HAL
open science

A New Formalism for Nonlinear and Non-Separable Multi-scale Representation

Cédric Gérot, Basarab Matei, Sylvain Meignen

To cite this version:

Cédric Gérot, Basarab Matei, Sylvain Meignen. A New Formalism for Nonlinear and Non-Separable Multi-scale Representation. 2010. hal-00529531v1

HAL Id: hal-00529531

https://hal.science/hal-00529531v1

Preprint submitted on 25 Oct 2010 (v1), last revised 31 Dec 2010 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A New Formalism for Nonlinear and Non-Separable Multi-scale Representation

Cédric Gérot[a], Basarab Matei[b], Sylvain Meignen[c]

[a] *GIPSA Laboratory, University of Grenoble, France*
Tel:0033-4-76-82-71-32
FAX:0033-4-76-82-63-84
E-mail: cedric.gerot@gipsa-lab.inpg.fr

[b] *LAGA Laboratory, Paris XIII University, France,*
Tel:0033-1-49-40-35-71
FAX:0033-4-48-26-35-68
E-mail: matei@math.univ-paris13.fr

[c] *LJK Laboratory, University of Grenoble, France*
Tel:0033-4-76-51-43-95
FAX:0033-4-76-63-12-63
E-mail: sylvain.meignen@imag.fr

Abstract

In this paper, we present a new formalism for nonlinear and non-separable multi-scale representations. We first show that most of the one-dimensional nonlinear multi-scale representations described in the literature are based on prediction operators which are the sum of a linear prediction operator and a perturbation defined using finite differences. We then extend this point of view to the multi-dimensional case where the scaling factor is replaced by a non-diagonal dilation matrix M. The new formalism we propose brings about similarities between existing nonlinear multi-scale representations and also enables us to alleviate the classical hypotheses made to prove the convergence of the multi-scale representations.

Keywords: Nonlinear Multiscale approximation, Besov Spaces, Stability.
2000 MSC: 41A46, MSC 41A60, MSC 41A63

1. Introduction

Multi-scale algorithms such as wavelet-type pyramid transforms for hierarchical data representation [1] and subdivision methods in computer-aided

geometric design [11] have completely changed the domain of data and geometry processing. Linear multi-scale representations of functions is a well understood subject in terms of approximation performance and limitations have also been brought about [6]. While in the univariate case the wavelet-type pyramid transforms provide optimal algorithms, in the multivariate case almost all algorithms fail in the treatment of nonlinear constraints that are inherent to the analyzed objects (e.g. singularities/edges in digital images). For complex objects such as digital images, where nonlinearities arise naturally, one clearly needs to develop data-adaptive algorithms.

Roughly speaking, the development of this emerging nonlinear theory is based on the design of four kinds of data-dependent multi-scale representations: the first one are *quasi-linear multi-scale representations* for piecewise smooth functions [20], the second one are median-interpolating schemes [23], the third one are normal multi-resolutions of curves and surfaces [10] and the fourth one are PPH or power-P representations [5].

The *quasi-linear multi-scale representations* were early introduced in [14], motivated by a better treatment of jumps which served as a simplified model for edges in image analysis. The theoretical analysis of these representations is available in [7]. The PPH and power-P schemes introduced in [5] to design multi-scale representations were also motivated by a better treatment of edges in image compression applications but do not lead to *quasi-linear multi-scale representations*. The median-interpolating scheme was motivated by applications to non-Gaussian noise removal (see [23]), while the normal multi-resolution was defined in [10] for optimal geometry compression of curves and surfaces. These last two multi-scale representations are examples of geometric nonlinear multi-scale transforms.

In general, the analysis of nonlinear representations naturally extends existing results for linear representations and is deeply related to the study of the difference operators associated to the underlying nonlinear prediction operators. The key point for such an analysis is the study of the joint spectral radius of these difference operators.

In what follows, we propose a new formalism for nonlinear prediction operator that enables to embed classical ENO (Essentially Non Oscillatory) prediction operators [20], the PPH and power-P schemes [4]. In a nutshell, the main idea is to write the classical nonlinear prediction operators as the sum of a linear one plus a perturbation term, which is a Lipschitz function of the differences. The order of the perturbation will be related to the polynomial reproduction order of the linear prediction operator. We call

Lipschitz-Linear these nonlinear prediction operators. With this formalism, the existence of the difference operators follows from the hypothesis made on the perturbation term.

After having introduced some generalities and notations for our model on *Lipschitz-Linear* prediction operators (section 2 and section 4), we show that ENO prediction operator and PPH scheme fit into this model in (section 5).

Nonlinear multi-scale representations are naturally defined using nonlinear prediction operators. These one-dimensional multi-scale representations are based on dyadic scales and therefore naturally extend to the multi-dimensional case by tensor product [20]. However, for applications it may be of interest to define multi-scale representations that are not based on a dyadic grid. Several examples exist in image processing where the use of representations built using non-dyadic grids significantly improves the compression performance [8],[19] and [18]. For that reason, we study the extension of the proposed formalism for nonlinear prediction operators to the case where the scales are defined using non-diagonal dilation matrices (section 6). We see that it is again possible to define *Lipschitz-Linear* prediction operators in that context, and we give several examples of such operators, namely the recently introduced PPH and power-P scheme. Sections 7 and 8 are devoted to the convergence and stability properties of the nonlinear multi-scale representations based on *Lipschitz-Linear* prediction operators both in L^p and Besov spaces. These results are very similar to that developed in [21], but some new important aspects for applications are brought about, especially the exact reproduction of polynomials is no longer necessary. A new aspect is introduced in section 9, namely the notion of prediction operators compatible with a set of differences. The results on the convergence and the stability of the corresponding multi-scale representations are identical to those obtained for representations based on *Lipschitz-Linear* prediction operators. For applications, the interesting aspect of this notion of compatibility is that it enables to reduce the complexity of the study of the joint spectral radius. In section 10, we conclude the paper showing the convergence of some nonlinear multi-scale representations associated with *Lipschitz-Linear* prediction operators, namely the PPH scheme.

2. Notations

Before we start, we need to introduce some standard multi-index notations. For example, for $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}_0^d$ we write $|\alpha| = \sum_{i=1}^d \alpha_i$ and for

$x \in \mathbb{R}^d$ we write $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$, monomial with degree $|\alpha|$. By $(e_1, \cdots, e_d)$ we denote the canonical basis on $\mathbb{Z}^d$. There are $r_N^d = \begin{pmatrix} N+d-1 \\ N \end{pmatrix}$ monomials x^α with degree N. We then introduce $\prod_N$ the space of polynomials of degree N generated by

$$\{x^\alpha = \prod_{i=1}^{d} x_i^{\alpha_i}, |\alpha| \leq N\}.$$

In what follows, we will write $\deg(p)$ for the degree of any polynomial p. With that in mind, we denote, for any multi-index α and any sequence $(v_k)_{k \in \mathbb{Z}^d}$:

$$\Delta^\alpha v_k = \Delta_{e_1}^{\alpha_1} \cdots \Delta_{e_d}^{\alpha_d} v_k$$

where $\Delta_{e_d}^{\alpha_d} v_k$ is defined recursively by:

$$\Delta_{e_d}^{\alpha_d} v_k = \Delta_{e_d}^{\alpha_d - 1} v_{k+e_d} - \Delta_{e_d}^{\alpha_d - 1} v_k.$$

For a given multi-index α, we will say that Δ^α is a difference of order $|\alpha|$. For any $N \in \mathbb{N}$, we will denote

$$\Delta^N v_k = \{\Delta^\alpha v_k, |\alpha| = N\}. \tag{1}$$

3. Multi-scale Representations

We assume that we are given the data v^j, associated to some grids Γ^j, $j \geq 0$. We also consider that there exist two interscale operators: (i) D-the downsampling operator and (ii) S-the prediction operator. The operator D gives a coarse version v^{j-1} of v^j. The operator S computes $\hat{v}^j = Sv^{j-1}$, an approximation of v^j. Then, we define the prediction error as $e^j = v^j - \hat{v}^j$. The information contained in v^j is completely equivalent to (v^{j-1}, e^j). By iterating this procedure from the initial data v^J, we obtain its *nonlinear multi-scale representation*

$$\mathcal{M}v^J = (v^0, e^1, \cdots, e^J). \tag{2}$$

Conversely, assume that the sequence $(v^0, (e^j)_{j \geq 0})$ is given, we are interested in studying the convergence of the following nonlinear iteration:

$$v^j = Sv^{j-1} + e^j, \tag{3}$$

to a limit function v, which is defined as the limit (when it exists) of:

$$v_j(x) = \sum_{k\in\mathbb{Z}^d} v_k^j \varphi_{j,k}(x),$$

where $\varphi_{j,k}(x)$ denotes $\varphi(M^j x - k)$ and φ is some specific function satisfying the scaling equation:

$$\varphi(x) = \sum_{n\in\mathbb{Z}^d} g_n \varphi(Mx - n) \text{ with } \sum_n g_n = m := |\det M|, \tag{4}$$

where M is a dilation matrix, (i.e. an invertible matrix in $\mathbb{Z}^d \times \mathbb{Z}^d$ satisfying $\lim_{n\to+\infty} M^{-n} = 0$). When the sequence of functions $(v_j)_{j\geq 0}$ is convergent to some limit function in some functional space, by abusing a little bit terminology, we say that the multi-scale representations $(v^0, (e^j)_{j\geq 0})$ is convergent in that space.

