $\mathcal{I}_k(x; F^k)$ is *any* interpolation of the values $F^k = \{F_i^k\}_{i=0}^{J_k}$ at the grid points of X^k in (3.2).

The prediction operator (1.6a) can be expressed in terms of the interpolation of the primitive values by

$$(P_{k-1}^k v^k - 1)_{2i-1} = \frac{1}{|c_{2i-1}^k|}[\mathcal{I}_{k-1}(x_{2i-1}^k; F^{k-1}) - F_{i-1}^{k-1}]. \tag{3.8}$$

The piecewise-polynomial reconstruction in (2.4) can be obtained via (3.7) from the following piecewise-polynomial interpolation: Let $\mathcal{S}_i^k$ denote a stencil of r consecutive *cells*

of C^k which includes c_i^k and let $q_i^k(x;F^k)$ denote the unique polynomial of degree r which interpolates F^k at the $r+1$ endpoints of the cells in this stencil . We define the piecewise-polynomial reconstruction $(\mathcal{R}_k v^k)(x)$ by

$$(\mathcal{R}_k v^k)(x) = p_i^k(x;v^k) =: \frac{d}{dx} q_i^k(x;F^k) \quad \text{for} \quad x \in c_i^k; \tag{3.9}$$

observe that $p_i^k(x;v^k)$ is of polynomial degree $(r-1)$. For $r=1$ we have $\mathcal{S}_i^k = \{c_i^k\}$ and $q_i^k(x;F^k)$ is the linear interpolation

$$q_i^k(x;F^k) = F_{i-1}^k + \frac{x - x_{i-1}^k}{x_i^k - x_{i-1}^k}(F_i^k - F_{i-1}^k) \quad for \quad x_{i-1}^k \le x \le x_i^k; \tag{3.10a}$$

in this case $\mathcal{R}_k v^k$ is the piecewise-constant reconstruction

$$(\mathcal{R}_k v^k)(x) = v_i^k \quad for \quad x \in c_i^k. \tag{3.10b}$$

Up to now we have not specified the stencil $\mathcal{S}_i^k$ of r consecutive cells of C^k that we assign to c_i^k. Clearly if we choose $\mathcal{S}_i^k$ independently of the data v^k then the most accurate choice is that of a centered stencil (away from the boundaries), i.e. for $r = 2s+1$ we take

$$\mathcal{S}_i^k = \{c_{i-s}^k, \ldots, c_{i+s}^k\} \quad \text{for} \quad s+1 \le i \le J_k - s, \tag{3.11a}$$

and near the boundaries

$$\begin{aligned} &\mathcal{S}_i^k = \{c_1^k, \ldots, c_r^k\} \quad \text{for} \quad 1 \le i \le s \\ &\mathcal{S}_i^k = \{c_{J_k-r+1}^k, \ldots, c_{J_k}^k\} \quad \text{for} \quad J_k - s + 1 \le i \le J_k. \end{aligned} \tag{3.11b}$$

When the grid X^k is uniform we get a particularly simple expression for the prediction $P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$ in (3.10)-(3.11) . For the centered stencil (3.11a) we get that the evaluation of the predicted value requires only s multiplications:

$$(P_{k-1}^k v^{k-1})_{2i-1} = v_i^{k-1} - \sum_{\ell=1}^{s} \gamma_\ell (v_{i+\ell}^{k-1} - v_{i-\ell}^{k-1}), \tag{3.12a}$$

where

$$\begin{cases} r = 3 \Longrightarrow \gamma_1 = -1/8 \\ \\ r = 5 \Longrightarrow \gamma_1 = -22/128, \ \gamma_2 = 3/128. \end{cases} \tag{3.12b}$$

In Figures 1 and 2 we show the subdivision limit of the piecewise-polynomial reconstruction for $r=3$. This is done by applying the inverse MR transform with $L=7$ levels and $J_0 = 8$ to $\hat{v} = \{0, \ldots, 0, \eta_i^0\}$,where $(\eta_i^0)_j = \delta_{i,j}$ for $j = 1, \ldots, J_0$, to obtain an approximation to

$$\varphi_i^0 = \lim_{L \to \infty} \prod_{\ell=0}^{L} (\mathcal{R}_\ell \mathcal{D}_\ell) \cdot (\mathcal{R}_0 \eta_i^0), \tag{3.13a}$$

and to $\hat{v} = \{0, \ldots, 0, \eta_i^0, 0\}$ to obtain an approximation to

$$\psi_i^1 = \lim_{L \to \infty} \prod_{\ell=1}^{L} (\mathcal{R}_\ell \mathcal{D}_\ell) \cdot (\mathcal{R}_1 \mu_i^1) \tag{3.13b}$$

(see appendix A). The subscripts a,b,c,d in Figures 1 and 2 stand for $i = 1, 2, 3, 4$, respectively. The limit functions for $i = 5, 6, 7, 8$ can be obtained from the former ones by an appropriate reflection.

In the next section we shall relate these limit functions to biorthogonal wavelets.

4. Biorthogonal wavelets

In this section we derive the MR schemes which correspond to the bases of biorthogonal wavelets in [CDF]. These MR schemes are obtained from nested discretization of functions in $L^2_{loc}(\mathbb{R})$ by taking weighted-averages on a nested dyadic sequence of uniform grids of $\mathbb{R}$, as follows:

$$(\mathcal{D}_k f)_i = \frac{1}{h_k} \int_{-\infty}^{\infty} f(x) w\left(\frac{x - x_i^k}{h_k}\right) dx, \ -\infty < i < \infty, \tag{4.1a}$$

where $w \in L^2(\mathbb{R})$ is a weight-funcntion

$$\int_{-\infty}^{\infty} w(x) dx = 1 \tag{4.1b}$$

of compact support and

$$x_i^k = i h_k, \ -\infty < i < \infty, \ h_k = 2^{-k} h_0. \tag{4.1c}$$

In [H2] we show that if $w(x)$ is a solution of a dilation equation

$$w(x) = 2 \sum_{\ell=0}^{N} \alpha_\ell w(2x - \ell). \tag{4.2a}$$

with coefficients that satisfy

$$\sum_{\ell} \alpha_{2\ell} = \sum_{\ell} \alpha_{2\ell-1} = 1/2, \tag{4.2b}$$

then $\{\mathcal{D}_k\}$ is a nested sequence of operators and its decimation operator is given by

$$(D_k^{k-1} v)_i =: \sum_{\ell} \alpha_\ell v_{2i-\ell}. \tag{4.2c}$$

In [Da] and [H3] it is shown that $\{\mu_j^k\}_{j=-\infty}^{\infty}$,

$$(\mu_j^k)_i = (-1)^{i+1} \alpha_{2j-i-1}, \ \ -\infty < i < \infty \tag{4.3}$$

is a basis of the null space of the decimation operator $\mathcal{N}(D_k^{k-1})$ in (1.8b).

In order to have a "wavelet basis" of $L^2(\mathbb{R})$ which consists of dilates and translates of a single function, we consider reconstruction of the form

$$(\mathcal{R}_k v)(x) = \sum_i v_i \varphi\left(\frac{x - x_i^k}{h_k}\right), \tag{4.4a}$$

where $\varphi(x)$ is a solution of the dilation equation

$$\varphi(x) = \sum_\ell \beta_\ell \varphi(2x - \ell) \tag{4.4b}$$

which is normalized by

$$\int \varphi(x) w(x) dx = 1. \tag{4.4c}$$

In [H2] we show that the coefficients $\{\beta_\ell\}$ have to satisfy the following relations

$$\sum_\ell \alpha_\ell \beta_{\ell+2m} = \delta_{m,0} \tag{4.5a}$$

$$\sum_\ell (-1)^\ell (\ell)^q \beta_\ell = 0 \quad \text{for} \quad q = 0, \ldots, (r-1), \quad r \geq 1. \tag{4.5b}$$

Relation (4.5a) is derived from the requirement that $\mathcal{R}_k$ in (4.4a) is indeed a reconstruction of the discretization (4.1), i.e. that $\mathcal{D}_k \mathcal{R}_k = I_k$. Relation (4.5b) is derived from the requirement that the reconstruction is exact for polynomial functions p of degree less or equal $(r-1)$, i.e. that $\mathcal{R}_k \mathcal{D}_k p = p, \quad deg(p) \leq (r-1)$.

It is easy to see that the corresponding prediction operator $P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$ is given by

$$(P_{k-1}^k v)_i = \sum_m \beta_{i-2m} v_m \tag{4.6}$$

and that the reconstruction $\mathcal{R}_k$ is hierarchial (see (A.4a) in appendix A). In this case we show in [H3] that for any $f \in \mathcal{F}$

$$\mathcal{R}_L \mathcal{D}_L f = \mathcal{R}_0 \mathcal{D}_0 f + \sum_{k=1}^{L} \sum_{j=-\infty}^{\infty} d_j^k \psi_j^k \ , \tag{4.7a}$$

where

$$\psi_j^k = \mathcal{R}_k \mu_j^k. \tag{4.7b}$$

and $d_j^k = d_j^k(f)$ are the scale coefficients in (1.10) :

$$d^k(f) = G_k e^k(f) = G_k \mathcal{D}_k (I - \mathcal{R}_{k-1} \mathcal{D}_{k-1}) f. \tag{4.7c}$$

Consequetly $\{\psi_j^k\}_{j,k}$ is a wavelet basis of any Banach space $\mathcal{F}$ in which $\{(\mathcal{R}_k \mathcal{D}_k)\}_{k=0}^{\infty}$ is a sequence of approximation (see appendix A).

Following the guidelines of the present paper, we first select a sequence of nested discretization and thus create a setting for MR schemes; in the framework of this section this amounts to a choice of coefficients $\{\alpha_\ell\}$ subject to conditions (4.2b). Once this is done relations (4.5) become a system of *linear* equations for the coefficients $\{\beta_\ell\}$. For

$$\alpha_\ell = \frac{1}{2}(\delta_{\ell,0} + \delta_{\ell,-1}) \tag{4.8a}$$

we get

$$w(x) = \begin{cases} 1 & -1 < x < 0 \\ 0 & otherwise \end{cases} \tag{4.8b}$$

which is square integrable, and the discretization (4.1a) becomes the cell-average discretization (2.1c) in the uniform grid (4.1c). It is easy to verify that if we take $2r$ nonzero coefficients $\{\beta_\ell\}_{\ell=-r}^{r-1}$ then we get r linear equations from (4.5a), and another set of r linear equations from (4.5b). The solution for this system of $2r$ equations is unique and the resulting prediction operator (4.6) is identical to that of the centered stencil in (3.12a). We conclude that the two MR schemes are the same in the case of uniform grids in $\mathbb{R}$, and that the reconstruction (4.4a) of the biorthogonal wavelets is the hierarchial form of the piecewise polynomial reconstruction in (3.9) (see appendix A). Thus in Figure 1d $\varphi_4^0(x) = \varphi(\frac{x}{h_0} - 4)$, where $\varphi(x)$ is the solution of the dilation equation (4.4b), and in Figure 2d $\psi_4^1(x) = \psi(\frac{x}{h_1} - 4)$, where $\psi(x)$ is the "mother wavelet function" in the expansion (4.7). The boundary limit functions in Figures 1 and 2 show how to augment the scaling functions and wavelets of the infinite domain in order to get a "wavelet basis" for the interval [0,1].

In [ADH] we take $w(x)$ in (4.1b) to be the "hat-function" and show how to derive MR schemes in this case.

