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-00529531v2

HAL Id: hal-00529531
https://hal.science/hal-00529531v2

Preprint submitted on 31 Dec 2010

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 Multiscale Representations

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

We present a new formalism for nonlinear and non-separable multiscale representations. We show that the most commonly used one-dimensional nonlinear multiscale representations can be defined using prediction operators which are the sum of a linear prediction operator and of 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 that we propose allows us to provide simple proofs of stability and convergence of the multiscale representations.

Keywords: Nonlinear multiscale approximation, Besov spaces, Stability.
2000 MSC: 41A46, MSC 41A60, MSC 41A63

1. Introduction

Multiscale algorithms such as wavelet-type pyramid transforms for hierarchical data representation [1] and subdivision methods for computer-aided geometric design [11] have completely changed the domains of data and geometry processing. Linear multiscale representations or equivalently wavelet

expansions of functions are now well understood in terms of approximation performance and limitations are well known [7]. While in the univariate case the wavelet-type pyramid transforms provide optimal algorithms [7], 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). This is directly reflected by the poor decay $\mathcal{O}(N^{-1/2})$ of the L^2 error of the best N-term approximation for a cartoon image, i.e. the characteristic function of bounded domain having smooth boundary. Improving this rate through a better representation of images near edges has motivated the study of ridgelets [5], curvelets [6] and bandelets [17]. These are bases or frames allowing anisotropic refinement close to the edges. Nonlinear multiscale representations are another possibility [8] to perform anisotropic refinement and are closely related to linear ones. In contrast to linear ones, here the detail coefficients are not computed using a linear rule, but a data dependent rules. These data dependent rules define a nonlinear prediction operator. Let us mention, in a nutshell, the main difference between ridgelets, curvelets or bandelets and the nonlinear multiscale representations: to define these bases/frames a functional point of view is used while to define the nonlinear multiscale representations a discrete point of view is adopted.

In some sense, these bases/frames representations inherit all the good functional properties of wavelet basis since they are wavelet basis having an angular selectivity. On the contrary, the analysis of prediction operators require to define a new and different mathematical framework. From the mathematical point of view, the analysis of the nonlinear multiscale representations is related to the analysis of subdivision algorithms.

Roughly speaking, the development of the theory of nonlinear prediction operators has enabled to design four kinds of data-dependent multiscale representations: the first one are quasi-linear multiscale representations for piecewise smooth functions [19], the second one are median-interpolating schemes [21], the third one are normal multi-resolutions of curves and surfaces [10] and the fourth one are PPH or power-P representations [4].

The quasi-linear multiscale 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 [8]. The PPH scheme, introduced in [4] to design multiscale representations, was also motivated by a better treatment of edges for image compression. The median-interpolating scheme was motivated by applications to non-Gaussian noise removal (see [21]), while the normal multi-

resolution was defined in [10] for optimal geometry compression of curves and surfaces. These last two representations are examples of nonlinear and multiscale geometric transforms.

In general, the analysis of such nonlinear multiscale representations naturally extends the results obtained in the linear case by studying 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 operators that enables to embed classical WENO (Weighted Essentially Non Oscillatory) prediction operators [19], and the PPH scheme [3]. 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 will call *Lipschitz-Linear* these nonlinear prediction operators.

After having introduced some notations and the definition of *Lipschitz-Linear* prediction operators (section 2 and section 4), we show that the WENO prediction operator, the PPH scheme and a modified version of the power-P scheme fit into that framework (section 5). These one-dimensional prediction operators are based on dyadic scales and naturally extend to the multi-dimensional case by tensor product [19]. However, in applications, it may be of interest to define multiscale 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 [9],[18] and [16]. For that reason, we study the extension of the *Lipschitz-Linear* formalism for prediction operators to the multi-dimensional case and where the scales are defined using non-diagonal dilation matrices (section 6). Sections 7 and 8 are then devoted to the study of the convergence and stability of nonlinear multiscale representations based on *Lipschitz-Linear* prediction operators both in L^p and Besov spaces. A new aspect is then introduced in section 9, through the notion of prediction operators compatible with a set of differences, which proves to be interesting in practice. The results on the convergence and the stability of the corresponding multiscale representations are identical to those based on *Lipschitz-Linear* prediction operators. In section 10, we conclude the paper showing some new results on the stability and the convergence in L^p of some nonlinear multiscale representations associated with *Lipschitz-Linear* prediction operators.