4. *Lipschitz-Linear* Prediction Operators

In the following a prediction operator is a map $v \in \ell^\infty(\mathbb{Z}^d) \mapsto Sv \in \ell^\infty(\mathbb{Z}^d)$. We need to recall the definition of a local and linear prediction operator S_l which is as follows:

$$(S_l v)_k = \sum_{l\in\mathbb{Z}^d} g_{k-Ml} v_l \tag{5}$$

where $\exists K > 0$ such that $g_{k-Ml} = 0$ if $\|k - Ml\|_\infty > K$ and we put $F(k) = \{l \in \mathbb{Z}^d, \|k - Ml\| < K\}$. In the following, we assume that the linear prediction operator S_l is such that $(g_n)_{n\in\mathbb{Z}^d}$ defines a scaling function φ satisfying the scaling equation (4). In this paper, we study a particular type of nonlinear prediction operator which is the sum of a linear prediction operators and a perturbation term. The linear prediction operator shall satisfy polynomial reproduction property which we now recall:

Definition 4.1. *We say that a prediction operator S reproduces polynomials of degree N if for $u_k = p(k)$ for any $p \in \prod_N$, we have*

$$Su_k = p(M^{-1}k) + q(k)$$

where q is a polynomial such that $\deg(q) < \deg(p)$. *When $q = 0$, we say that the prediction operator exactly reproduces polynomials.*

With this in mind, we introduce the definition of *Lipschitz-Linear* prediction operator:

Definition 4.2. *A prediction operator S is Lipschitz-Linear of order $N+1$, if there exists a local and linear prediction operator S_l reproducing polynomials of degree N, and Lipschitz functions Φ_i $i = 0, \cdots, m-1$ such that:*

$$(Sv)_{Mk+i} = (S_l v)_{Mk+i} + \Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}) \quad \forall i \in \ coset(M)$$

where $\{p_1, \cdots, p_q\}$ is a fixed set and where a M is a dilation matrix.

Remark 4.1. *From the above definition, we remark that when S_l reproduces polynomials of degree N so does S.*

5. One-Dimensional *Lipschitz-Linear* Prediction Operators

5.1. Preliminaries

We start by considering the one-dimensional case with $M = 2$. Given a set of embedded grids $\Gamma^j = \{2^{-j}k,\ k \in \mathbb{Z}\}$ we consider discrete values v_k^j defined on each vertex of these grids. These quantities shall represent a certain function v at level j. Typical examples of such discretizations are: point-values, where $v_k^j = v(2^{-j}k)$ for some function v and cell-averages, where v_k^j is the average of v over a neighborhood of $2^{-j}k$. We call them point-values (resp. cell-average) *multi-scale representation*. Assuming a certain type of multi-scale representation, we define a nonlinear prediction operator that in turn leads to a nonlinear multi-scale representation.

Let us now recall some properties of Lagrange interpolation useful to show that some classical nonlinear prediction operators are *Lipschitz-Linear*. Consider the interpolation polynomial p_N of degree N of v at $x_0, \cdots, x_N$ and $p_{N,r}$ (where r stands for right) the interpolation polynomial of v at the same set as p_N but shifted to the right (i.e. $x_1, \cdots, x_{N+1}$). Using standard arguments, we write the difference between the two polynomials as:

$$p_{N,r}(x) - p_N(x) \quad = \quad \Delta^{N+1} v_0 \frac{1}{N! h^N} \prod_{i=1}^{N} (x - x_i), \tag{6}$$

where $h = x_{i+1} - x_i$. The same kind of result can be obtained considering $p_{N,l}$, the interpolation polynomial at $x_{-1}, \cdots, x_{N-1}$.

5.2. Quasi-Linear Prediction Operators Using Polynomial Interpolation

Here, we use identity (6) to analyze nonlinear prediction operators in the context of point-values *multi-scale representations*. These operators compute the approximation $\hat{v}_k^j$ of $v_k^j = v(2^{-j}k)$ using only $v_k^{j-1} = v(2^{-j+1}k) \ \in \mathbb{Z}$. In this framework, since $v_{2k}^j = v_k^{j-1}$ (also called *consistency* property), only $\hat{v}_{2k+1}^j$ need be computed. To do so, we consider the Lagrange polynomial p_{2N+1} of degree $2N+1$ defined on the $2N+2$ closest neighbors of $2^{-j}(2k+1)$ on Γ^{j-1}, i.e.

$$p_{2N+1}(2^{-j+1}(k+n)) = v_{k+n}^{j-1} = v(2^{-j+1}(k+n)), \quad n = -N, \cdots, N+1.$$

This polynomial is used to compute $\hat{v}_{2k+1}^j$ through the so-called *centered* prediction as follows:

$$\hat{v}_{2k+1}^j = p_{2N+1}(2^{-j}(2k+1)). \tag{7}$$

When $N = 1$, we obtain the four points scheme:

$$\hat{v}_{2k+1}^j = \frac{9}{16}(v_k^{j-1} + v_{k+1}^{j-1}) - \frac{1}{16}(v_{k-1}^{j-1} + v_{k+2}^{j-1})$$

which is exact for cubic polynomials. The four point scheme was widely studied in litterature (see [12]). Now, consider the polynomial $p_{2N+1,1}$ whose interpolation set is that of p_{2N+1} shifted by 2^{-j+1} to the right. This leads, for instance, when $N = 1$, to the prediction:

$$\hat{v}_{2k+1,1}^j := p_{3,1}(2^{-j}(2k+1)) = \frac{5}{16}v_k^{j-1} + \frac{15}{16}v_{k+1}^{j-1} - \frac{5}{16}v_{k+2}^{j-1} + \frac{1}{16}v_{k+3}^{j-1}. \tag{8}$$

Now, if we compute the difference between the above predictions we obtain:

$$\hat{v}_{2k+1,1}^j - \hat{v}_{2k+1}^j = \frac{1}{16}\Delta^4 v_{k-1}^{j-1}, \tag{9}$$

which corresponds to (6), with $x_i = 2^{-j+1}(k+i-1)$, $i = 0, \cdots, 2$ and $x = 2^{-j}(2k+1)$.

The same conclusion is also valid for the polynomial $p_{2N+1,-1}$, for $N = 1$, whose interpolation set is that of p_{2N+1} but shifted to the left by 2^{-j+1}. We now show how to generalize the above formula to any N:

Proposition 5.1. *For any N, assume that $\hat{v}_k^j$ (resp. $\hat{v}_{k,1}^j$) is obtained using the polynomial p_{2N+1} (resp. $p_{2N+1,1}$), then:*

$$\hat{v}_{2k+1,1}^j - \hat{v}_{2k+1}^j = (-1)^{N-1}\Delta^{2N+2}v_{k-N}^{j-1}\frac{1}{2^{4N}}\begin{pmatrix} 2N-1 \\ N \end{pmatrix}$$

Proof. Let us put $y_0 = v_{k-N}^{j-1}, \cdots, y_{2N+1} = v_{k+N+1}^{j-1}$, and $x_0 = 2^{-j+1}(k-N), \cdots, x_{2N+1} = 2^{-j+1}(k+N+1)$. Then, using (6) the difference between p_{2N+1} and $p_{2N+1,1}$ evaluated at $2^{-j}(2k+1)$, reads as follows:

$$\begin{aligned} \hat{v}_{2k+1,1}^j - \hat{v}_{2k+1}^j &= -\Delta^{2N+2}y_0\frac{1}{(2N+1)!2^{2N+1}}\prod_{i=-N+1}^{N+1}(2i-1) \\ &= (-1)^{N-1}\Delta^{2N+2}v_{k-N}^{j-1}\frac{1}{2^{4N}}\frac{(2N-1)!}{N!(N-1)!} \end{aligned}$$

□

Remark 5.1. *Note that we can define other polynomials $p_{2N+1,q}$ for $-N \leq q \leq N$, that are obtained by shifting the centered interpolation set by $q2^{-j+1}$, and then predict using one of these polynomials. In any case, the difference between this prediction and the centered one will be a linear function of the differences of order $2N+2$, since we can write (assuming $q > 0$, but still true for any q) that:*

$$\hat{v}_{2k+1,q}^j - \hat{v}_{2k+1}^j = \sum_{l=1}^{q-1}\hat{v}_{2k+1,l+1}^j - \hat{v}_{2k+1,l}^j + \hat{v}_{2k+1,1}^j - \hat{v}_{2k+1}^j,$$

and then apply Proposition 5.1.

The different prediction operators defined above satisfy the exact polynomial reproduction of degree $2N+1$. To define a so-called *quasi-linear* prediction operator, at each location k we choose between the $2N+1$ predictions according to some criterion. The nonlinearity is thus contained in the choice of the stencil used to predict. An example of such prediction operator is the essentially non-oscillatory (ENO) prediction operator introduced by Harten in [15]. Given the prediction operator defined using $p_{2N+1,q}$ and by considering as its linear part the centered prediction, Proposition 5.1 says that it is a *Lipschitz-linear* prediction operator.

Assume that $p_{2N+1,\tilde{q}}$ is the chosen polynomial by ENO criterion for prediction at $2^{-j}(2k+1)$ and note that the support of interpolation polynomial contains $2^{-j}(2k+1)$. By using shifted stencils, we improve the precision in the vicinity of the singularity and avoid Gibbs oscillations. Nevertheless, the prediction is still unsatisfactory within the interval containing a potential singularity. This is the origin of the subcell resolution strategy, which improves the prediction within the interval by using extrapolation techniques. Since the polynomials used in the extrapolation are built using a shifted stencil as previously, it leads again to a *Lipschitz-linear* prediction operator (cf. Remark 5.1).