Remark 4.1. Daubechies' orthonormal wavelets are obtained from the above formulation by adding the requirement $\beta_\ell = 2\alpha_\ell$ which couples the systems of equations for $\{\alpha_\ell\}$ and $\{\beta_\ell\}$, and then the biortogonality condition (4.5a) becomes a *nonlinear* orthogonality condition. Observe that as we change the order of accuracy r , we necessarily change $\{\alpha_\ell\}$ and consequently the nature of the discretization. In this respect Daubechies' orthonormal wavelets deviate from our set up in which the discretization is fixed, and we change the order of the reconstruction if desired. Also observe that we did not use the notion of orthogonality or biortogonality in our formulation. The latter is an automatic consequence of the relation $\mathcal{D}_k\mathcal{R}_k = I_k$ and it does not impose an independent requirement.

5. Data compression and error control .

In this section we consider strategies for data compression and apriori bounds on the compression error.

We can obtain data compression by setting to zero all scale coefficients which fall below a prescribed tolerance. Let us denote

$$\tilde{d}_j^k = tr(d_j^k;\varepsilon_k) =: \begin{cases} 0 & if\ |d_j^k| \le \varepsilon_k \\ \\ d_j^k & if\ |d_j^k| > \varepsilon_k, \end{cases} \tag{5.1a}$$

and refer to this operation as truncation. This type of data compression is used primarily to reduce the "dimensionality" of the data. Another aspect of data compression is to reduce the digital representation of the data for purposes of storage or transmission. In this case we use "quantization" which we model by

$$\tilde{d}_j^k = qu(d_j^k;\varepsilon_k) =: 2\varepsilon_k \cdot ROUND\left[\frac{d_j^k}{2\varepsilon_k}\right], \tag{5.1b}$$

where $ROUND\,[\,\cdot\,]$ denotes the integer which is obtained by rounding of the number. For example if $|d_j^k| \le 256$ and $\varepsilon_k = 4$ then we can represent $|d_j^k|$ by an integer which is not larger than 32 and commit a maximal error of 4. Observe that $|d_j^k| < \varepsilon_k \Longrightarrow qu(d_j^k;\varepsilon_k) = 0$ and that in both cases

$$|d_j^k - tr(d_j^k;\varepsilon_k)| \le \varepsilon_k, \tag{5.2a}$$

$$|d_j^k - qu(d_j^k;\varepsilon_k)| \le \varepsilon_k. \tag{5.2b}$$

Let

$$\tilde{v}^L = M^{-1} \cdot \{\tilde{d}^L, \ldots, \tilde{d}^1, v^0\}, \tag{5.3a}$$

denote the "decompressed" data; then, under the circmstances which are described in appendix A, we typically get the following bound on the compression error

$$|\tilde{v}^L - v^L|_L \le C \cdot \sum_{k=1}^{L} \varepsilon_k \tag{5.3b};$$

here $|\,\cdot\,|_L$ is the discrete norm which is defined in (A.2b) in appendix A and C is a constant independent of L . It should be noted that we assume here that M^{-1} is a linear operator. Given any $\varepsilon > 0$ we can take for example

$$\varepsilon_k = (1-q)q^{L-k}\varepsilon \quad \text{for some} \quad 0 < q < 1, \tag{5.4a}$$

in which case we get from (5.3b) that

$$|\tilde{v}^L - v^L|_L \le C \cdot \varepsilon \tag{5.4b}.$$

Only in very simple situations we can choose a sequence of tolerances ε_k which ensures that for a prescribed ε

$$\max_{0 \le i \le J_L} |v_i^L - \tilde{v}_i^L| < \varepsilon. \tag{5.5}$$

Thus the typical situation in this set up is that we can specify the rate of compression, but we cannot specify an upper bound on the compression error.

In the following we describe a technique which enables us to specify a desired level of accuracy in the decompressed signal, independent of the particular form of the prediction operator; most importantly it also applies to nonlinear (adaptive) prediction methods. This is accomplished by using a modification of the encoding algorithm (1.10) which keeps track of the cumulative compression error in the predetermined decoding procedure (1.11) and compresses accordingly. As is to be expected , we cannot specify compression rate at the same time. This algorithm applies to both the truncation (5.1a) and the quantization (5.1b), and we denote their operation by the generic name $cm(d^k;\varepsilon_k)$, where cm stands for compression.

We denote our error control encoding by M_ε and describe its operation on the input v^L by

$$\hat{v}_\varepsilon =: M_\varepsilon \cdot v^L =: \{\tilde{d}^L, \ldots, \tilde{d}^1, v^0\}. \tag{5.6}$$

First we apply successive decimation to the input v^L

$$\begin{cases} DO\ k = L, \ldots, 1 \\ v^{k-1} = D_k^{k-1} v^k \end{cases} \tag{5.7a}$$

These values of v^k are used in the following to monitor and control the accumulation of the compression error: Set

$$\tilde{v}^0 = v^0 \tag{5.7b}$$

$$\begin{cases} DO\ k = 1, \ldots, L \\ v^P = P_{k-1}^k \tilde{v}^{k-1} \\ \tilde{d}^k = cm[G_k(v^k - v^P);\varepsilon_k] \\ \tilde{v}^k = v^P + E_k \tilde{d}^k. \end{cases} \tag{5.7c}$$

Clearly, if we apply the decoding M^{-1} :

$$\begin{cases} DO\ k = 1, \ldots, L \\ v^P = P_{k-1}^k v^{k-1} \\ v^k = v^P + E_k d^k \end{cases} \tag{5.8a}$$

to the encoded data $\hat{v}_\varepsilon$ we get that

$$M^{-1} \cdot \{\tilde{d}^L, \ldots, \tilde{d}^1, v^0\} = \tilde{v}^L \tag{5.8b}$$

where $\tilde{v}^L$ is the quantity which is computed in (5.7c) for $k = L$. Our task is now to find G_k and ε_k such that (5.5) holds.

Next we apply this program to the case of cell-average discretization in a uniform grid of $[0,1]$; here because of the simplicity of the expressions we can precisely monitor the cumulative compression error in the decoding algorithm and show how to control it. For simplicity we consider compression by truncation (5.1a). The predetirmined decoding procedure is (3.6), which we now rewrite in the form

$$\begin{cases} DO\ k = 1, \ldots, L \\ DO\ i = 1, \ldots, J_{k-1} \\ v^P_{2i-1} = (P^k_{k-1} v^{k-1})_{2i-1} \\ v^k_{2i-1} = v^P_{2i-1} + d^k_i \\ v^k_{2i} = 2v^{k-1}_i - v^k_{2i-1} \end{cases} \tag{5.9}$$

The modified encoding procedure is described algorithmically by the following:

(i) Apply decimation to the input v^L

$$\begin{cases} DO\ k = L, \ldots, 1 \\ DO\ i = 1, \ldots, J_{k-1} \\ v^{k-1}_i = \frac{1}{2}(v^k_{2i-1} + v^k_{2i}) \end{cases} \tag{5.10a}$$

(ii) Set

$$\tilde{v}^0 = v^0 \tag{5.10b}$$

(iii) Calculate

$$\begin{cases} DO\ k = 1, \ldots, L \\ DO\ j = 1, \ldots, J_{k-1} \\ v^P_{2j-1} = (P^k_{k-1} \tilde{v}^{k-1})_{2j-1} \\ \tilde{d}^k_j = tr(v^k_{2j-1} - v^P_{2j-1} - (v^{k-1}_j - \tilde{v}^{k-1}_j); \varepsilon_k) \\ \tilde{v}^k_{2j-1} = v^P_{2j-1} + \tilde{d}^k_j \\ \tilde{v}^k_{2j} = 2\tilde{v}^{k-1}_j - \tilde{v}^k_{2j-1}. \end{cases} \tag{5.10c}$$

Note that we used the relation

$$v^P_{2j-1} + v^P_{2j} = 2\tilde{v}^{k-1}_j$$

in order to rewrite the expression for $\tilde{d}_j^k$ in (5.7b) as

$$[G_k(v^k - v^P)]_j =: \frac{1}{2}[(v_{2j-1}^k - v_{2j-1}^P) - (v_{2j}^k - v_{2j}^P)] = v_{2j-1}^k - v_{2j-1}^P - (v_j^{k-1} - \tilde{v}_j^{k-1}).$$

Let us denote the cumulative compression error by

$$\mathcal{E}_j^k = v_j^k - \tilde{v}_j^k \tag{5.11a}$$

and the prediction error

$$e_{2j-1}^P = v_{2j-1}^k - v_{2j-1}^P. \tag{5.11b}$$

With this notation we get from (5.10c) that

$$\mathcal{E}_{2j-1}^k = e_{2j-1}^P - tr(e_{2j-1}^P - \mathcal{E}_j^{k-1}; \varepsilon_k), \tag{5.11c}$$

$$\frac{1}{2}(\mathcal{E}_{2j-1}^k + \mathcal{E}_{2j}^k) = \mathcal{E}_j^{k-1}. \tag{5.11d}$$

Subtracting (5.11c) from (5.11d) we get

$$\frac{1}{2}(\mathcal{E}_{2j}^k - \mathcal{E}_{2j-1}^k) = \mathcal{E}_j^{k-1} - e_{2j-1}^P + tr(e_{2j-1}^P - \mathcal{E}_j^{k-1}; \varepsilon_k). \tag{5.12}$$

Let us now examine the two possibilities in (5.12):

$$|\mathcal{E}_j^{k-1} - e_{2j-1}^P| \geq \varepsilon_k \Rightarrow \frac{1}{2}(\mathcal{E}_{2j-1}^k - \mathcal{E}_{2j}^k) = 0 \Rightarrow \mathcal{E}_{2j-1}^k = \mathcal{E}_{2j}^k = \mathcal{E}_j^{k-1}, \tag{5.13a}$$

$$|\mathcal{E}_j^{k-1} - e_{2j-1}^P| < \varepsilon_k \Rightarrow \begin{cases} \frac{1}{2}(\mathcal{E}_{2j}^k - \mathcal{E}_{2j-1}^k) = \mathcal{E}_j^{k-1} - e_{2j-1}^P \\ \frac{1}{2}(\mathcal{E}_{2j}^k + \mathcal{E}_{2j-1}^k) = \mathcal{E}_j^{k-1}. \end{cases} \tag{5.13b}$$

From (5.13) we get the following inequalities

$$\max(|\mathcal{E}_{2j-1}^k|, |\mathcal{E}_{2j}^k|) = \frac{1}{2}|\mathcal{E}_{2j}^k + \mathcal{E}_{2j-1}^k| + \frac{1}{2}|\mathcal{E}_{2j}^k - \mathcal{E}_{2j-1}^k| \leq |\mathcal{E}_j^{k-1}| + \varepsilon_k, \tag{5.14a}$$

$$(|\mathcal{E}_{2j-1}^k| + |\mathcal{E}_{2j}^k|) = \max(|\mathcal{E}_{2j}^k + \mathcal{E}_{2j-1}^k|, |\mathcal{E}_{2j}^k - \mathcal{E}_{2j-1}^k|) \leq 2\max(|\mathcal{E}_j^{k-1}|, \varepsilon_k). \tag{5.14b}$$

Recalling that $\mathcal{E}^0 = 0$ we get from (5.14a)

$$\|\mathcal{E}^k\|_\infty \leq \|\mathcal{E}^{k-1}\|_\infty + \varepsilon_k \leq \cdots \leq \sum_{\ell=1}^{k} \varepsilon_\ell; \tag{5.15a}$$