2. Notations

Before we start, we need to introduce some standard multi-index notations. For $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}^d$, 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 of degree $|\alpha|$. 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 $\deg(p)$ the degree of any polynomial p. By $(e_1, \cdots, e_d)$ we denote the canonical basis on $\mathbb{Z}^d$. 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 $\Delta_a^{\alpha_d} v_k$, for any vector a in $\mathbb{Z}^d$ is defined recursively by:

$$\Delta_a^{\alpha_d} v_k = \Delta_a^{\alpha_d - 1} v_{k+a} - \Delta_a^{\alpha_d - 1} v_k.$$

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. Multiscale Representations

We assume that the data v^j, associated to some grids Γ^j, $j \geq 0$ are given. We also assume the existence of a prediction operator S which 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 multiscale 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 $\varphi(M^j x - k)$ with φ some compactly supported 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 multiscale representation $(v^0, (e^j)_{j\geq 0})$ is convergent in that space. The construction of the multiscale representation is thus based on the definition of the prediction operator.

4. *Lipschitz-Linear* Prediction Operators

A prediction operator is a map $v \in \ell^\infty(\mathbb{Z}^d) \mapsto Sv \in \ell^\infty(\mathbb{Z}^d)$. Let us first recall the definition of a local and linear prediction operator S_l which is:

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

where the sequence $(g_n)_{n\in\mathbb{Z}^d}$ defining S_l satisfies $\exists K > 0$ such that $g_{k-Ml} = 0$ if $\|k - Ml\|_\infty > K$. In the following, we assume that the linear prediction operator S_l constructed trough the sequence $(g_n)_{n\in\mathbb{Z}^d}$ 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.

For our study we first need to recall the polynomial reproduction property.

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.

Note that the linear prediction operator shall satisfy polynomial reproduction property. 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.*

We recall that in [13], the concept of one-dimensional prediction operators that are local, r-shift invariant and off-set invariant for polynomials of degree N is introduced. It consists in prediction operators S defined for any $v \in \ell^\infty(\mathbb{Z}^d)$ as follows:

$$Sv_{rk+i} = \Psi_i(v_{k+p_1}, \cdots, v_{k+p_2}) \tag{6}$$

where $r \in \mathbb{N}$ and Ψ_i is a Lipschitz function, S being also such that for any polynomial sequence p of degree N there exists a polynomial q of degree at most $N-1$ satisfying:

$$(S(v+p))_i = (Sv)_i + p(i/r) + q(i) \quad i \in \mathbb{Z}. \tag{7}$$

It is clear that the above notion can easily be extended to the multi-dimensional case through the concept of local, M-shift and off-set invariant prediction operators, replacing r by the matrix M. It is then easy to check that *Lipschitz-Linear* prediction operators are local, M-shift invariant and off-set invariant for polynomial of degree N. Indeed, to show the off-set invariance for polynomial sequences p of degree N, we write for all $v \in \ell^\infty(\mathbb{Z}^d)$:

$$\begin{aligned}
(S(v+p))_{Mk+i} &= (S_l(v+p))_{Mk+i} + \Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}) \\
(S(v+p))_{Mk+i} &= (S_l(v))_{Mk+i} + \Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}) \\
&+ p(k + M^{-1} i) + q(k) \\
(S(v+p))_{Mk+i} &= (S(v))_{Mk+i} + p(k) + q_i(k).
\end{aligned}$$

Lipschitz-Linear prediction operators make up a sub-class of local, M-shift and off-set invariant prediction operators. To consider such a restriction on prediction operators will allow us to write simpler proofs of convergence and stability for the multiscale representations.

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: (i) point-value, where v_k^j are points values of some function v, $v_k^j = v(2^{-j}k)$ and (ii) cell-average, where v_k^j are the average of some function v over a neighborhood of $2^{-j}k$. Assuming a certain type of discretization, we define a nonlinear prediction operator that in turn leads to a nonlinear multiscale representation. We call them point-value (resp. cell-average) *multiscale representation*.

Let us now recall some useful properties on Lagrange interpolation. Consider the interpolation polynomial p_N of degree N of v at $x_0, \cdots, x_N$ and $p_{N,1}$ the interpolation polynomial of v at $x_1, \cdots, x_{N+1}$. Using standard arguments, we write the difference between the two polynomials as:

$$p_{N,1}(x) - p_N(x) \quad = \quad \Delta^{N+1} v_0 \frac{1}{N! h^N} \prod_{i=1}^{N} (x - x_i), \tag{8}$$