5.3. Quasi-Linear Prediction Operators Using Cell-Average Interpolation

We now show how relation (6) still enables to prove that *quasi-linear* prediction operators designed on cell-average *multi-scale representations* are *Lipschitz-Linear*. In the cell-average setting the data v_k^j is the average of some function v over the interval $I_{j,k} = [2^{-j}k, 2^{-j}(k+1)]$ as follows:

$$v_k^j = 2^j \int_{I_{j,k}} v(t)dt \tag{10}$$

In that framework, we have the so-called *consistency* property :

$$v_k^{j-1} = \frac{1}{2}(v_{2k}^j + v_{2k+1}^j). \tag{11}$$

Now, we design a nonlinear prediction operator on this multi-scale representation considering the interpolation polynomial p_{2N} of degree $2N$ defined as follows:

$$2^{j-1} \int_{I_{j-1,k+n}} p_{2N}(t)dt = v_{k+n}^{j-1} \quad n = -N, \cdots, N.$$

We then define the *centered* prediction by:

$$\hat{v}_{2k}^j = 2^j \int_{I_{j,2k}} p_{2N,k}(t)dt \text{ and } \hat{v}_{2k+1}^j = 2^j \int_{I_{j,2k+1}} p_{2N,k}(t)dt.$$

For instance, when $N = 1$, this leads to:

$$\hat{v}_{2k}^j = v_k^{j-1} + \frac{1}{8}(v_{k-1}^{j-1} - v_{k+1}^{j-1}) \text{ and } \hat{v}_{2k+1}^j = v_k^{j-1} - \frac{1}{8}(v_{k-1}^{j-1} - v_{k+1}^{j-1}).$$

Still for $N = 1$, the prediction operator built using the polynomial $p_{2N,1}$ that interpolates the average on intervals $I_{j-1,k}, I_{j-1,k+1}, I_{j-1,k+2}$ leads to the following predictions:

$$\hat{v}^j_{2k,1} = \frac{11}{8}v^{j-1}_k - \frac{1}{2}v^{j-1}_{k+1} + \frac{1}{8}v^{j-1}_{k+2} \text{ and } \hat{v}^j_{2k+1,1} = \frac{15}{8}v^{j-1}_k + \frac{1}{2}v^{j-1}_{k+1} - \frac{1}{8}v^{j-1}_{k+2}.$$

Now, if we compute the difference between this shifted prediction and the *centered* one, we get:

$$\hat{v}^j_{2k+1,1} - \hat{v}^j_{2k+1} = -\frac{1}{8}\Delta^3 v^{j-1}_{k-1} \text{ and } \hat{v}^j_{2k,1} - \hat{v}^j_{2k} = \frac{1}{8}\Delta^3 v^{j-1}_{k-1}. \tag{12}$$

Similarly, we can define a prediction using the set of intervals shifted to the left and obtain the same kind of result. The equality (12) can then be generalized to any N:

Proposition 5.2. *Consider the prediction $\hat{v}^j_k$ (resp. $\hat{v}^j_{k,1}$) obtained using p_{2N} (resp. $p_{2N,1}$), then we may write:*

$$\begin{aligned}
\hat{v}^j_{2k,1} - \hat{v}^j_{2k} &= (-1)^{N-1}\Delta^{2N+1}v^{j-1}_{k-N}\frac{1}{2^{4N-1}}\begin{pmatrix} 2N-1 \\ N \end{pmatrix} \\
\hat{v}^j_{2k+1,1} - \hat{v}^j_{2k+1} &= -(\hat{v}^j_{2k,1} - \hat{v}^j_{2k})
\end{aligned}$$

Proof. To consider the interpolation of the average on $I_{j-1,k+n}$, $n = -N, \cdots, N$ using the polynomial p_{2N} is equivalent to consider the primitive P_{2N} of p_{2N} such that $\bar{P}_{2N} = 2^{j-1}P_{2N}$ interpolates $y_0 = 0, y_1 = v^{j-1}_{k-N}, y_2 = y_1 + v^{j-1}_{k-N+1}, \cdots, y_{2N+1} = y_{2N} + v^{j-1}_{k+N}$ respectively at $x_0 = 2^{-j+1}(k-N), x_1 = 2^{-j+1}(k-N+1), x_2 = 2^{-j+1}(k-N+2), \cdots, x_{2N+1} = 2^{-j+1}(k+N+1)$. Similarly, the interpolation of the average computed on the intervals $I_{j-1,k+n}$, $n = -N+1, \cdots, N+1$ using polynomial $p_{2N,1}$ is equivalent to consider its primitive $P_{2N,1}$ such that $\bar{P}_{2N,1} = 2^{j-1}P_{2N,1}$ interpolates $\tilde{y}_1 = 0, \tilde{y}_2 = v^{j-1}_{k-N+1}, \tilde{y}_3 = \tilde{y}_2 + v^{j-1}_{k-N+2}, \cdots, \tilde{y}_{2N+2} = \tilde{y}_{2N+1} + v^{j-1}_{k+N+1}$ respectively at $x_1, x_2, \cdots, x_{2N+2} = 2^{-j+1}(k+N+2)$. Using the Newton form for each polynomial $\bar{P}_{2N}$ and $\bar{P}_{2N,1}$ and remarking that the divided differences are such that: $[\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_k] = [y_1, y_2, \cdots, y_k]$ for all $k \leq 2N+2$, we write:

$$\begin{aligned}
\bar{P}_{2N,1}(x) - \bar{P}_{2N}(x) &= -v^{j-1}_{k-N} + [y_0, \cdots, y_{2N+2}](x_{2N+2} - x_0)\prod_{i=1}^{2N+1}(x - x_i) \\
&= -v^{j-1}_{k-N} + \Delta^{2N+1}v^{j-1}_{k-N}\frac{1}{(2N+1)!(2^{-j+1})^{2N+1}}\prod_{i=1}^{2N+1}(x - x_i).
\end{aligned}$$

In that framework, we also have:

$$v_k^{j-1} = \bar{P}_{2N}(2^{-j+1}(k+1)) - \bar{P}_{2N}(2^{-j+1}k) = \bar{P}_{2N,1}(2^{-j+1}(k+1)) - \bar{P}_{2N,1}(2^{-j+1}k).$$

The *centered* prediction following (10) is:

$$\begin{aligned} \hat{v}_{2k}^j &= 2\left(\bar{P}_{2N}(2^{-j+1}(k+1/2)) - \bar{P}_{2N}(2^{-j+1}k)\right) \\ \hat{v}_{2k+1}^j &= 2\left(\bar{P}_{2N}(2^{-j+1}(k+1)) - \bar{P}_{2N}(2^{-j+1}(k+1/2)\right). \end{aligned}$$

Considering the leading coefficient of the polynomial P_{2N}, one can check that the corresponding prediction operator reproduces polynomials of degree $2N+1$. The definition of $\hat{v}_{2k,1}^j$ and $\hat{v}_{2k+1,1}^j$ are identical to that of $\hat{v}_{2k}^j$ and $\hat{v}_{2k+1}^j$ replacing P_{2N} by $P_{2N,1}$. Then, computing the difference between $P_{2N,1}$ and P_{2N} and applying it at putting $x = 2^{-j}k$, we get:

$$\hat{v}_{2k,1}^j - \hat{v}_{2k}^j = \Delta^{2N+1} v_{k-N}^{j-1} (-1)^{N-1} \frac{1}{2^{4N-1}} \begin{pmatrix} 2N-1 \\ N \end{pmatrix}$$

□

As in the point-values setting, we can define $p_{2N,q}$, for any q, by shifting the computation intervals and then predict using this polynomial. Following the same argument as previously, it leads to a *Lipschitz-Linear* prediction operator.

5.4. PPH scheme as Lipschitz-Linear Prediction Operator

In this section, we show that PPH and power-P schemes are examples of *Lipschitz-Linear* prediction operators ([4]). The PPH scheme is defined by:

$$\left\{ \begin{array}{c} \hat{v}_{2k+1}^j = \frac{v_{k+1}^{j-1} + v_k^{j-1}}{2} - \frac{1}{8} H(\Delta^2 v_{k-1}^{j-1}, \Delta^2 v_k^{j-1}) \\ \hat{v}_{2k}^j = v_k^{j-1} \end{array} \right. \tag{13}$$

where $H(x,y) := \frac{xy}{x+y}(\mathrm{sign}(xy)+1)$. Since H satisfies $|H(x,y) - H(x',y')| \leq 2\max\{|x-x'|, |y-y'|\}$, it is Lipschitz with respect to (x,y) and since the linear scheme $\frac{v_{k+1}^{j-1} + v_k^{j-1}}{2}$ reproduces polynomials of degree 1, the PPH-scheme is a *Lipschitz-Linear* prediction operator of order 2. The power-P scheme [26], is a generalization of the PPH-scheme replacing H by

$$H_p(x,y) = \left\{ \begin{array}{c} \frac{x+y}{2}\left(1 - |\frac{x-y}{x+y}|^p\right), \quad xy > 0 \\ 0, \quad xy \leq 0, \end{array} \right.$$

where H_p is still Lipschitz (see [26] for details on power-P scheme), and it is also a *Lipschitz-Linear* prediction operator of order 2.