Recalling that here $|c_i^k| = h_k = \frac{1}{2}h_{k-1}$ we get from (5.14b)

$$\begin{aligned} \|\mathcal{E}^k\|_{\ell_1} = h_k \textstyle\sum_{i=1}^{J_k} |\mathcal{E}_i^k| &= \tfrac{1}{2}h_{k-1} \textstyle\sum_{j=1}^{J_{k-1}} (|\mathcal{E}_{2j-1}^k| + |\mathcal{E}_{2j}^k|) \\ &\leq h_{k-1} \textstyle\sum_{j=1}^{J_{k-1}} \max(|\mathcal{E}_j^{k-1}|, \varepsilon_k). \end{aligned} \tag{5.15b}$$

It follows from (5.15) that

$$\|\mathcal{E}^L\|_\infty \le \sum_{\ell=1}^{L} \varepsilon_\ell, \tag{5.16a}$$

and if we choose $\{\varepsilon_\ell\}_{\ell=1}^{L}$ such that for all $\ell \ge 2$

$$\varepsilon_\ell \ge \sum_{m=1}^{\ell-1} \varepsilon_m \tag{5.16b}$$

then the ℓ_1-error is

$$\|\mathcal{E}^L\|_{\ell_1} \le \varepsilon_L. \tag{5.16c}$$

Given ε it makes good sense to choose the tolerance-levels ε_k to be

$$\varepsilon_k = (1-q)q^{L-k}\varepsilon \quad \text{for some} \quad 0 < q \le \frac{1}{2}; \tag{5.17a}$$

in this case we get

$$\|v^L - \tilde{v}^L\|_\infty = \|\mathcal{E}^L\|_\infty \le \varepsilon, \tag{5.17b}$$

$$\|v^L - \tilde{v}^L\|_{\ell_1} = \|\mathcal{E}^L\|_{\ell_1} \le (1-q)\varepsilon. \tag{5.17c}$$

6. ENO reconstruction and subcell resolution

In [HEOC] we presented a data-dependent piecewise-polynomial reconstruction technique which avoids the Gibbs-phenomenon by an adaptive selection of stencil $\mathcal{S}_i^k$ in (3.9) ; we refer to this technique as Essentially Non-Oscillatory (ENO) reconstruction. The basic idea of ENO reconstruction is to assign to a cell c_i^k which is in the smooth part of the sampled function, a stencil $\mathcal{S}_i^k = \{c_{i_0}^k, \ldots, c_{i_0+r-1}^k\}$ with $i_0 = i_0(i)$, which is likewise in the smooth part of the function (provided that this is possible ,i.e. that discontinuities are well separated and are far enough from the boundaries). This is done by choosing $\mathcal{S}_i^k$ to be the stencil for which the reconstruction polynomial $p_i^k(x; v^k)$ in (3.9) is the "smoothest" among all candidate-stencils, i.e. those of r consecutive cells of C^k (starting with $c_{i_0}^k$) which contain the cell c_i^k ,e.g. by taking $i_0(i)$ to be the index for which

$$\min_{i_0} \left| \frac{d^{r-1}}{dx^{r-1}} p_i^k(x; v^k) \right| \tag{6.1}$$

is attained among all candidate-stencils. This enables us to get a good approximation everywhere except in the cells which contain a discontinuity.

Next we show that cell-average discretization enables us to get a good approximation even in cells which contain discontinuity by using "subcell resolution" (see [H6]). Let $\mathcal{I}_k(x; F^k)$

denote the piecewise-polynomial interpolation of the primitive function in (3.9). Since it has formal order of accuracy $r+1$, we get in regions of smoothness of f that

$$\mathcal{I}_k(x;F^k) = F(x) + O((h_k)^{r+1}\|F^{(r+1)}\|) \tag{6.2a}$$

then

$$\begin{aligned}(\mathcal{R}_k v^k)(x) = \tfrac{d}{dx} I_k(x;F^k) &= \tfrac{d}{dx}F(x) + O((h_k)^r\|F^{(r+1)}\|) \\ &= f(x) + O((h_k)^r\|f^{(r)}\|).\end{aligned} \tag{6.2b}$$

Assume now that $f(x)$ has $(p-1)$ continuous derivatives and that $f^{(p)}(x)$ is discontinuous but bounded. It is clear from relations (6.2) that the maximal accuracy that can be achieved from either point-values or cell-averages is $O(h^p\|f^{(p)}\|)$: Using cell-averages we gain one order of smoothness in the primitive function (3.5a) but we lose it in the differentiation (3.5b). Consequently there is no advantage in using cell-averages rather than point-values of $f(x)$ for smooth data.

There is a significant advantage however in using cell-averages rather than pointvalues of f when $f(x)$ is discontinuous in a finite number of points . To see that let us assume that $f(x)$ is discontinuous at $x_d \in (x^k_{j-1}, x^k_j)$ and that in $[a,x_d)\cup(x_d,b]$, $0 \le a < x_d < b \le 1$, f has at least r continuous derivatives. Let $\mathcal{I}^L$ and $\mathcal{I}^R$ denote interpolation of either $f(x)$ or $F(x)$ at grid points in $[a,x_d)$ and $(x_d,b]$, respectively. We note that $F(x)$ is continuous in $[a,b]$, but has a discontinuous derivative at x_d. Consequently, if $F(x)$ is properly resolved on the k-th grid $\mathcal{I}^L(x;F^k)$ and $\mathcal{I}^R(x;F^k)$ will intersect at some point $\tilde{x}_d \in c^k_j$. From (6.2) we get that this intersection point is a good approximation to the location of the discontinuity within the cell c^k_j, i.e.

$$\tilde{x}_d - x_d = O((h_k)^r\|f^{(r)}\|)). \tag{6.3}$$

On the other hand, having knowledge of point-values $\{f(x^k_i)\}$ in $[a,b]$, there is nothing much we can say about the location of the discontinuity *within* the cell c^k_j.

We describe now how to use the subcell-resolution technique of [H6] in the prediction (3.8) in order to get an accurate prediction in a cell c^{k-1}_j which contains a discontinuity. Let us denote by $\tilde{F}^k_{2j-1}$ our approximation to $F(x^k_{2j-1})$, and let $\mathcal{I}^L(x;F^{k-1})$ and $\mathcal{I}^R(x;F^{k-1})$ denote the ENO interpolation in the neighboring cells c^{k-1}_{j-1} and c^{k-1}_{j+1} on the left and right, respectively, and define

$$D(x) = \mathcal{I}^R(x;F^{k-1}) - \mathcal{I}^L(x;F^{k-1}). \tag{6.4a}$$

Since $D(\tilde{x}_d) = 0$ we assume that

$$D(x^{k-1}_{j-1}) \cdot D(x^{k-1}_j) < 0. \tag{6.4b}$$

$\tilde{F}^k_{2j-1}$ is now computed as follows

$$\tilde{F}^k_{2j-1} = \begin{cases} \mathcal{I}^L(x^k_{2j-1}; F^{k-1}) & \text{if} \quad D(x^k_{2j-1}) \cdot D(x^{k-1}_j) \le 0 \\ \\ \mathcal{I}^R(x^k_{2j-1}; F^{k-1}) & \text{otherwise} \end{cases} \tag{6.5}$$

It is easy to see that if $f(x)$ is a piecewise-polynomial function

$$f(x) = \begin{cases} P_L(x) & a \le x < x_d \\ \\ P_R(x) & x_d < x \le b \end{cases} \tag{6.6a}$$

with

$$\deg(P_L) \le r-1, \quad \deg(P_R) \le r-1, \tag{6.6b}$$

then

$$\tilde{F}^k_{2j-1} = F(x^k_{2j-1}), \tag{6.6c}$$

i.e. the procedure (6.5) is exact . More generally, if $f(x)$ has p continuous derivatives to the left and the right of the discontinuity, we get that

$$\tilde{F}^k_{2j-1} = F(x^k_{2j-1}) + O((h_k)^{\bar{p}+1}\|f^{(\bar{p})}\|), \qquad \bar{p} = \min(p, r); \tag{6.7}$$

Remark 6.1. If we know that $f(x)$ has $q-1$ continuous derivatives and a discontinuity of the q-th derivative in x_d, $x^{k-1}_{j-1} < x_d < x^{k-1}_j$ we can extend the subcell resolution technique of (6.4) - (6.5) to this case as follows: $\frac{d^q}{dx^q} F(x)$ has a discontinuous first derivative at x_d. If it is sufficiently resolved on the grid, we expect $\frac{d^q}{dx^q} I_L(x; F^{k-1})$ and $\frac{d^q}{dx^q} I_k(x; F^{k-1})$ to intersect at $\tilde{x}_d$ in c^{k-1}_j,

$$\tilde{x}_d - x_d = O(h^{r-q}). \tag{6.8a}$$

It follows therefore that if we replace $D(x)$ in (6.4) by

$$D(x) = \frac{d^q}{dx^q} I^R(x; F^{k-1}) - \frac{d^q}{dx^q} I^L(x; F^{k-1}), \tag{6.8b}$$

we get a subcell-resolution technique which is exact for the corresponding piecewise-polynomial problem (6.6); this implies (6.7).

Remark 6.2. Extrapolating the analysis of the information contents in cell-averages vs. point-values, we get that weighted-averages with respect to the hat-function contain information that will enable us to obtain subcell resolution of δ-distributions; this may be useful for compression of digital images and propagation of singularities (see [ADH]).

7. Numerical experiments and conclusions.

In Figure 3 and Table 1 we present results of data compression for the MR schemes which are desribed in this paper. We generate the discrete data v^L for $J_L = 512$ by

$$v_i^L = f_1(\xi_i^L), \quad \xi_i^L = -1 + (i - \frac{1}{2})h_L, \quad h_L = 2/J_L, \quad i = 1, \ldots, J_L, \tag{7.1}$$

where

$$f_1(x) = \begin{cases} -xsin(\frac{3\pi}{2}x^2) & -1 < x \leq -\frac{1}{3} \\ |sin(2\pi x)| & |x| < \frac{1}{3} \\ 2x - 1 - sin(3\pi x)/6 & \frac{1}{3} \leq x < 1 \end{cases} . \tag{7.2}$$

This input data is displayed in Figure 3a by drawing a circle around the points (ξ_i^L, v_i^L) for $i = 1, \ldots, J_L$.

The MR schemes are used with 6 levels of resolution, i.e. $L = 6$ and $J_0 = 8$. The scale coefficients are truncated by (5.1a) with $\varepsilon = 10^{-3}$ and $q = \frac{1}{2}$ in (5.4a). The result of this truncation is displayed in the $x - k$ plane by drawing a circle around (ξ_j^k, k), where $\xi_j^k = -1 + (j - \frac{1}{2})h_k, \quad h_k = 2^{L-k}h_L$ for each d_j^k which is above the tolerance ε_k.

In Figure 3b we show the data compression of the MR scheme which is based on Daubechies' orthonormal wavelets with $r = 5$, i.e. five vanishing moments.

In Figure 3c we show the results corresponding to the biorthogonal wavelets of [CDF] with $r = 5$. In both Figures 3b and 3c we used periodic extension at the boundaries.

In Figure 3d we show the results of the piecewise-polynomial reconsrtuction (3.11)–(3.12) with $r = 5$, where we used one-sided stencils near the boundaries.

In Figures 3e and 3f we used the error control algorithm (5.10) with the same choice of $\{\varepsilon_k\}$ as above. In Figure 3e we show the results of the ENO reconstruction with $r = 6$. The particular technique that we used here is described in [HEOC]; it is the one-dimensional version of the "crystal-growth" algorithm in [HC].

In Figure 3f we append this ENO reconstruction with subcell resolution (6.5). The subcell resolution is applied to cells c_j^k for which $S_{j-1}^k \cap S_{j+1}^k = \emptyset$, i.e. where the selection of ENO stencils "shies away" from c_j^k.

The corresponding rates of compression and compression errors are listed in Table 1 . We would like to make the following observations regarding these numerical results:

(1) The compression error is about the same for all the MR schemes in this experiment.

(2) The "smooth part" of the data is resolved to the prescribed tolerance at level 3; thus the significant scale-coefficients of levels 4 to 6 are the "signatures" of the various prediction operators of the two discontinuities in the function (7.2) and the discontinuity in

the derivative. In the case of wavelets where we used periodic extension, we also get the "signature" of the discontinuity which is so introduced at the boundaries.

(3) Biorthogonal wavelets perform better than orthonormal wavelets; their "signature" is narrower because of the smaller stencil and it is symmetric. We remark that the lack of symmetry in the coefficients of Daubechies' orthonormal wavelets increses the number of operations in various applications (see [HY], [H7] and [GKK]).

(4) As to be expected, the piecewise-polynomial reconstruction in Figure 3d is identical to the biorthogonal wavelets in the interior; the only improvement is the removal of the "signature" of the boundary discontinuity which is accomplished by using one-sided stencils near the boundaries.

(5) Since the main difference between the various MR schemes in this experiment is in their "signature" of discontinuities, there is a considerable improvement in using the ENO reconstruction, because then the signature is typically 1 point. This signature can be completely eliminated at times by using subcell resolution (after all the only information in a discontinuity is its location). Observe that the particular version of ENO reconstruction that is used here misses a bit at the discontinuous derivative which is located at $x = 0$. This can be fixed by using special procedures which are described in [Do] ; see also Remark 6.1.

Next we repeat the numerical experiments in Figure 3 for two-dimensional data corresponding to a tensor-product of the grid in (7.1). We generate the discrete data $\bar{f}^L$ by

$$\bar{f}^L_{i,j} = f_2(\xi^L_i, \xi^L_j), \qquad 1 \le i, j \le 512 \tag{7.3}$$

where the function $f_2(x, y)$, which is defined in $[-1, 1]^2$, is the following two-dimensional "variation" on $f_1(x)$ in (7.2) :

$$f_2(x,y) = \begin{cases} u_{\sqrt{\pi/2}}(x,y) & \text{if } \; x \le \frac{1}{2}\cos(\pi y) \\ \\ u_{-\sqrt{\pi/2}}(x,y) + \cos(2\pi y) & \text{if } \; x > \frac{1}{2}\cos(\pi y), \end{cases} \tag{7.4a}$$

where u_ϕ is

$$u_\phi(x,y) = \begin{cases} -r\sin(\frac{\pi}{2}r^2) & \text{if } r \le -\frac{1}{3} \\ \\ |\sin(2\pi r)| & \text{if } |r| < \frac{1}{3} \\ \\ 2r - 1 + \frac{1}{6}\sin(3\pi r) & \text{if } r \ge \frac{1}{3} \end{cases} \tag{7.4b}$$

and

$$r = x + \tan(\phi) y. \tag{7.4c}$$

We display this function in terms of its isolines in Figure 4a.

The MR schemes are used with $L = 6$ levels of resolution, i.e. the coarsest level is an 8×8 matrix.

In Figure 4b we show the data compression of the MR scheme which is based on Daubechies' orthonormal wavelets with $r = 3$, i.e. three vanishing moments.

In Figure 4c we show the results corresponding to the biorthogonal wavelets of [CDF] with $r = 3$. In both Figures 4b and 4c we used periodic extension at the boundaries.

In Figure 4d we show the results of the piecewise-polynomial reconsrtuction (3.11)–(3.12) with $r = 3$, where we used one-sided stencils near the boundaries.

The results in Figures 4b – 4e where obtained with a tensor-product extension of the corresponding one-dimensional MR scheme which is described in appendix B. As in the one-dimensional case the scale coefficients are truncated by (5.1a) with $\varepsilon = 10^{-3}$ and $q = \frac{1}{2}$ in (5.4a). The result of this truncation is displayed in the form which is described in (B.10).

In Figures 4e and 4f we used a tensor-product extension of the error control algorithm (5.10) with the same choice of $\{\varepsilon_k\}$ as above; this algorithm is described in appendix C.

In Figure 4e we show the results of the ENO reconstruction with $r = 6$. In Figure 4f we append this ENO reconstruction with subcell resolution (6.5) which is applied in its one-dimensional form.

The corresponding rates of compression and compression errors are listed in Table 2 . We would like to make the following observations regarding these numerical results:

(1) Comparing these results to their one-dimensional counterpart we see that the compression error is somewhat larger, but of the same order. The compression ratio is larger – this is of course data dependent, but nevertheless "typical".

(2) The order of "efficiency" between the various MR schemes remains as in the one-dimensional case, and the differences are somewhat magnified. We used here $r = 3$ for the non-adaptive schemes because it gives better results than $r = 5$; e.g. using $r = 5$ for Daubechies' orthonormal wavelets yields a compression ratio of 7.79 with maximal error of 1.226×10^{-3} .

(3) The addition of subcell resolution to the ENO reconstruction improves the compression, but only slightly. We can get a more significant improvement if we use a genuinely two-dimensional extension of the algorithm in (6.5).

We would like to remark that one can use the error control algorithm also with the non-adaptive MR schemes in Figures 4b – 4d , in which case we always get an improvement.

In Figure 5 we demonstrate the geometrical flexibility of using cell-average discretization by showing the results of data compression in a domain which is carved in the form of a smiling face from $[-1, 1]^2$. The finest mesh has 12,962 cells and the input data is the cell-averages of $f_2(x, y)$ in (7.4); this data is displayed in terms of isolines as the left plot in Figure

5b; to the right of this plot we show the result of its decimation on the next two coarser levels of resolution, which are obtained by successive agglomeration. These levels contain 2,122 and 369 cells, respectively; the cells of the coarsest level are shown in Figure 5a. In this calculation we used a piecewise-polynomial reconstruction with $r = 3$, and truncated scale-coefficients with $\varepsilon = 10^{-2}$. In Figure 5c we display this truncation by marking the coarser-level cells in which one of the prediction errors in (2.3) is larger than the tolerance. The compression ratio in this experiment is 2.65 , the maximal error is 2.15×10^{-2} and the L^1-error is 6.48×10^{-4}. In [AH], from which these results are taken, we show that the performance of this MR scheme is comparable to corresponding tensor-product schemes in similar circumstances. We refer the reader to [AH] for more details.

In building the case to promote the use of cell-average discretization we have not yet discussed the issue of suitability to applications. As we have mentioned earlier, for smooth input data the performance of all MR schemes is about the same, and then pointwise discretization is usually the best choice because of its simplicity. However, when the input data is only piecewise-smooth there is a clear advantage to using cell-average discretization – this is true in the numerical solution of hyperbolic conservation laws where the solution typically contains shock waves (see [H4-5]), and in the solution of integral equations where the kernel is usually integrably singular (see [BCR],[HY],[H7] and[GKK]). Another important feature of cell-average discretization is that if the input data contains noise or local high-frequency components, then these are averaged out in the coarser levels. This enables us to to develop techniques to reduce noise (see [Mi]) and to design special techniques to handle piecewise-smooth data which carries some local high frquency components (see [DH]).

These characteristics of cell-average discretization make it ideal for the purpose of image compression, where the coarser levels of discretization behave like those of piecewise-smooth functions. Furthermore, the strategy of truncation in (5.4a) and (5.17a) dictates increase in accuracy as we go to larger cells. This is suitable to the averaging properties of the human eye and makes the decompressed image look "pleasant"; in some respects this works like cosmetic retouching, e.g. in compressing an image of a face it may eliminate freckles, but at the same time keep the larger features sharp.

Acknowledgements.

Many thanks to my collaborators Remi Abgrall, Paco Arandiga, Barna Bihari, Rosa Donat and Itai Yad-Shalom.

This research was supported in UCLA by Grants ONR-N00014-92-J-1890 and NSF-DMS91-03104.

Appendix A. Stability analysis and existence of MR bases.

We assume that sequence $\{(\mathcal{R}_k\mathcal{D}_k)\}_{k=0}^{\infty}$,

$$(\mathcal{R}_k\mathcal{D}_k) : \mathcal{F} \to \mathcal{F} \ , \tag{A.1a}$$

is a sequence of (discrete) approximation in the Banach space $\mathcal{F}$, i.e. that for any $f \in \mathcal{F}$

(i)

$$\|\mathcal{R}_k\mathcal{D}_k f\| \le C_A^k \|f\| \tag{A.1b}$$

(ii)

$$\|\mathcal{R}_k\mathcal{D}_k f - f\| \to 0 \quad as \quad k \to \infty \ . \tag{A.1c}$$

Using the principle of uniform boundedness we conclude that there exists a constant C_A such that for all k

$$C_A^k \le C_A. \tag{A.1d}$$

.

If (A.1a) is a *nested* sequence of discretization, we get that the direct MR transform (2.11) is stable with respect to perturbations in the input data v^L, and that

$$\begin{aligned} &\langle \delta(d^k) \rangle_k = |\delta(e^k)|_k \le C_A(1 + C_A)|\delta(v^L)|_L \\ &|\delta(v^0)|_0 \le C_A|\delta(v^L)|_L \ . \end{aligned} \tag{A.2a}$$

where $\delta(\cdot)$ denotes the perturbation, and the discrete norms above are defined as follows:

$$|v^k|_k = \|\mathcal{R}_k v^k\| \ ; \tag{A.2b}$$

$$\langle d^k \rangle_k = |E_k d^k|_k \ . \tag{A.2c}$$

The "natural" function space for cell-average discretization is $\mathcal{F} = L_1(\Omega)$, and there (A.1b)-(A.1c) take the following form:

$$\int_\Omega |(\mathcal{R}_k\mathcal{D}_k f)(x)|dx \le C_A^k \int_\Omega |f(x)|dx, \tag{A.3a}$$

$$\lim_{k\to\infty} \int_\Omega |(\mathcal{R}_k\mathcal{D}_k f)(x) - f(x)|dx = 0. \tag{A.3b}$$

In [H3] we investigate the stability of the inverse MR transform (2.12) and the related question of existence of MR bases for mappings in $\mathcal{F}$. In the following we present a summary of the results:

Case 1.