6. Multi-Dimensional *Lipschitz-Linear* Prediction Operators on Non-Dyadic Grids

To illustrate the notion of *Lipschitz-Linear* prediction operators in the multivariate case, we introduce the concept of nonlinear prediction on non-dyadic grids. The motivation to consider this type of grids are, for instance, better image compression results (see [8] and [19]). Having defined the grid $\Gamma^j = \left\{M^{-j}k, \quad k \in \mathbb{Z}^d\right\}$ using a dilation matrix M, we consider discrete quantities v_k^j defined on each of these grids. They shall represent a certain approximation of a function v at level j. As in the one-dimensional case, we will first evoke two kinds of examples: the point-values setting, corresponding to $v(M^{-j}k)$ and the cell-average setting where v_k^j corresponds to a local average of v over a neighborhood of $M^{-j}k$.

6.1. Quasi-Linear Prediction Operators Using Point-Values Interpolation

In what follows, we consider point-values *multi-scale representations* based on scales defined by the quincunx matrix:

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{14}$$

To build a nonlinear prediction operator, we first consider the polynomial p_2 of degree 2 which interpolates v on Γ^{j-1} on the stencil $V^1 = M^{-j+1}\{k, k+e_1, k+e_2, k+e_1+e_2, k+2e_1, k+2e_2\}$. Note that such a polynomial exists but this may not be the case with another stencil. Since $m = 2$, there are two cosets associated with M (i.e. $\mathbb{Z}^2 = \{Mk, k \in \mathbb{Z}^2\} \bigcup \{Mk + e_1, k \in \mathbb{Z}^2\}$) and since we consider point-values *multi-scale representations*, we have $\hat{v}_{Mk}^j = v_k^{j-1}$. We predict $v_{Mk+e_1}^j$ for $k \in \mathbb{Z}^2$ using p_2:

$$\hat{v}_{Mk+e_1}^{j,1} = \frac{1}{2}\left(v_{k+e_1}^{j-1} + v_{k+e_2}^{j-1}\right) - \frac{1}{8}\left(v_{k+2e_1}^{j-1} + v_{k+2e_2}^{j-1}\right) + \frac{1}{4}v_{k+e_1+e_2}^{j-1}. \tag{15}$$

Now, if we move the stencil by $-e_1$ or by $-e_2$, we obtain the respective stencils:

$$\begin{aligned} V^2 &= M^{-j+1}\{k-e_1, k, k+e_1, k-e_1+e_2, k+e_2, k-e_1+2e_2\} \\ V^3 &= M^{-j+1}\{k-e_2, k-e_2+e_1, k-e_2+2e_1, k, k+e_1, k+e_2\}, \end{aligned}$$

and the corresponding predictions:

$$\begin{aligned} \hat{v}_{Mk+e_1}^{j,2} &= \frac{3}{8}v_{k+e_1}^{j-1} + \frac{3}{4}v_{k+e_2}^{j-1} - \frac{1}{8}v_{k-e_1+2e_2}^{j-1} \\ \hat{v}_{Mk+e_1}^{j,3} &= \frac{3}{8}v_{k+e_2}^{j-1} + \frac{3}{4}v_{k+e_1}^{j-1} - \frac{1}{8}v_{k+2e_1-e_2}. \end{aligned}$$

After some simple computations, we obtain:

$$\hat{v}^{j,1}_{Mk+e_1} - \hat{v}^{j,2}_{Mk+e_1} = -\frac{1}{8}\Delta_{e_1}\Delta^2_{e_2}v^{j-1}_k + \frac{1}{8}\Delta^2_{e_1}\Delta_{e_2}v^{j-1}_{k+e_1-e_2} + \frac{1}{8}\Delta^2_{e_1}\Delta_{e_2}v_{k-e_1} - \frac{1}{8}\Delta^3_{e_1}v^{j-1}_{k-e_1}$$
$$\hat{v}^{j,1}_{Mk+e_1} - \hat{v}^{j,3}_{Mk+e_1} = -\frac{1}{8}\Delta_{e_2}\Delta^2_{e_1}v^{j-1}_k + \frac{1}{8}\Delta^2_{e_2}\Delta_{e_1}v^{j-1}_{k+e_1-e_2} + \frac{1}{8}\Delta^2_{e_2}\Delta_{e_1}v^{j-1}_{k-e_2} - \frac{1}{8}\Delta^3_{e_1}v^{j-1}_{k-e_2}. \quad (16)$$

Note that the prediction operators just introduced, exactly reproduce polynomials of degree 2 and that the difference between $\hat{v}^{j,1}$ and $\hat{v}^{j,2}$ (resp. $\hat{v}^{j,3}$) is a linear function of the differences of order 3. Therefore these prediction operators are *Lipschitz-Linear*.

We generalize (16) through the following Proposition:

Property 6.1. *Let S_l and $\tilde{S}_l$ be two local and linear prediction operators satisfying exact polynomial reproduction of degree N and defined by using g and $\tilde{g}$ respectively. Then the difference between predictions S_l and $\tilde{S}_l$ can be expressed as a linear combination of differences of order $N+1$*

Proof. If $v^{j-1}_l = p(l)$ for some $p \in \prod_N$ we have:

$$\sum_{l\in\mathbb{Z}^d}(g_{k-Ml} - \tilde{g}_{k-Ml})v^{j-1}_l = 0 \quad \forall k \in \mathbb{Z}^d.$$

From this, we deduce that $(g_{k-Ml} - \tilde{g}_{k-Ml})_{l\in\mathbb{Z}^d}$ is orthogonal to $\prod_N$. According to [17], Theorem 4.3, a potential basis is $\{\Delta^\mu\delta_{.-l}, |\mu| = N+1, l \in \mathbb{Z}^d\}$, so that we may write:

$$g_{k-Ml} - \tilde{g}_{k-Ml} = \sum_{|\mu|=N+1}\sum_{l\in\mathbb{Z}^d} c^\mu_l \Delta^\mu \delta_{k-l}$$

and then by composition we obtain, for any v^{j-1}:

$$\sum_{l\in F(k)}(g_{k-Ml} - \tilde{g}_{k-Ml})v^{j-1}_l = \sum_{l\in F(k)}\sum_{|\mu|=N+1} c^\mu_k \Delta^\mu v^{j-1}_l.$$

□

6.2. Quasi-Linear Prediction Operators Using Cell-Average Interpolation

We now introduce a special version of multi-dimensional cell-average interpolation that leads to a *Lipschitz-Linear* prediction operator. Indeed, let us consider the following discretization:

$$v^j_k = m^j \int v(x)\tilde{\varphi}(M^j x - k)dx,$$

where M is the quincunx matrix and $\tilde{\varphi}$ is the characteristic function of $\{\lambda e_1 + \beta M e_1, 0 \leq \lambda \leq 1, 0 \leq \beta \leq 1\}$. We define the polynomial p_2 of degree 2 such that:

$$v_{k+n}^{j-1} = m^{j-1} \int p_2(x)\tilde{\varphi}(M^{j-1}x - k - n)dx \quad n = -e_1 - e_2, 0, e_1 + e_2 \quad (17)$$

and then, we choose three other values for n in the set $\{-e_1, -2e_1 - e_2, e_2, e_1, 2e_1 + e_2, -e_2\}$ to define the polynomial p_2 completely (following the same interpolation condition as (17)). The prediction is then computed by:

$$\hat{v}_k^j = m^j \int p_2(x)\tilde{\varphi}(M^j x - k)dx.$$

In this particular case, regardless of the last three points chosen to define p_2, we obtain the so-called *centered* prediction:

$$\hat{v}_{Mk}^j = v_k^{j-1} + \frac{1}{8}\left(v_{k-e_1-e_2}^{j-1} - v_{k+e_1+e_2}^{j-1}\right) \quad (18)$$

and then by *consistency*:

$$\hat{v}_{Mk+e_1}^j = v_k^{j-1} - \frac{1}{8}\left(v_{k-e_1-e_2}^{j-1} - v_{k+e_1+e_2}^{j-1}\right). \quad (19)$$

Now, let us consider the following stencils for prediction:

$$\begin{aligned}
V^1 &= \{-2e_1 - e_2, -e_1, e_2, 0, e_1 + e_2, q\} \text{ for any } q \in \{-e_2, e_1, 2e_1 + e_2\} \\
V^2 &= \{-2e_1 - e_2, -e_1, e_2, -e_1 - e_2, 0, q\} \text{ for any } q \in \{-e_2, e_1, 2e_1 + e_2\} \\
V^3 &= \{q, -e_1 - e_2, 0, -e_2, e_1, 2e_1 + e_2\} \text{ for any } q \in \{-2e_1 - e_2, -e_1, e_2\} \\
V^4 &= \{q, 0, e_1 + e_2, -e_2, e_1, 2e_1 + e_2\} \text{ for any } q \in \{-2e_1 - e_2, -e_1, e_2\}.
\end{aligned}$$