We assume that $\{(\mathcal{R}_k\mathcal{D}_k)\}_{k=0}^{\infty}$ is an hierarchic sequence of approximation, i.e. in addition to (A.1) it satisfies the following for all $k > 0$

$$(\mathcal{R}_k\mathcal{D}_k)\mathcal{R}_{k-1} = \mathcal{R}_{k-1}; \tag{A.4a}$$

note that another way to express (A.4a) is

$$\mathcal{R}_k P_{k-1}^k = \mathcal{R}_{k-1}. \tag{A.4b}$$

In this case we show in [H3] that the inverse MR transform is stable with respect to perturbations

$$|\delta(v^L)|_L \leq |\delta(v^0)|_0 + \sum_{k=1}^{L} \langle \delta(d^k) \rangle_k \ , \tag{A.5}$$

and that for for any $f \in \mathcal{F}$

$$f = \mathcal{R}_0\mathcal{D}_0 f + \sum_{k=1}^{\infty} \sum_{j=1}^{J_k - J_{k-1}} d_j^k \psi_j^k \ , \tag{A.6a}$$

where the coefficients $d_j^k = d_j^k(f)$ are the scale coefficients in (1.10)

$$d^k(f) = G_k e^k(f) = G_k\mathcal{D}_k(I - \mathcal{R}_{k-1}\mathcal{D}_{k-1})f. \tag{A.6b}$$

The functions of the MR basis of $\mathcal{F}$ are obtained by reconstruction of the basis (1.8b) of the null space of the decimation operator

$$\psi_j^k = \mathcal{R}_k \mu_j^k. \tag{A.6c}$$

Case 2.

We assume that $\{\mathcal{R}_k\mathcal{D}_k\}_{k=0}^{\infty}$ is a sequence of approximation which is σ-contractive, i.e. that there exist $0 < q < 1$ and a convergent series $\sum_{\ell=0}^{\infty} \Delta_\ell < \infty$ of positive numbers such that for all $k \geq 0$ and any $f \in \mathcal{F}$

$$\sigma_{k+1}(\mathcal{R}_k\mathcal{D}_k f) \leq q \cdot (1 + \Delta_k) \cdot \sigma_k(f) \ , \tag{A.7a}$$

where

$$\sigma_\ell(f) =: \|(I - \mathcal{R}_{\ell+1}\mathcal{D}_{\ell+1})\mathcal{R}_\ell\mathcal{D}_\ell f\|. \tag{A.7b}$$

In this case we show in [H3] that the following limit exists

$$\mathcal{R}_k^H\mathcal{D}_k f =: \lim_{L\to\infty} \prod_{\ell=k}^{L} (\mathcal{R}_\ell\mathcal{D}_\ell) \cdot f, \tag{A.8}$$

that $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ is a hierarchic sequence of approximation, and that its MR scheme is exactly the same as that of $\{(\mathcal{R}_k \mathcal{D}_k)\}$. Hence we can use the results of case 1 to conclude that the MR scheme is stable with respect to perturbations, and that the compression error is bounded by (A.2a) where now the discrete norms in (A.2b)-(A.2c) are defined with the hierarchic form $\mathcal{R}_k^H$. Furthermore, any $f \in \mathcal{F}$ has the expansion (A.6a), where $\mathcal{R}_0$ is replaced by $\mathcal{R}_0^H$ and

$$\psi_j^k =: \lim_{L\to\infty} \prod_{\ell=k}^{L} (\mathcal{R}_\ell \mathcal{D}_\ell) \cdot (\mathcal{R}_k \mu_j^k). \tag{A.9}$$

Remark A.1.

Let $\{\eta_i^k\}_{i=1}^{J_k}$ denote any basis of the linear space V^k in (1.1)

$$V^k = \ span\{\eta_i^k\}, \tag{A.10a}$$

and denote

$$\varphi_i^k =: \mathcal{R}_k^H \eta_i^k = \lim_{L\to\infty} \prod_{\ell=k}^{L} (\mathcal{R}_\ell \mathcal{D}_\ell) \cdot (\mathcal{R}_k \eta_i^k). \tag{A.10b}$$

In [H2-3] we show that

$$\varphi_\ell^{k-1} = \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \varphi_i^k \tag{A.10c}$$

where $\hat{P}_{k-1}^k$ is the matrix representation of the *original* prediction operator (1.6a). The sequence $\{\varphi_i^k\}$ is related to $\{\psi_j^k\}$ in (A.9) by

$$\psi_j^k = \sum_i (\hat{E}_k)_{i,j} \varphi_i^k, \tag{A.11}$$

where $\hat{E}_k$ is the matrix representation of the operator E^k in (1.9a).

Note that when $\hat{P}_{k-1}^k$ is the Töplitz-like matrix in (4.6) , then all $\varphi_i^k(x)$ are generated from a single function $\varphi(x)$ which is the solution of the dilation equation (4.4b). From this point of view $\{\psi_j^k\}$ in (A.9) can be thought of as "generalized wavelets".

Remark A.2.

Observe that the functionals σ_ℓ vanish if the sequence is hierarchic, and thus their size measures the deviation of the sequence from being hierarchic. We remark that wherever condition (A.7) has been verified, this was done with "hard" analysis; nevertheless, the design problem seems to be "soft". Although there are known examples of divergence of the above limit, most "natural" approximations give rise to stable MR schemes (as is evident from the experiments in this paper and others not reported here).

Appendix B. Matrix form of tensor-product MR Schemes.

In this appendix we show how to use tensor product extension of one-dimensional operators in [0,1] in order to obtain MR schemes for data which is obtained from absolutely integrable functions f

$$f : [0,1] \times [0,1] \longrightarrow I\!R$$

which are discretized on the tensor-product mesh

$$C^k = \{c_i^k \times c_j^k\}_{i,j=1}^{N_k}$$

by

$$\bar{f}^k =: \{\bar{f}_{i,j}^k\}_{i,j=1}^{N_k} \ , \ \bar{f}_{i,j}^k = \frac{1}{|c_i^k| \cdot |c_j^k|} \int_{c_i^k} \int_{c_j^k} f(x_1, x_2) dx_1 dx_2$$

where $\{c_i^k\}_{i=1}^{N_k}$ are the one-dimensional cells in (3.2a); note that here we replaced J_k by N_k, since we reserve J_k for the dimension of the vector space V^k in (1.1), i.e. $J_k = (N_k)^2$. Although this case is covered by the general framework , it is more convenient to describe the tensor-product extension of the one-dimensional MR scheme by using matrix notation. Therefore we represent the two-dimensional array $\{\bar{f}_{i,j}^k\}_{i,j=1}^{N_k}$ as the $N_k \times N_k$ matrix A^k and describe the tensor-product MR scheme for the input-matrix A^L.

Let us denote the matrix representation of the various one-dimensional operators by

$$\begin{aligned} & D_k^{k-1} \longrightarrow (D)_{N_{k-1} \times N_k} \\ & P_{k-1}^k \longrightarrow (P)_{N_k \times N_{k-1}} \\ & G_k^D = G_k(I_k - P_{k-1}^k D_k^{k-1}) \longrightarrow (G^D)_{N_{k-1} \times N_k} = G(I - PD) \\ & E_k \longrightarrow (E)_{N_k \times N_{k-1}}. \end{aligned} \tag{B.1}$$

These matrix representations are obtained by taking v^k and d^k to be column-vectors, e.g.

$$v_i^{k-1} = (D_k^{k-1} v^k)_i =: \sum_{j=1}^{N_k} D_{ij} v_j^k, \qquad 1 \le i \le N_{k-1};$$

for simplicity we drop the index k.

Starting with A^L we decimate to get

$$A^{k-1} = D A^k D^*, \qquad k = L, \ldots, 1; \tag{B.2}$$

here $(\cdot)^*$ denotes the transpose. Given A^{k-1} we get an approximation to A^k by

$$A^k \approx P A^{k-1} P^*$$

and observe that the prediciton error *matrix* e^k

$$e^k = A^k - PA^{k-1}P^* \tag{B.3a}$$

satisfies

$$De^kD^* = 0. \tag{B.3b}$$

The dimension of the null space of the decimation operators here is

$$J_k - J_{k-1} = (2N_{k-1})^2 - (N_{k-1})^2 = 3(N_{k-1})^2$$

and we store the scale-coefficients d^k in three $N_{k-1} \times N_{k-1}$ matrices $\Delta_1^k, \Delta_2^k, \Delta_3^k$.

Using the matrix identity

$$I = PD + EG^D, \qquad G^D = G(I - PD), \tag{B.4}$$

we show that if we take A^k and A^{k-1} from the sequence (B.2) and define

$$\Delta_1^k = G^D A^k (G^D)^*, \qquad \Delta_2^k = DA^k (G^D)^*, \qquad \Delta_3^k = G^D A^k D^* \tag{B.5}$$

then A^k can be recovered from A^{k-1} and the above by

$$A^k = PA^{k-1}P^* + E\Delta_1^k E^* + P\Delta_2^k E^* + E\Delta_3^k P^*. \tag{B.6}$$

This follows immediately from the identity

$$\begin{aligned} A^k &= (PD + EG^D)A^k(PD + EG^D)^* = P(DA^kD^*)P^* \\ &+ E[G^DA^k(G^D)^*]E^* + P[DA^k(G^D)^*]E^* + E(G^DA^kD^*)P^*. \end{aligned}$$

We conclude from (B.5)-(B.6) that A^L has a multiresolution representation $\hat{A}_M$,

$$A^L \overset{1:1}{\longleftrightarrow} \hat{A}_M =: \left(\left\{ \Delta_m^L \right\}_{m=1}^3, \ldots, \left\{ \Delta_m^1 \right\}_{m=1}^3, A^0 \right), \tag{B.7}$$

where the direct MR transform is given by

$$\begin{cases} DO \quad k = L, \ldots, 1 \\ \\ A^{k-1} = DA^kD^* \\ \\ \Delta_1^k = G^DA^k(G^D)^*, \quad \Delta_2^k = DA^k(G^D)^*, \quad \Delta_3^k = G^DA^kD^*, \end{cases} \tag{B.8}$$

and the inverse MR transform is

$$\begin{cases} DO \quad k = 1, \ldots, L \\ \\ A^k = PA^{k-1}P^* + E\Delta_1^k E^* + P\Delta_2^k E^* + E\Delta_3^k P^*. \end{cases} \tag{B.9}$$

We display the results of data compression by writing $\hat{A}_M$ in (B.7) as the matrix

$$\widehat{A}_M = \left[\begin{array}{cc|cc} \Delta_1^L & \Delta_2^L & & \\ & & & \\ \hline & & & \\ \Delta_3^L & & \ddots & \\ & & & \begin{array}{cc} \Delta_1^1 & \Delta_2^1 \\ \Delta_3^1 & A^0 \end{array} \end{array}\right] \tag{B.10}$$

and mark each entry which is larger in absolute value than the prescribed tolerance by a black dot.

Appendix C. Error-control algorithms for tensor-product grids.

In this appendix we describe our error-control algorithm for a rectangular grid. For this purpose let us introduce two FORTRAN-like subroutines:

(1) Let DEC1D(U, J, W) describe the one-dimensional decimation in (5.10a), i.e. its input is an array U of size $2J$, and its output is the array W of size J which is computed as follows:

$$\begin{cases} DO\ i = 1, \ldots, J \\ \\ W_i = \frac{1}{2}(U_{2i-1} + U_{2i}) \end{cases} \tag{C.1}$$

(2) Let ERRCON1D$(\tilde{U}, J, V, D, W, \varepsilon)$ describe the one-dimensional error-control in (5.10c), i.e. it operates on the array $\tilde{U}$ of size J and its output is D of size J (the scale-coefficients) and W of size $2J$; here V is the "control" which is also of size $2J$. We denote by P the one-dimensional prediction operator.

$$\begin{cases} DO\ j = 1, \ldots, J \\ \\ v^P = (P \cdot \tilde{U})_{2j-1} \\ \\ D_j = tr(V_{2j-1} - v^P - (V_j - \tilde{U}_j); \varepsilon) \\ \\ W_{2j-1} = v^P + D_j \\ \\ W_{2j} = 2\tilde{U}_j - W_{2j-1}. \end{cases} \tag{C.2}$$

Next we describe the encoding procedure for a $J_L \times J_L$ array $\bar{f}^L = \{\bar{f}^L_{i,j}\}$:

(i) Apply decimation to the input $\bar{f}^L$

$$\begin{cases} DO\ k = L, \ldots, 1 \\ \\ \quad DO\ i = 1, \ldots, J_k \\ \quad \text{CALL DEC1D}(\bar{f}^k_{i,\cdot}, J_{k-1}, \bar{f}^{k-\frac{1}{2}}_{i,\cdot}) \\ \quad \text{END DO} \\ \\ \quad DO\ j = 1, \ldots, J_{k-1} \\ \quad \text{CALL DEC1D}(\bar{f}^{k-\frac{1}{2}}_{\cdot,j}, J_{k-1}, \bar{f}^{k-1}_{\cdot,j}) \\ \quad \text{END DO} \\ \\ \text{END DO} \end{cases} \tag{C.3a}$$

(ii) Set

$$\tilde{f}^0 = \bar{f}^0 \tag{C.3b}$$

(iii) Calculate

$$\begin{cases} DO\ k = L, \ldots, 1 \\ \quad DO\ i = 1, \ldots, J_{k-1} \\ \quad \text{CALL ERRCON1D}(\tilde{f}^{k-1}_{i,\cdot}, J_{k-1}, \bar{f}^{k-\frac{1}{2}}_{i,\cdot}, (\Delta^k_1)_{i,\cdot}, \tilde{f}^{k-\frac{1}{2}}_{i,\cdot}, \varepsilon_{k-\frac{1}{2}}) \\ \quad \text{END DO} \\ \\ \quad DO\ j = 1, \ldots, J_{k-1} \\ \quad \text{CALL ERRCON1D}(\tilde{f}^{k-\frac{1}{2}}_{\cdot,2j-1}, J_{k-1}, \bar{f}^{k}_{\cdot,2j-1}, (\Delta^k_2)_{\cdot,j}, \tilde{f}^{k}_{\cdot,2j-1}, \varepsilon_k) \\ \quad \text{CALL ERRCON1D}(\tilde{f}^{k-\frac{1}{2}}_{\cdot,2j}, J_{k-1}, \bar{f}^{k}_{\cdot,2j}, (\Delta^k_3)_{\cdot,j}, \tilde{f}^{k}_{\cdot,2j}, \varepsilon_k) \\ \quad \text{END DO} \\ \\ \text{END DO} \end{cases} \tag{C.3c}$$

Observe that as in appendix B, the scale-coefficients are stored in three $J_{k-1} \times J_{k-1}$ matrices $\Delta^k_1, \Delta^k_2, \Delta^k_3$. In the numerical experiments of this paper we used

$$\varepsilon_{k-\frac{1}{2}} = \varepsilon_k = 2^{k-L-1}\varepsilon. \tag{C.4}$$

The significant scale-coefficients are displayed in Figures 4e – 4f as in (B.10).

To describe the decoding algorithm we introduce the subroutine RECOVER1D$(\tilde{U}, J, D, W)$ for the one-dimensional algorithm (5.9). It operates on the array $\tilde{U}$ of size J and its output is W of size $2J$; here D is the array of scale-coefficients, which is of size J.Again we denote by P the one-dimensional prediction operator. The operation of RECOVER1D$(\tilde{U}, J, D, W)$ is as follows:

$$\begin{cases} DO\ j = 1, \ldots, J \\ \\ v^P = (P \cdot \tilde{U})_{2j-1} \\ \\ W_{2j-1} = v^P + D_j \\ \\ W_{2j} = 2\tilde{U}_j - W_{2j-1}. \end{cases} \tag{C.5}$$

To apply the decoding we

(i) Set

$$\tilde{f}^0 = \bar{f}^0 \tag{C.6a}$$

(ii) Calculate

$$
\left\{
\begin{array}{l}
DO\ k = L, \ldots, 1 \\
\\
\quad DO\ i = 1, \ldots, J_{k-1} \\
\quad \text{CALL RECOVER1D}(\tilde{f}^{k-1}_{i,\,\cdot}, J_{k-1}, (\Delta^k_1)_{i,\,\cdot}, \tilde{f}^{k-\frac{1}{2}}_{i,\,\cdot}) \\
\quad \text{END DO} \\
\\
\quad DO\ j = 1, \ldots, J_{k-1} \\
\quad \text{CALL RECOVER1D}(\tilde{f}^{k-\frac{1}{2}}_{\cdot\,,2j-1}, J_{k-1}, (\Delta^k_2)_{\cdot\,,j}, \tilde{f}^{k}_{\cdot\,,2j-1}) \\
\quad \text{CALL RECOVER1D}(\tilde{f}^{k-\frac{1}{2}}_{\cdot\,,2j}, J_{k-1}, (\Delta^k_3)_{\cdot\,,j}, \tilde{f}^{k}_{\cdot\,,2j}) \\
\quad \text{END DO} \\
\\
\text{END DO}
\end{array}
\right. \qquad (C.6b)
$$

References

[Ab] R. Abgrall, "Design of an essentially nonoscillatory reconstruction procedure on finite element type meshes", ICASE report 91-84, (Dec. 1991), in revised form INRIA report No. 1592 (February 1992), submitted to Math. of Comp.

[ADH] F. Arandiga, R. Donat and A. Harten, "Multiresolution based on weighted averages of the hat function", UCLA CAM Report 93-34, September 1993.

[AH] R. Abgrall and A. Harten, "Multiresolution representation in unstructured meshes. I. Preliminary report", UCLA CAM Report 94-20, July 1994.

[BCR] G. Beylkin, R. Coifman and V. Rokhlin, "Fast wavelet tranform and numerical algorithms. I," Comm. Pure Appl. Math. V. 44, pp. 141-183, 1991.

[CDF] A. Cohen, I. Daubechies and J.-C. Feauveau, "Biorthogonal bases of compactly suported wavelets", Comm. Pure Appl. Math., Vol. 45, pp. 485-560, 1992.

[Da] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Comm. Pure Appl. Math. V. 41, pp. 909-996, 1988.

[Do] R. Donat "Studies on error propagation for certain nonlinear approximations to hyperbolic equations: Discontinuities in derivatives," SIAM Jour. Num. Anal., V. 31, pp. 655-679, 1994.

[DH] R. Donat and A. Harten, "Data compression algorithms for locally oscillatory data: A preliminary report", UCLA Preprint, October 1992.

[GKK] K. Guru Prasad, D.E. Keyes and J.H. Kane, "Generalized wavelet bases for boundary element matrices", Preprint, 1994.

[H1] A. Harten, "Discrete multiresolution analysis and generalized wavelets," J. Appl. Num. Math. V. 12, pp. 153-193, 1993; also UCLA CAM Report 92-08, February 1992.

[H2] A. Harten, "Multiresolution representation of data. I. Preliminary Report," UCLA CAM Report No. 93-13, June 1993.

[H3] A. Harten, "Multiresolution representation of data. II. General Framework," UCLA CAM Report No. 94-10, April 1994; also Technical Report 3-94, Dept. of Applied Mathematics, Tel Aviv University, April 1994.

[H4] A. Harten, "Multiresolution algorithms for the numerical solution of hyperbolic conservation laws," UCLA CAM Report 93-03, March 1993; to appear in Comm. Pure Appl. Math.

[H5] A. Harten, "Adaptive multiresolution schemes for shock computations," UCLA CAM Report 93-06, April 1993; to appear in Jour. Comput. Phys.

[H6] A. Harten, "ENO schemes with subcell resolution", J. Comp. Phys., Vol. 83, pp. 148-184, 1989; also ICASE Report No. 87-56.

[H7] A. Harten, "Multiresolution representation and numerical algorithms: A brief review", UCLA CAM Report 94-12, June 1994.

[HC] A. Harten and S. R. Chakravarthy, "Multi-dimensional ENO schemes for general geometries", ICASE Report 91-76, September 1991; also submitted to Jour. Comput. Phys.

[HY] A. Harten and I. Yad-Shalom, "Fast multiresolution algorithms for matrix-vector multiplication," ICASE Report 92-55, October 1992; to appear in SIAM Jour. Num. Anal.

[HEOC] A. Harten, B. Engquist, S. Osher, and S. Chakravarthy, "Uniformly high order accurate essentially non-oscillatory schemes, III," J. Comput. Phys. V. 71, pp. 231-303, 1987.

[Ma] S. Mallat, "Multiresolution approximation and wavelets orthonormal bases of $L^2(R)$," Trans. Amer. Math. Soc. V. 315, pp. 69-87, 1989.

[Me] Y. Meyer, "Ondelettes et Opérateurs," Hermann, 1990.

[Mi] D. Mizrachi, "Removing noise from discontinuous data", M. Sc. Thesis (1991) Tel-Aviv University.

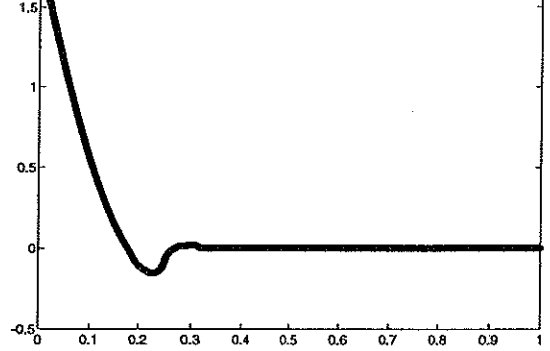
Table 1: Data compression of one-dimensional input.

Method	Order r	Compression	$L_\infty - error$	$L_1 - error$	$L_2 - error$
Orthonormal wavelets	5	4.163	3.524×10^{-4}	1.081×10^{-5}	3.574×10^{-5}
Biorthogonal wavelets	5	4.923	7.659×10^{-5}	1.114×10^{-5}	2.141×10^{-5}
Piecewise-polynomial	5	5.626	2.967×10^{-4}	1.455×10^{-5}	2.824×10^{-5}
ENO reconstruction	6	9.143	2.525×10^{-4}	2.002×10^{-5}	3.514×10^{-5}
ENO +Subcell resol.	6	10.039	2.525×10^{-4}	2.088×10^{-5}	3.611×10^{-5}

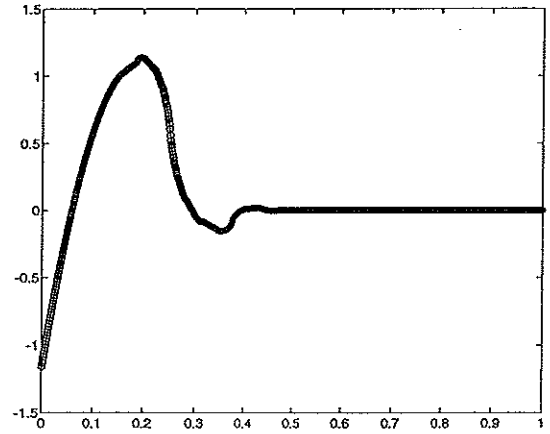
Table 2: Data compression of two-dimensional input.