This leads to the predictions:

$$\begin{aligned}
v_{Mk,1}^j &= \frac{1}{8}(v_{k-2e_1-e_2}^{j-1} + v_{k+e_2}^{j-1}) - \frac{1}{4}(v_{k-e_1}^{j-1} + v_{k+e_1+e_2}^{j-1}) + \frac{5}{4}v_k^{j-1} \\
v_{Mk,2}^j &= -\frac{1}{8}(v_{k-2e_1-e_2}^{j-1} + v_{k+e_2}^{j-1}) + \frac{1}{4}(v_{k-e_1}^{j-1} + v_{k-e_1-e_2}^{j-1}) + \frac{3}{4}v_k^{j-1} \\
v_{Mk,3}^j &= -\frac{1}{8}(v_{k-e_2}^{j-1} + v_{k+2e_1+e_2}^{j-1}) + \frac{1}{4}(v_{k+e_1}^{j-1} + v_{k-e_1-e_2}^{j-1}) + \frac{3}{4}v_k^{j-1} \\
v_{Mk,4}^j &= \frac{1}{8}(v_{k+2e_1+e_2}^{j-1} + v_{k-e_2}^{j-1}) - \frac{1}{4}(v_{k+e_1+e_2}^{j-1} + v_{k+e_1}^{j-1}) + \frac{5}{4}v_k^{j-1}.
\end{aligned}$$

Then, if we compute the difference between these predictions and the *centered* one, we get:

$$\begin{aligned}
v^j_{Mk} - v^j_{Mk,1} &= \frac{1}{8}\Delta_{e_1}\Delta^2_{Me_1} v^{j-1}_{k-2e_1-e_2} \\
v^j_{Mk} - v^j_{Mk,2} &= -\frac{1}{8}\Delta_{e_1}\Delta^2_{Me_1} v^{j-1}_{k-2e_1-e_2} \\
v^j_{Mk} - v^j_{Mk,3} &= \frac{1}{8}\Delta_{e_1}\Delta^2_{Me_1} v^{j-1}_{k-e_1-e_2} \\
v^j_{Mk} - v^j_{Mk,4} &= -\frac{1}{8}\Delta_{e_1}\Delta^2_{Me_1} v^{j-1}_{k-e_1-e_2}. \qquad (20)
\end{aligned}$$

The differences at $Mk + e_1$ are derived using the consistency property. Note that these prediction operators reproduce polynomials of degree 2, and considering the *centered* one as their linear part we deduce that any of these prediction operators are *Lipschitz-Linear*.

6.3. Multi-Dimensional PPH Scheme as Lipschitz-Linear Prediction Operators

In this section, we consider the bi-dimensional PPH-scheme which is:

$$\hat{v}^j_{Mk+e_1} = \begin{cases} \frac{v^{j-1}_k + v^{j-1}_{k+Me_1}}{2} - \frac{1}{8}H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) \\ \hat{v}^j_{Mk} = v^{j-1}_k. \end{cases}$$

Note that the linear part of the prediction operator is obtained by considering an affine interpolation polynomial at v^{j-1}_k, $v^{j-1}_{k+e_1}$ and $v^{j-1}_{k+e_1+e_2}$ and thus reproduces polynomials of degree 1. Since the perturbation is a Lipschitz function of the differences of order 2, this multi-dimensional prediction operator is *Lipschitz-Linear* of order 2. In [2], another generalization of the PPH-scheme to the bidimensional case is proposed:

$$\begin{aligned}
\hat{v}^{j,1}_{Mk+e_1} &= \frac{v^{j-1}_k + v^{j-1}_{k+Me_1}}{2} - \frac{1}{8}H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) \\
\hat{v}^{j,2}_{Mk+e_1} &= \frac{v^{j-1}_{k+e_1} + v^{j-1}_{k+e_2}}{2} - \frac{1}{8}H(\Delta^2_{Me_2} v^{j-1}_{k+e_2}, \Delta^2_{Me_2} v^{j-1}_{k+e_2-Me_2}), \qquad (21)
\end{aligned}$$

and $v^j_{Mk} = v^{j-1}_k$. The choice between the first and the second prediction is a function of the finite differences of order 3 in the direction Me_1 or Me_2. Note that

$$\frac{v^{j-1}_k + v^{j-1}_{k+Me_1}}{2} = \frac{v^{j-1}_{k+e_1} + v^{j-1}_{k+e_2}}{2} + \frac{1}{2}\Delta_{e_2}\Delta_{e_1} v^{j-1}_k,$$

which again means that the prediction (2) corresponds to a *Lipschitz-Linear* prediction operator assuming its linear part corresponds to prediction (1).

7. Convergence Theorems

In what follows, for two positive quantities A and B depending on a set of parameters, the relation $A \lesssim B$ implies the existence of a positive constant C, independent of the parameters, such that $A \leq CB$. Also $A \sim B$ means $A \lesssim B$ and $B \lesssim A$.

The convergence theorems are obtained by studying the difference operators associated to *Lipschitz-Linear* prediction operators. The existence of such difference operators is ensured by the following theorem:

Theorem 7.1. *Let S be a Lipschitz-Linear prediction operator of order $N+1$ then there exists a multi-dimensional local operator $S^{(N+1)}$ such that:*

$$\Delta^{N+1} S v = S^{(N+1)} \Delta^{N+1} v$$

Proof. Since S reproduces polynomials of degree N, the existence of $S^{(N+1)}$ was already proved in [21]. What is particular here is the form for the differences of order $N+1$:

$$\begin{aligned} \Delta^{N+1}(S v_{Mk+i}) &= \Delta^{N+1}(S_l v_{Mk+i}) + \Delta^{N+1}\Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}) \\ &= (S_l^{(N+1)})_i \Delta^{N+1} v_k + \Delta^{N+1}\Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}). \end{aligned}$$

From which, we deduce:

$$S_i^{(N+1)} w_k = (S_l^{(N+1)})_i w_k + \Delta^{N+1}\Phi_i(w_{k+p_1}, \cdots, w_{k+p_q})$$

□

Note that the previous theorem shows the existence of the operator for the differences of order k for all $k \leq N+1$. To study the convergence of the iteration (3), we introduce the definition of the joint spectral radius for difference operators:

Definition 7.1. *Let us consider a Lipschitz-Linear prediction operator S of order $N+1$. The joint spectral radius in $(\ell^p(\mathbb{Z}^d))^{r_k^d}$ of S^k, for $k \leq N+1$ is*

given by

$$\begin{aligned}\rho_p(S^k) &:= \inf_{j\geq 0} \|(S^k)^j\|^{1/j}_{(\ell^p(\mathbb{Z}^d))^{r_k^d}\to(\ell^p(\mathbb{Z}^d))^{r_k^d}} \\ &= \inf\{\rho, \|\Delta^k S^j v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \\ &\lesssim \rho^j \|\Delta^k v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}, \forall v \in \ell^p(\mathbb{Z}^d)\}. \end{aligned} \tag{22}$$

In all the theorems that follows $v_j(x) = \sum_{k\in\mathbb{Z}^d} v_k^j \varphi_{j,k}(x)$, where φ satisfies (4) with g associated to the linear prediction operator S_l (see (5)). We first need to establish some extensions to the non-separable case of results obtained in [20]:

Lemma 7.1. *Let S be a Lipschitz-linear prediction operator of order $N+1$. Then, for any $k \leq N+1$*

$$\|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p}\left(\|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} + \|e^{j+1}\|_{\ell^p(\mathbb{Z}^d)}\right). \tag{23}$$

Moreover, if $\rho > \rho_p(S^{(k)})$,

$$\|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \lesssim \rho^j \left(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j} \rho^{-l}\|e^l\|_{\ell^p(\mathbb{Z}^d)}\right). \tag{24}$$

Proof. Using the definition of functions $v_j(x)$ and the scaling equation (4), we get that $v_{j+1}(x) - v_j(x)$ is given by:

$$\begin{aligned} &= \sum_k v_k^{j+1}\varphi_{j+1,k}(x) - \sum_k v_k^j \varphi_{j,k}(x) \\ &= \sum_{i\in \text{ coset } (M)} \sum_k ((Sv^j)_{Mk+i} + e^{j+1}_{Mk+i})\varphi_{j+1,Mk-i}(x) - \sum_k v_k^j \sum_l g_{l-Mk}\varphi_{j+1,l}(x) \\ &= \sum_{i\in \text{ coset } (M)} \sum_k ((Sv^j)_{Mk+i} - \sum_l g_{M(k-l)+i} v_l^j)\varphi_{j+1,Mk-i}(x) + \sum_k e_k^{j+1}\varphi_{j+1,k}(x). \end{aligned}$$

Since S is a *Lipschitz-Linear* prediction operator of order $N+1$, we get:

$$\begin{aligned} \|\sum_k ((Sv^j)_{Mk+i} &- \sum_l g_{M(k-l)+i} v_l^j)\varphi_{j+1,Mk-i}(x)\|_{L^p(\mathbb{R}^d)} \\ &\lesssim m^{-j/p}\|\Phi_i(\Delta^{N+1} v^j_{\cdot+p_1}, \cdots, \Delta^{N+1} v^j_{\cdot+p_q})\|_{\ell^p(\mathbb{Z}^d)} \\ &\lesssim m^{-j/p}\|\bar{\Phi}_i(\Delta^k v^j_{\cdot+\bar{p}_1}, \cdots, \Delta^k v^j_{\cdot+\bar{p}_{\bar{q}}})\|_{\ell^p(\mathbb{Z}^d)} \lesssim m^{-j/p}\|\Delta^k v^j\|_{\ell^p(\mathbb{Z}^d)}. \end{aligned}$$