Method	Order r	Compression	$L_\infty - error$	$L_1 - error$	$L_2 - error$
Orthonormal wavelets	3	9.46	1.264×10^{-3}	5.882×10^{-5}	8.716×10^{-5}
Biorthogonal wavelets	3	10.67	9.717×10^{-4}	5.889×10^{-5}	8.225×10^{-5}
Piecewise-polynomial	3	12.09	9.717×10^{-4}	5.933×10^{-5}	8.255×10^{-5}
ENO reconstruction	6	37.03	3.542×10^{-3}	3.122×10^{-5}	6.190×10^{-5}
ENO +Subcell resol.	6	41.83	1.042×10^{-3}	3.043×10^{-5}	5.794×10^{-5}

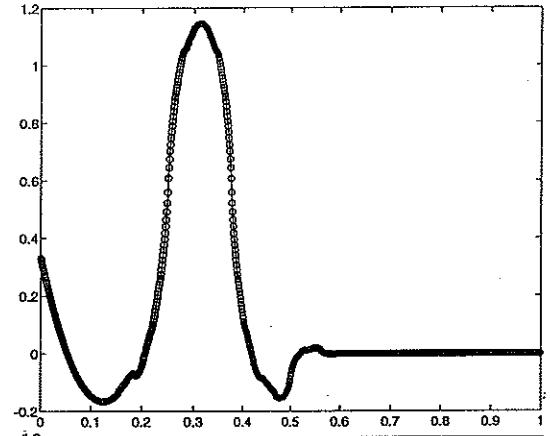
(a) φ_1^0



(b) φ_2^0



(c) φ_3^0



(d) φ_4^0

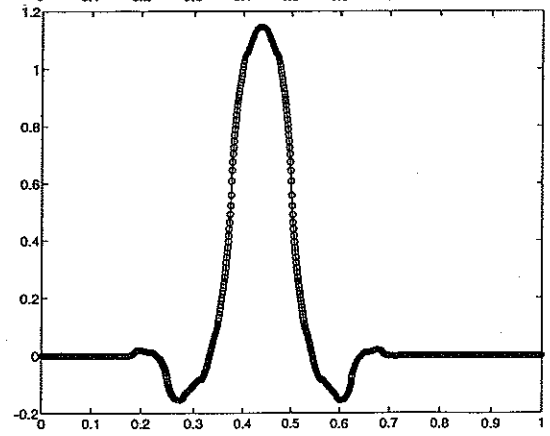
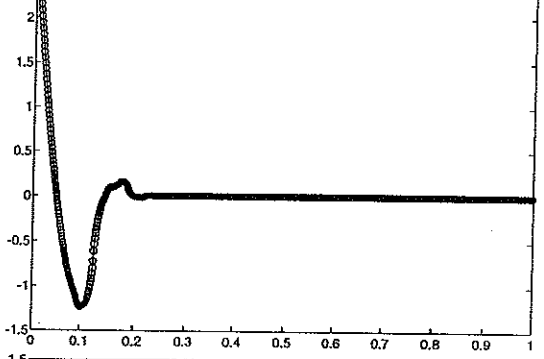
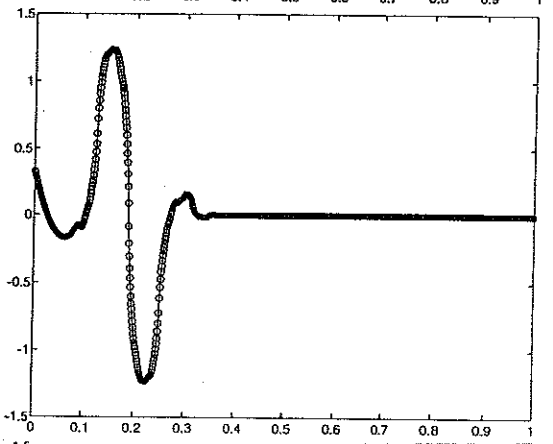


Figure 1. Subdivision limit for $\{\varphi_i^0\}_{i=1}^4$

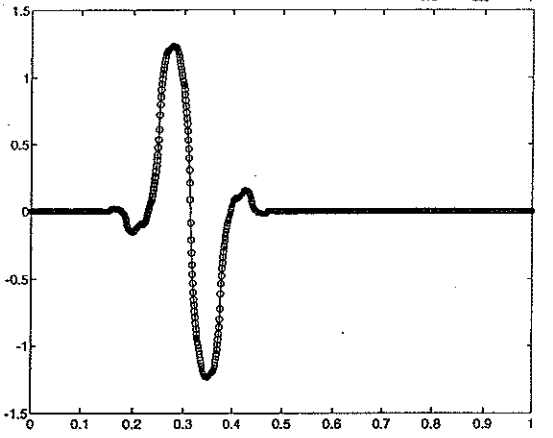
(a) ψ_1^1



(b) ψ_2^1



(c) ψ_3^1



(d) ψ_4^1

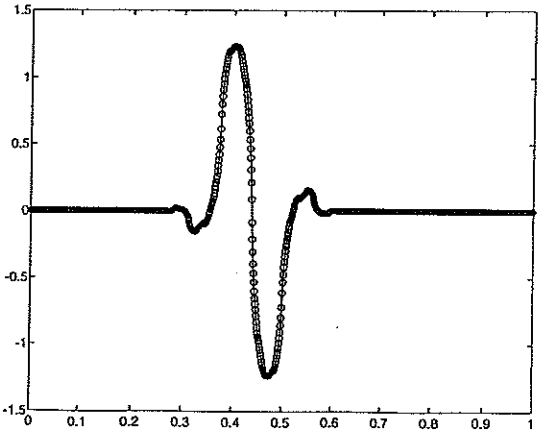
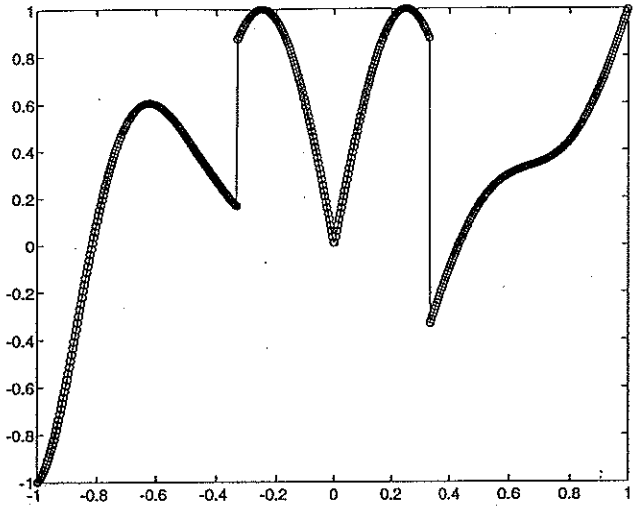
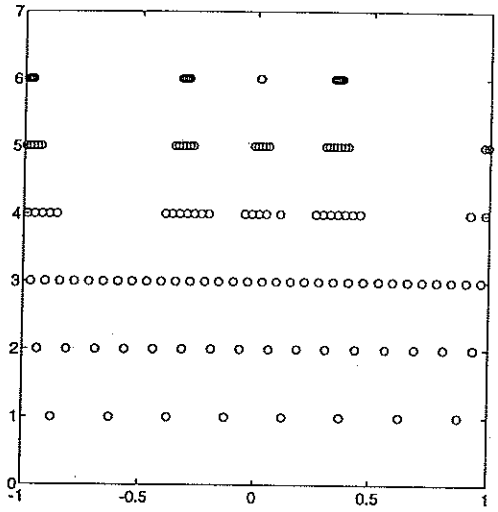


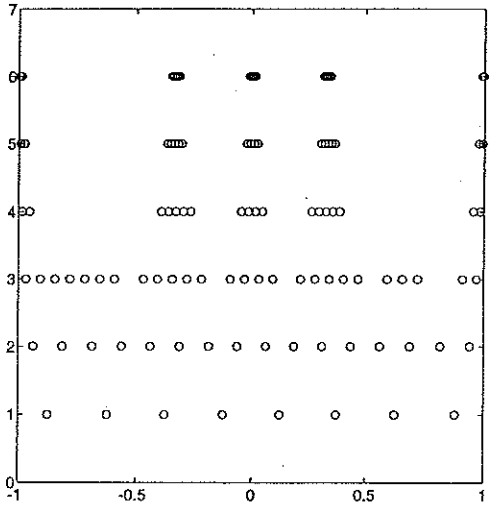
Figure 2. Subdivision limit for $\{\psi_j^1\}_{j=1}^4$



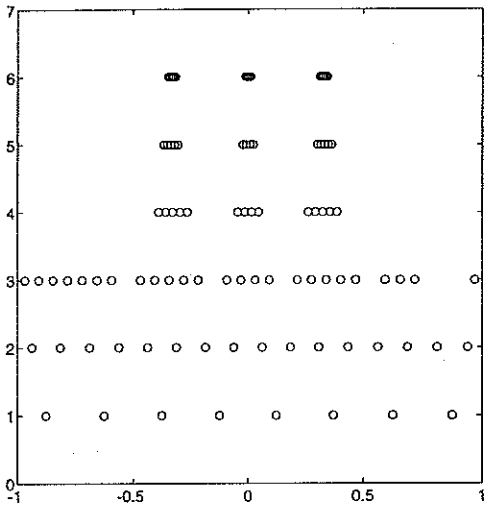
(a) Input data.



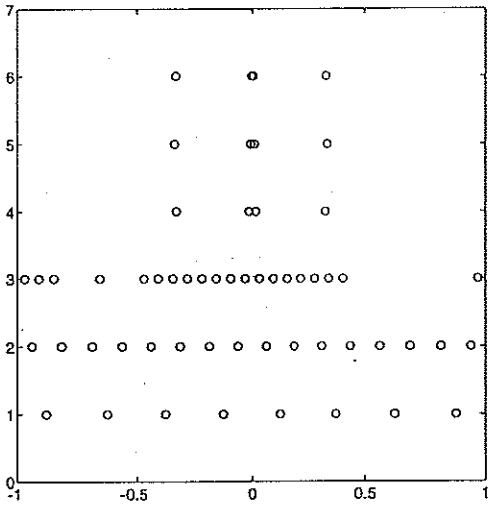
(b) Orthonormal wavelets.



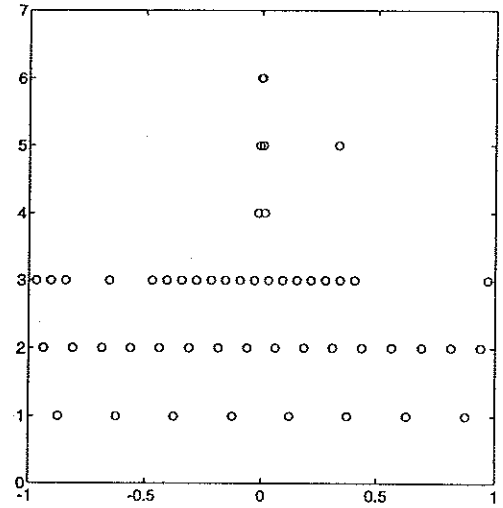
(c) Biorthogonal wavelets.



(d) Piecewise-polynomial.