The proof of (23) is thus complete. Note that we have used

$$\|\Phi_i(\Delta^{N+1}v^j_{\cdot+p_1},\cdots,\Delta^{N+1}v^j_{\cdot+p_q})\|_{\ell^p(\mathbb{Z}^d)} = \|\bar{\Phi}_i(\Delta^k v^j_{\cdot+\bar{p}_1},\cdots,\Delta^k v^j_{\cdot+\bar{p}_{\bar{q}}})\|_{\ell^p(\mathbb{Z}^d)}, \tag{25}$$

where $\bar{\Phi}_i$ is Lipschitz with respect to its argument. This is due to the fact that higher order finite differences can be expressed as linear combinations of lower order differences. To prove (24), we note that:

$$\begin{aligned}
\|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} &\lesssim \|S^{(k)}\Delta^k v^{j-1}\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} + \|\Delta^k e^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \\
&\lesssim \rho\|\Delta^k v^{j-1}\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} + \|\Delta^k e^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \\
&\lesssim \rho^j\left(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j}\rho^{-l}\|e^l\|_{\ell^p(\mathbb{Z}^d)}\right).
\end{aligned}$$

□

Now, using the above lemma, we are able to prove:

Theorem 7.2. *Let S be a Lipschitz-Linear prediction operator of order $N+1$. Assume that $\rho_p(S^{(k)}) < m^{1/p}$, for some $k \leq N+1$ and that*

$$\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p}\|e^j\|_{\ell^p(\mathbb{Z}^d)} < \infty.$$

Then, the limit function v belongs to $L^p(\mathbb{R}^d)$ and

$$\|v\|_{L^p(\mathbb{R}^d)} \leq \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p}\|e^j\|_{\ell^p(\mathbb{Z}^d)} \tag{26}$$

Proof. From estimates (23) and (24) one has, in particular

$$\|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p}\rho^j\left(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j+1}\rho^{-l}\|e^l\|_{\ell^p(\mathbb{Z}^d)}\right) \tag{27}$$

for $\rho > \rho_p(S^{(k)})$. Considering k such that $\rho_p(S^{(k)}) < m^{1/p}$ and then by

choosing ρ such that $\rho_p(S^{(k)}) < \rho < m^{1/p}$ we get successively:

$$\begin{aligned}
\|v\|_{L^p(\mathbb{R}^d)} &\leq \|v_0\|_{L^p(\mathbb{R}^d)} + \sum_{j\geq 0} \|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \\
&\lesssim_{(27)} \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j\geq 0} m^{-j/p}\rho^j \left(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j+1} \rho^{-l} \|e^l\|_{\ell^p(\mathbb{Z}^d)} \right) \\
&\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} \left(\sum_{j\geq 0} (m^{-1/p}\rho)^j + 1 \right) + \sum_{l>0} \|e^l\|_{\ell^p(\mathbb{Z}^d)} \rho^{-l} \sum_{j>l} (m^{-1/p}\rho)^j \\
&\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l>0} \|e^l\|_{\ell^p(\mathbb{Z}^d)} m^{-l/p}.
\end{aligned}$$

□

Remark 7.1. *Usually the convergence in L^p is associated to the condition $\rho(S^{(1)}) < m^{1/p}$. With a Lipschitz - Linear prediction operator of order $N+1$, the convergence in $L^p(\mathbb{R}^d)$ is ensured provided $\rho(S^{(k)}) < m^{1/p}$, for some $k \leq N+1$. This remark is of interest since there is no link between upper bounds of $\rho(S^{(k)})$ and of $\rho(S^{(k+1)})$.*

The above remark, leads to a new inverse theorem in Besov spaces.

Theorem 7.3. *Let S be a Lipschitz-Linear prediction operator of order $N+1$. Assume that $\rho_p(S^{(k)}) < m^{1/p-s/d}$ for some $s \geq N$ and some $k \leq N+1$, and also that $(v^0, e^1, e^2, \ldots)$ satisfies*

$$\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j} \|(e^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j>0}\|_{\ell^q(\mathbb{Z}^d)} < \infty.$$

Then, the limit function v belongs to $B^s_{p,q}(\mathbb{R}^d)$ and

$$\|v\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j} \|(e^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j>0}\|_{\ell^q(\mathbb{Z}^d)}. \tag{28}$$

The proof of (28) is similar to that of Theorem 3 [21] replacing Δ^{N+1} by Δ^k, so we will not expand on this here. The novelty of the approach is, on the one hand, that the property on the spectral radius has to be verified only for some $k \leq N+1$ but not necessarily for $k = N+1$ and, on the other hand, that the prediction operator does not necessarily reproduce exactly polynomials (this is directly due to Lemma 7.1).

8. Stability in L^p and Besov spaces

In applications, the multi-scale data may be corrupted by some process. Since our model is nonlinear the inverse theorems does not ensure the stability. We develop here the stability results for our new nonlinear formalism. To this end, we consider two data sets $(v^0, e^1, e^2, \cdots)$ and $(\tilde{v}^0, \tilde{e}^1, \tilde{e}^2, \cdots)$ corresponding to two reconstruction processes:

$$v^j = Sv^{j-1} + e^j \text{ and } \tilde{v}^j = S\tilde{v}^{j-1} + \tilde{e}^j.$$

In that context, we recall the definition of v as the limit of $v_j(x) = \sum_{k\in\mathbb{Z}^d} v_k^j \varphi_{j,k}(x)$, with $\varphi_{j,k}(x) = \varphi(M^j x - k)$ (and similarly for $\tilde{v}$).

8.1. Stability in L^p spaces

First, we study the stability of the multi-scale representation in $L^p(\mathbb{R}^d)$, which is stated by the following theorem:

Theorem 8.1. *Let S be a Lipschitz-Linear prediction operator of order $N+1$, and suppose that $\rho_p(S^{(k)}) < m^{1/p}$ for some $k \leq N+1$. Assume that v_j and $\tilde{v}_j$ converge to v and $\tilde{v}$ in $L^p(\mathbb{R}^d)$ respectively and also that:*

$$\|\Delta^k(v^j - \tilde{v}^j)\|_{\ell^p(\mathbb{Z}^d)^{r_k^d}} \lesssim \rho^j \left(\|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=0}^{j-1} \rho^{-l} \|e^l - \tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)} \right) \quad (29)$$

holds for some $\rho_p(S^{(k)}) < \rho < m^{1/p}$. Then, we have:

$$\|v - \tilde{v}\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{L^p(\mathbb{R}^d)} + \sum_{l=1}^{j} \rho^{-l} \|e^l - \tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)} \quad (30)$$

Proof. Note that the difference $v_j - \tilde{v}_j$ is given by:

$$\sum_{i\in \text{ coset}(M), p\in\mathbb{Z}^d} \left(S_l(v^{j-1} - \tilde{v}^{j-1})_{Mp+i} + \bar{\Phi}_i(\Delta^k v^{j-1}_{p+\bar{p}_1}, \cdots, \Delta^k v^{j-1}_{p+\bar{p}_{\bar{q}}}) + (e^j_{Mp+i} - \tilde{e}^j_{Mp+i}) \right) \varphi_{j,Mp-i}(x),$$

where we have used the identity (25). It follows that $v_j - \tilde{v}_j$ is

$$
\begin{aligned}
&= \sum_{i\in\mathbb{Z}^d, p\in\mathbb{Z}^d} g_{k-Mi}(v_i^{j-1} - \tilde{v}_i^{j-1})\varphi_{j,p}(x) \\
&+ \sum_{i\in \text{ coset}(M), p\in\mathbb{Z}^d} \left(\tilde{\Phi}_i(\Delta^k v_{p+p_1}^{j-1}, \cdots, \Delta^k v_{p+p_q}^{j-1}) + (e_{Mp+i}^j - \tilde{e}_{Mp+i}^j)\right)\varphi_{j,Mp-i}(x) \\
&= \sum_{p\in\mathbb{Z}^d} (v_p^{j-1} - \tilde{v}_p^{j-1})\varphi_{j-1,p}(x) \\
&+ \sum_{i\in \text{ coset}(M), p\in\mathbb{Z}^d} \left(\tilde{\Phi}_i(\Delta^k v_{p+p_1}^{j-1}, \cdots, \Delta^k v_{p+p_q}^{j-1}) + (e_{Mp+i}^j - \tilde{e}_{Mp+i}^j)\right)\varphi_{j,Mp-i}(x).
\end{aligned}
$$

From this we deduce that:

$$
\|v_j - \tilde{v}_j\|_{L^p(\mathbb{R}^d)} \leq \|v_{j-1} - \tilde{v}_{j-1}\|_{L^p(\mathbb{R}^d)} + m^{-j/p}(\|\Delta^k(v^{j-1} - \tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} + \|e^j - \tilde{e}^j\|_{\ell^p(\mathbb{Z}^d)}),
$$

using the fact that $\bar{\Phi}_i$ is lipschitz and the inequality

$$
m^{(-j+1)/p}\|v^{j-1} - \tilde{v}^{j-1}\|_{\ell^p(\mathbb{Z}^d)} \leq \|v_{j-1} - \tilde{v}_{j-1}\|_{L^p(\mathbb{R}^d)}.
$$

Then using hypothesis (29), we get:

$$
\|v_j - \tilde{v}_j\|_{L^p(\mathbb{R}^d)} \leq \|v_{j-1} - \tilde{v}_{j-1}\|_{L^p(\mathbb{R}^d)} + m^{-j/p}\rho^j \left(\|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j} \rho^{-l}\|e^l - \tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}\right).
$$

Since $\rho m^{-1/p} < 1$:

$$
\|v - \tilde{v}\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l>0} \rho^{-l}\|e^l - \tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}
$$

□

Remark 8.1. *Similarly to the convergence theorem, to use a Lipschitz - Linear operator of order $N+1$ implies that the stability in $L^p(\mathbb{R}^d)$ is ensured as soon as $\rho(S^{(k)}) < m^{1/p}$ for some $k \leq N+1$. Let us stress that we have the same condition $\rho(S^{(k)}) < m^{1/p}$ for inverse and stability theorems (see Remark 7.1). Furthermore, we still not require the exact polynomial reproduction of degree N contrary to the results in [21].*

8.2. Stability in Besov spaces

In view of the inverse inequality (28), to show the stability, it seems natural to seek an inequality of type:

$$\|v-\tilde{v}\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\|(m^{(s/d-1/p)j}\|e^j_\cdot-\tilde{e}^j_\cdot\|_{\ell^p(\mathbb{Z}^d)})_{j>0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (31)$$

We now state without proof a stability theorem in Besov space $B^s_{p,q}(\mathbb{R}^d)$:

Theorem 8.2. *Let us assume that S is a Lipschitz-Linear prediction operator of order $N+1$ such that $\rho_p(S^{(k)}) < m^{1/p-s/d}$ for some $s > N$ and some $k \leq N+1$ and assume that*

$$\|\Delta^k(v^j-\tilde{v}^j)\|_{(\ell^p(\mathbb{Z}^d))^{r^d_k}} \lesssim \rho^j \left(\|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j} \rho^{-l}\|e^l-\tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)} \right) \quad (32)$$

for some $\rho_p(S^{(k)}) < \rho < m^{1/p-s/d}$. Then, the function $v-\tilde{v}$ belongs to $B^s_{p,q}(\mathbb{R}^d)$. Moreover, we obtain for that ρ:

$$\|v-\tilde{v}\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\|(\rho^{-j}\|(e^j_k-\tilde{e}^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j>0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (33)$$

The proof is the same as that of Theorem 3 of [21] except that we do not require the exact polynomial reproduction.

9. $(\mathcal{A}, I)$-Compatible Nonlinear Prediction Operators

Given families of multi-indices I and of vectors $\mathcal{A}$, we define:

$$\Delta^{\mathcal{A}_I} = \left\{ \Delta^{i_1}_{a_1} \cdots \Delta^{i_p}_{a_p},\ a_k \in \mathcal{A}, i_k \in I \right\}.$$

In other words, $\Delta^{\mathcal{A}_I}$ is a difference operator computed with respect to the family of vectors $\mathcal{A}$ and orders given by I. Then, we have the definition of $(\mathcal{A}, I)$-compatible nonlinear prediction operator:

Definition 9.1. *A nonlinear prediction operator S is called $(\mathcal{A}, I)$-compatible if there exists a local linear prediction operator S_l and if it satisfies*

$$(Sv)_{Mk+i} = (S_l v)_{Mk+i} + \Phi_i(\Delta^{\mathcal{A}_I} v_{k+p_1}, \cdots, \Delta^{\mathcal{A}_I} v_{k+p_q}) \quad \forall i \in \ coset(M)$$

where $\{p_1, \cdots, p_q\}$ is a fix set, Φ_i are Lipschitz functions and if there exists an operator $S^{\mathcal{A}_I}_l$ satisfying:

$$\Delta^{\mathcal{A}_I} S_l v = S^{\mathcal{A}_I}_l \Delta^{\mathcal{A}_I} v.$$

From its definition, the operator S admits an operator $S^{\mathcal{A}_I}$. We also remark that *Lipschitz - Linear* operators of order $N+1$ are $(\mathcal{A}, I)$-compatible with $I = \{i; |i| = N+1\}$ and $\mathcal{A} = \{e_1, \ldots, e_d\}$.

Note that we can extend all the notions described in the previous sections for *Lipschitz-Linear* prediction operators to $(\mathcal{A}, I)$-compatible prediction operators (i.e. multi-scale representation, joint spectral radius of $S^{\mathcal{A}_I}$, convergence and stability theorems). For instance, in Theorem 7.3, if the prediction operator is $(\mathcal{A}, I)$-compatible, then the result is true provided that $\rho_p(S^{\mathcal{A}_I}) < m^{1/p-s/d}$.

The interest of using the notion of $(\mathcal{A}, I)$-compatibility is to provide proofs of convergence where the classical approach fails, as shown in the next section. The $(\mathcal{A}, I)$-compatibility also enables to significantly reduce the number of computed differences in the study of joint spectral radius for convergence theorems. Note also that the compatibility notion is not related to polynomial reproduction for prediction operator, which makes it a new tool for analysis. From practical point of view, given a prediction operator we first identify its type (i.e. *Lipschitz-Linear* or $(\mathcal{A}, I)$-compatible) and then proceed to the analysis of the corresponding multi-scale representation.

10. Applications

10.1. Convergence of One-Dimensional Scheme

In one dimension, the notion of $(\mathcal{A}, I)$-compatibility does not make sense. Our point is to give an illustration of the new convergence Theorem 7.1 on the multi-scale representation based on PPH *Lipschitz-Linear* prediction operator. Since the PPH prediction operator is *Lipschitz-Linear* of order 2, the convergence in L^∞ occurs when $\rho_\infty(S^{(k)}) < 1$ for $k = 1$ or $k = 2$. It is shown in [4] Proposition 2, that the PPH-scheme is uniformly convergent because $\rho_\infty(S^{(1)}) \leq \frac{1}{2}$. However, the proof involves complex computations whereas the expression of $S^{(2)}$ is particularly simple:

$$\begin{aligned}
(S^{(2)}w)_{2i} &= \frac{1}{4}H(w_{i-1}, w_i) \\
(S^{(2)}w)_{2i+1} &= \frac{w_i}{2} - \frac{1}{8}\left(H(w_{i-1}, w_i) + H(w_i, w_{i+1})\right).
\end{aligned}$$

As in [5], since $|H(x, y)| \leq |\max(x, y)|$, we immediately get: $\rho_\infty(S^{(2)}) < \frac{3}{4}$, which is a much simpler proof for L^∞ convergence.

Based on this simple expression of $S^{(2)}$ we propose a new proof of L^p convergence. Indeed, we may write assuming that $p \geq 1$:

$$\begin{aligned}
|(S^{(2)}w)_{2i}|^p &\leq \frac{1}{4^p}\max(|w_{i-1}|,|w_i|)^p \\
|(S^{(2)}w)_{2i+1}|^p &\leq \left(\frac{1}{2}|w_i| + \frac{1}{8}\max(|w_{i-1}|,|w_i|) + \frac{1}{8}\max(|w_i|,|w_{i+1}|\right)^p \\
&\leq \left(\frac{1}{2}|w_i| + \frac{1}{4}(\frac{1}{2}\max(|w_{i-1}|,|w_i| + \frac{1}{2}\max(|w_i|,|w_{i+1}|)\right)^p \\
&\leq \frac{1}{2}|w_i|^p + \frac{1}{4}\frac{1}{2^p}\max(|w_{i-1}|,|w_i|)^p + \frac{1}{4}\frac{1}{2^p}\max(|w_i|,|w_{i+1}|)^p.
\end{aligned}$$

The last inequality being obtained because we have a convex combination. Now, to obtain an upper bound, we note:

$$\sum_{i\in\mathbb{Z}} \frac{1}{4^p}\max(|w_{i-1}|,|w_i|)^p + \frac{1}{2}|w_i|^p + \frac{1}{4}\frac{1}{2^p}\max(|w_{i-1}|,|w_i|)^p + \frac{1}{4}\frac{1}{2^p}\max(|w_i|,|w_{i+1}|)^p.$$

The larger coefficient in the sum is obtained when $|w_i|$ is larger than $|w_{i-1}|$ and $|w_{i+1}|$, leading to the coefficient $\frac{1}{4^p} + \frac{1}{2} + \frac{1}{2}\frac{1}{2^p}$. The multi-scale is L^p convergent for $p \geq 1$, if p satisfies $\frac{1}{2} + \frac{1}{2}\frac{1}{2^p} + \frac{1}{4^p} < 2$, which is true for any $p \geq 1$.