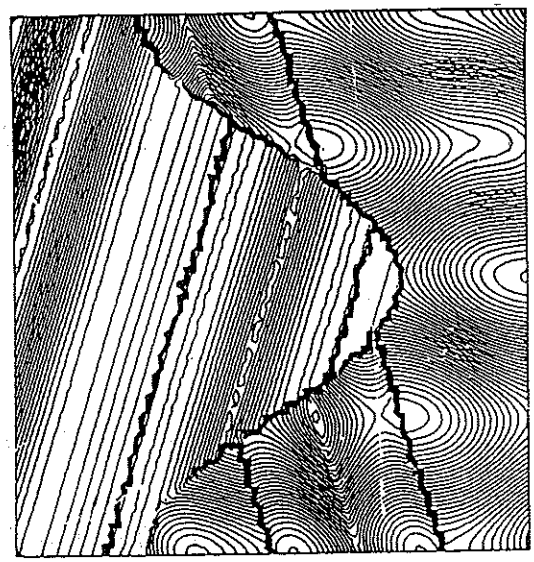


(e) ENO reconstruction.

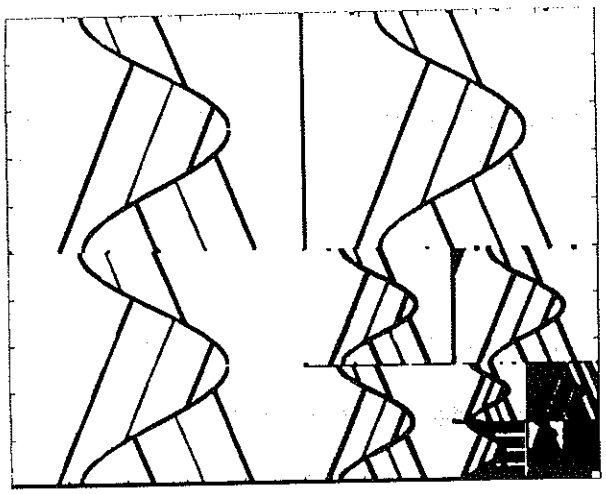


(f) ENO + Subcell resolution.

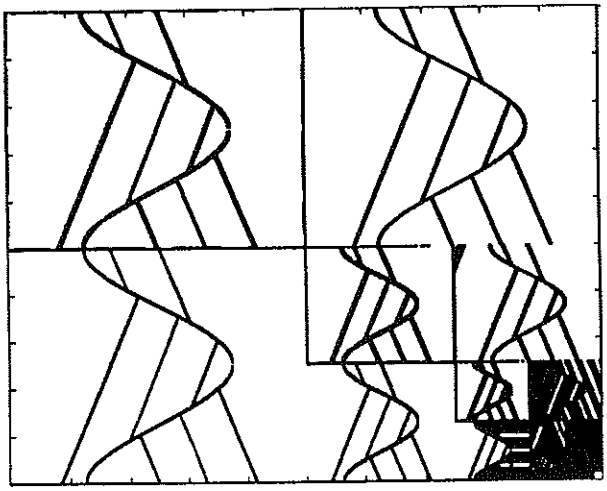
Figure 3. Data compression.



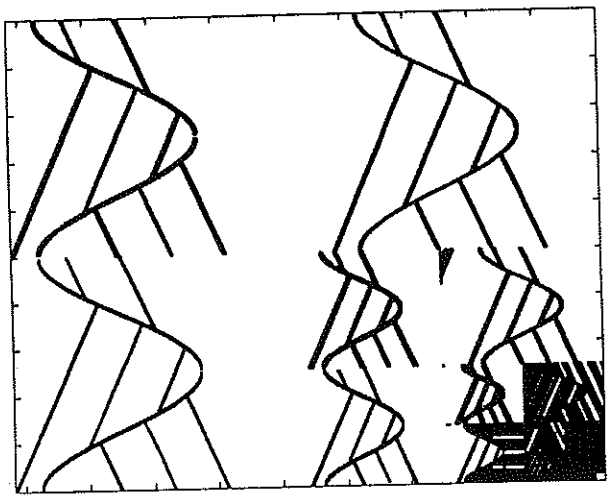
(a) Input data.



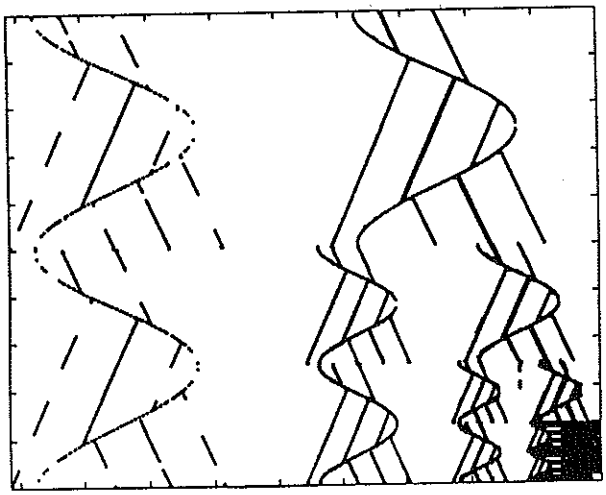
(b) Orthonormal wavelets.



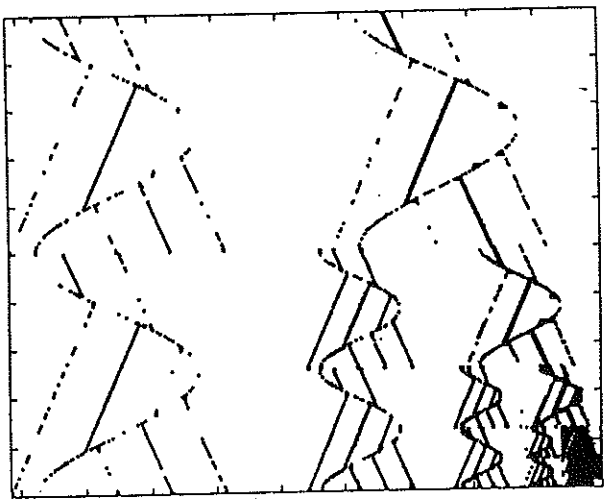
(c) Biorthogonal wavelets.



(d) Piecewise-polynomial.

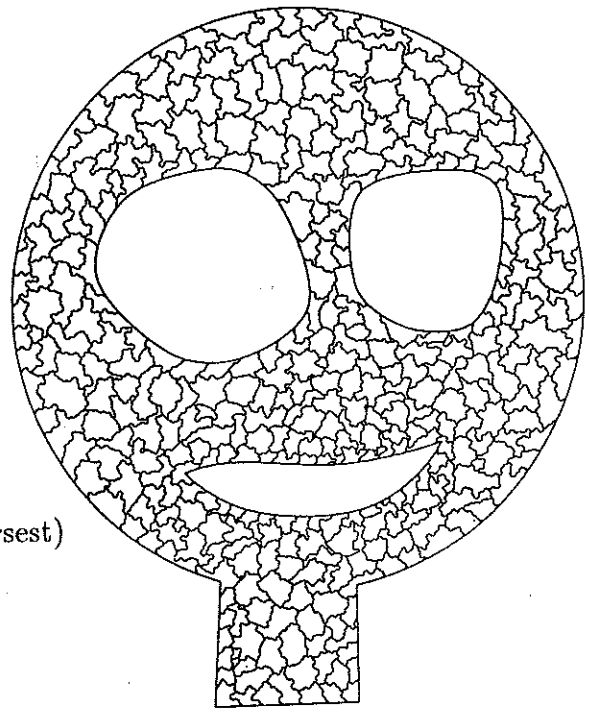


(e) ENO reconstruction.

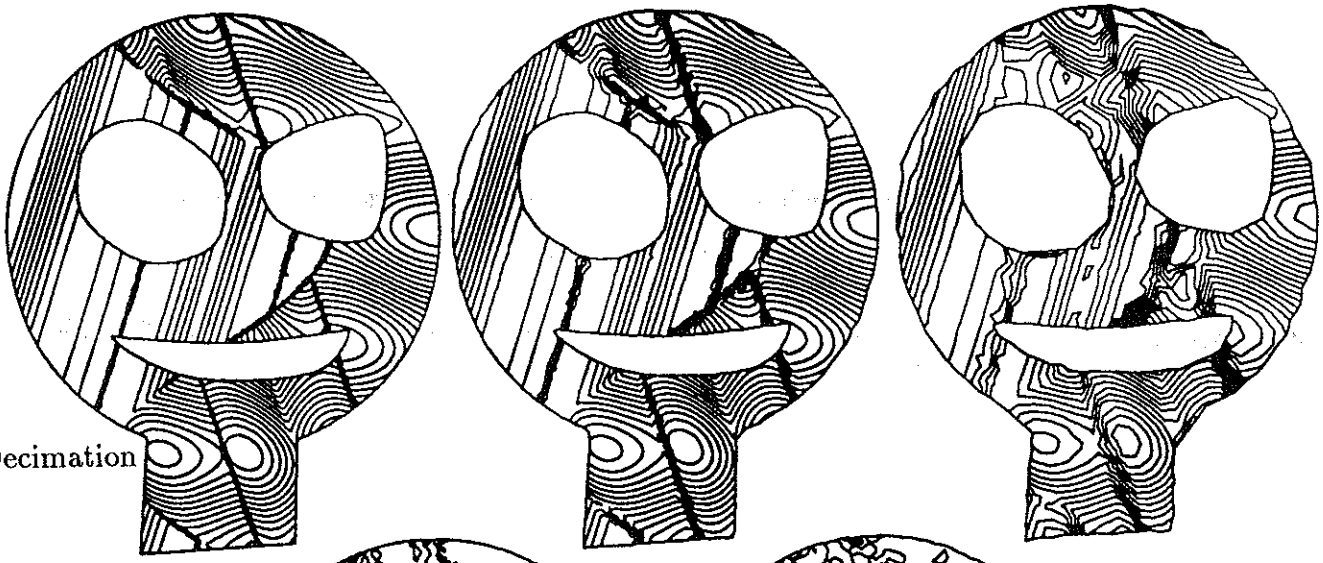


(f) ENO + Subcell resolution.

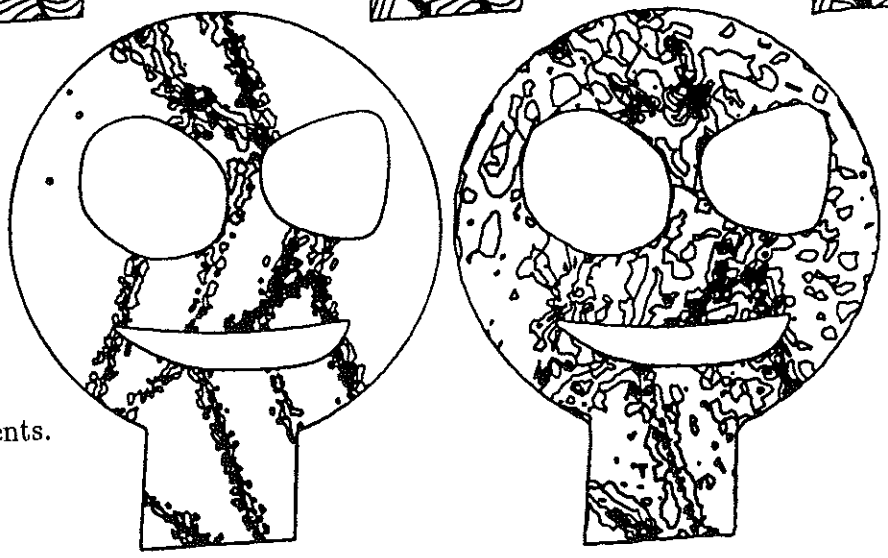
Figure 4. Data compression of 2-D array .



(a) Agglomerated mesh (coarsest)



(b) Decimation



(c) Scale coefficients.

Figure 5. Data compression in unstructured mesh.