10.2. Convergence of Bidimensional PPH Multi-scale Representations

We study the convergence of bidimensional PPH multi-scale representations with prediction operator given by (34). We already noticed that the prediction operator is *Lipschitz-Linear*, we now notice that it is $(\mathcal{A}, I)$-compatible with $\mathcal{A} = \{e_1, Me_1\}$ and $I = \{(0,2),(2,0)\}$, where M is the quincunx matrix. Therefore, to prove the convergence we only study the joint spectral radius of $S^{\mathcal{A}_I}$. To this end, we compute the differences of order 2 in the directions $\{e_1, Me_1\}$, which are given by:

$$\begin{aligned}
\Delta^2_{e_1} v^j_{Mk} &= \frac{1}{4}H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) \\
\Delta^2_{e_1} v^j_{Mk+e_1} &= \frac{1}{2}\Delta^2_{Me_1} v^{j-1}_k - \frac{1}{8}\left(H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) + H(\Delta^2_{Me_1} v^{j-1}_{k+Me_1}, \Delta^2_{Me_1} v^{j-1}_k)\right) \\
\Delta^2_{Me_1} v^j_{Mk} &= \Delta^2_{e_1} v^{j-1}_k \\
\Delta^2_{Me_1} v^j_{Mk+e_1} &= \frac{1}{2}(\Delta^2_1 v^{j-1}_k + \Delta^2_1 v^{j-1}_{k+Me_1}) + \frac{1}{4}H(\Delta^2_{Me_1} v^{j-1}_{k+e_1}, \Delta^2_{Me_1} v^{j-1}_{k+e_1-Me_1}) \\
&- \frac{1}{8}\left(H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) + H(\Delta^2_{Me_1} v^{j-1}_{k+2e_1}, \Delta^2_{Me_1} v^{j-1}_{k+2e_1-Me_1})\right).
\end{aligned}$$

We then compute $\rho_\infty(S^{\mathcal{A}_I})$. Remarking as previously that $|H(x,y)| \leq \max(|x|,|y|)$ and also that this upper bound is attained for $x = y$, we get $\rho_\infty(S^{\mathcal{A}_I}) = \sqrt{(\frac{1}{2}+\frac{1}{4})(1+\frac{1}{2})} = \sqrt{\frac{9}{8}} > 1$. The first term of the product corresponds to the difference in the direction e_1, while the second to the direction Me_1. It is essential to note that the direction e_1 at level j becomes the direction Me_1 at level $j+1$ which makes, in this case, the computations particularly simple.

Going further, we modify the PPH scheme defined in (34) as follows

$$\hat{v}^j_{Mk+e_1} = \begin{cases} \frac{v^{j-1}_k + v^{j-1}_{k+Me_1}}{2} - \omega H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) \\ \hat{v}^j_{Mk} = v^{j-1}_k, \end{cases} \tag{34}$$

where $0 < \omega < \frac{\sqrt{2}-1}{4}$. With this choice for ω, the convergence in L^∞ of the multi-scale representation immediately follows. Note that it does not involve the mixed differences $\Delta_{e_1}\Delta_{Me_1}$.

For the L^p convergence, we study:

$$a := \sum_{k\in\mathbb{Z}^2} (\Delta^2_{Me_1} v^j_{Mk+e_1})^p + (\Delta^2_{Me_1} v^j_{Mk})^p + (\Delta^2_{e_1} v^j_{Mk})^p + \Delta^2_{Me_1} v^j_{Mk+e_1})^p. \tag{35}$$

As in the one-dimensional study, straightforward computations give the following upper bound for a:

$$(\frac{1}{2} + 2\frac{1}{2^p}) \sum_{k\in\mathbb{Z}^2} (\Delta^2_1 v^j_k)^p + (6\frac{1}{8^p} + 4\frac{1}{4^p} + \frac{1}{2^p} + (\frac{3}{4})^p) \sum_{k\in\mathbb{Z}^2} (\Delta^2_{Me_1} v^j_k)^p.$$

Recalling that $m = 2$, we get a L^p convergent representation when $\rho_p(S^{\mathcal{A}_I}) < 2^{\frac{1}{p}}$, for which a sufficient condition is:

$$\max(\frac{1}{2} + 2\frac{1}{2^p}, 6\frac{1}{8^p} + 4\frac{1}{4^p} + \frac{1}{2^p} + (\frac{3}{4})^p) < 2.$$

This is always true when $p \geq 1.37$. To get the convergence property for any p, it suffices to make the same reasoning as for L^∞ convergence, and replace the factor $\frac{1}{8}$ in the definition of the PPH scheme by an appropiate ω.

11. Conclusion

In this paper, we have introduced a new formalism for nonlinear and non-separable multi-scale representations. The introduced formalism includes

classical nonlinear multi-scale representations such as ENO and those based on PPH or power-P schemes. In our context, the nonlinear prediction operators are perturbations of some linear prediction operator. These perturbations are modeled by Lipschitz functions depending on finite differences whose order depends on the degree of the polynomials reproduced by the linear prediction operator plus one. In the paper, we called these particular kind of prediction operators *Lipschitz-Linear*. We gave several illustrations in one and multi-dimensional cases of such prediction operators and also enounced the convergence and stability theorems associated to these representations. We put forward new conditions on the joint spectral radii of difference operators for both convergence and stability. We also introduced the notion of $(\mathcal{A}, I)$-compatible prediction operators which behaves like *Lipschitz-Linear* ones in terms of convergence and stability. We saw in applications that the $(\mathcal{A}, I)$-compatibility of prediction operators sometimes makes the proofs of convergence easier. In spite the convergence and stability theorems apply to non-interpolatory prediction operators, future works should involve the search for convergent examples of non-interpolatory multi-scale representations based on either *Lipschitz-Linear* or $(\mathcal{A}, I)$-compatible prediction operators.

[1] E. H. Adelson, E. P. Simoncelli, and W. T. Freeman, Pyramids for Early Vision, Representations of Vision, Cambridge University Press, 3-16, (1991).

[2] S. Amat, S. Busquier, J.C Trillo, Nonlinear Harten's Multiresolution on the Quincunx Pyramid J. Comput. Appl. Math., 189, 555-567, (2003).

[3] Amat, S., F. Arandiga, A. Cohen, R. Donat, G. Garcia, and M. Von Oehsen Data compression with ENO Schemes: a case study Appl. Comput. Harmon. Anal. , (1999).

[4] S. Amat, R. Donat, J. Liandrat, J.C. Trillo, Analysis of a New Non-Linear Subdivision Scheme: Applications to Image Processing, Found. Comput. Math., 6, 193-226 (2006).

[5] S. Amat and J. Liandrat, On the stability of the PPH Nonlinear Multiresolution Applied and Computational Harmonc Analysis, 18, 198-206, (2006).

[6] A. Cohen Wavelets in numerical analysis, Handbook of Numerical Analysis, VII, P.G. Ciarlet and J.L. Lions, eds., Elsevier, Amsterdam, (2003).

[7] A. Cohen, N. Dyn and B. Mateï, Quasi-linear Subdivision Schemes with Applications to ENO Interpolation, Appl. Comput. Harmon. Anal., 15, 89-116, (2003).

[8] V. Chappelier and C. Guillemot, Oriented Wavelet for Image Compression and Denoising, IEEE Transactions on Image Processing, 15, 2892-2903, (2006).

[9] T. F. Chan, H.M. Zhou, ENO-Wavelet Transforms for Piecewise Smooth Functions, SIAM J. Numer. Anal., 40, 1369-1404, (2002).

[10] Daubechies, I., Runborg, O., Sweldens, W., Normal multiresolution approximation of curves, Constr. Approx. 20, 399-463, (2004).

[11] N. Dyn, Subdivision Scheme in CAGD, Advances in numerical analysis II, Wavelets, Subdivision Algorithms and Radial Basis Functions W. Light (ed.) Clarendon Press, Oxford, 36-104, (1992).

[12] N. Dyn, J. A. Gregory and D. Levin, A four-point interpolatory subdivision scheme for curve design, Computer Aided Geometric Design 4, 257-268, (1987).

[13] N. Dyn and P. Oswald, Univariate Subdivision and Multi-scale Transforms: the nonlinear case, in Multiscale, Nonlinear and Adaptive Approximation, Springer, 203-249, (2009).

[14] A. Harten, Discrete Multiresolution Analysis and Generalized Wavelets, J. Appl. Num. Math. 12,153-193, (1993) .

[15] A. Harten, B. Enquist, S. Osher and S. Chakravarthy, Uniformly high order accurate essentially non-oscillatory schemes III, J. Comput. Phys. 71, 231-303, (1987).

[16] Harten, A. ENO Schemes with Subcell Resolution, Journal of Applied Numerical Mathematics, 12, 153-193, (1993).

[17] R-Q. Jia, Characterization of Smoothness of Multivariate Refinable Functions in Sobolev Spaces, Transactions on the American Mathematical Society, 49, 4089-4112, (1999).

[18] Kompatsiaris, I. Tzovaras D., Strintzis M.G., Hierarchical Representation and Coding of Surfaces Using 3D Pollygon Meshes, IEEE Transactions on Image Processing, 10, 1133-1151, (2001).

[19] Malassiotis S., Strintzis M.G., Optimal Biorthogonal Wavelet Decomposition of Wire-Frame Meshes Using Box Splines and its Application to the Hierarchical Coding of 3D Surfaces, IEEE Transactions on Image Processing, 8, 41-57, (1999).

[20] B. Mateï, Smoothness Characterization and Stability in Nonlinear multiscale Framework: Theoretical Results Asymptotic Analysis, 46, 277-309, (2005).

[21] B. Mateï, S. Meignen and A. Zakarova, Smoothness Characterization and Stability of Nonlinear and Non-Separable multi-scale Representation, submitted, ref. hal-0047176.

[22] B. Mateï, Méthodes Multirésolutions Non-Linéaires, Applications au Traitement d'Images Thèse de Doctorat de L'université Paris VI, (2002).

[23] Oswald, P., Smoothness of nonlinear median-interpolation subdivision, Adv. Comput. Math. 20, 401-423, (2004).

[24] T. Sauer and Y. Xu, On Multivariate Lagrange Interpolation, Mathematics of Computation, 64, 211, 1147-1170, (1995).

[25] T. Sauer and Y. Xu, A Case Study in Multivariate Lagrange Interpolation, Preprint 2004.

[26] S. Serna and A. Marquina, Power ENO Methods: a Fifth-order Accurate Weighted Power ENO method, Journal of Computational Physics, 194, 632-658, (2004).