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

5.2. Quasi-Linear Prediction Operators Using Polynomial Interpolation

Here, we use identity (8) to analyze nonlinear prediction operators in the context of point-value *multiscale 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 that framework, since $v_{2k}^j = v_k^{j-1}$ (also called *consistency* property), only $\hat{v}_{2k+1}^j$ needs to 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 $\hat{v}^j_{2k+1}$ through the so-called *centered* prediction as follows:

$$\hat{v}^j_{2k+1} = p_{2N+1}(2^{-j}(2k+1)). \tag{9}$$

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

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

which is exact for cubic polynomials. The four point scheme was widely studied in literature (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:

$$\hat{v}^j_{2k+1,1} := p_{3,1}(2^{-j}(2k+1)) = \frac{5}{16}v^{j-1}_k + \frac{15}{16}v^{j-1}_{k+1} - \frac{5}{16}v^{j-1}_{k+2} + \frac{1}{16}v^{j-1}_{k+3}. \tag{10}$$

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

$$\hat{v}^j_{2k+1,1} - \hat{v}^j_{2k+1} = \frac{1}{16}\Delta^4 v^{j-1}_{k-1}, \tag{11}$$

which corresponds to (8), with $x_i = 2^{-j+1}(k+i-1)$, $i = 0, \cdots, 2$ and $x = 2^{-j}(2k+1)$.

The same conclusion holds 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 can generalize the above formula to any N through the following proposition:

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

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

Proof. Let us put $x_0 = 2^{-j+1}(k-N), \cdots, x_{2N+1} = 2^{-j+1}(k+N+1)$. Then, using (8) the difference between p_{2N+1} and $p_{2N+1,1}$ evaluated at $2^{-j}(2k+1)$, reads as follows:

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

□

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:*

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

and then apply Proposition 5.1.

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

We now show how Proposition 5.1 extends to cell-average *multiscale representations*. In the cell-average setting, the data v^j_k is the average of some function v over the interval $I_{j,k} = [2^{-j}k, 2^{-j}(k+1)]$ as follows:

$$v^j_k = 2^j \int_{I_{j,k}} v(t)dt \tag{12}$$

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

$$v^{j-1}_k = \frac{1}{2}(v^j_{2k} + v^j_{2k+1}). \tag{13}$$

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

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

We then define the *centered* prediction by:

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

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

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

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:

$$\hat{v}^j_{2k+1,1} - \hat{v}^j_{2k+1} = -\frac{1}{8}\Delta^3 v^{j-1}_{k-1} \text{ and } \hat{v}^j_{2k,1} - \hat{v}^j_{2k} = \frac{1}{8}\Delta^3 v^{j-1}_{k-1}. \tag{14}$$

Similarly, we can define a prediction using the set of intervals shifted to the left and obtain the same kind of result. The equality (14) 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:*

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

The proof is available in Appendix A.

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.

5.4. WENO-Prediction as Lipschitz-Linear Prediction Operator

Given a type a multiscale representation (i.e. either point-value or cell-average), the ENO prediction consists in predicting at a given point using a polynomial $p_{2N+1,l}$ (in the point value case or $p_{2N,l}$ in the cell average case) defined on one of the potential stencils. This kind of prediction operator introduced by Harten in [15] is known to be numerically unstable. The instability of the method can be overcome by the use of weighted-ENO (WENO) interpolation which provides a smooth transition between different prediction rules . The WENO formulation is based on a convex combination of the potential prediction operators given by the ENO method, that is we write: $\hat{v}^j_k := \sum_{r=-m,r\neq 0}^{m} \alpha^j_{r,k}\hat{v}^j_{k,r}$ with $\alpha^j_{r,k} \geq 0$ and $\sum_{r=-m,r\neq 0}^{m} \alpha^j_{r,k} = 1$. The weights depends on v^{j-1} and on the corresponding rule r.

As an illustration, let us consider the point values setting (using cubic polynomials to interpolate), for which we have:

$$\hat{v}^j_{2k+1} - v^j_{2k+1} \;=\; \frac{1}{16}\left(\alpha^j_{1,k}\Delta^4 v^{j-1}_{k-1} + \alpha^j_{-1,k}\Delta^4 v^{j-1}_{k-2}\right).$$

To fit into the proposed *Lipschitz-Linear* model, we consider that $\alpha^j_{1,k}$ is a function $\alpha(\Delta^4 v^{j-1}_{k-2}, \Delta^4 v^{j-1}_{k-1})$ and that $\alpha^j_{-1,k}$ equals to $1-\alpha^j_{1,k}$. We then determine $\alpha^j_{1,k}$ such that $\alpha(x,y)(x-y)$ is a Lipschitz function which implies that α is bounded on $\mathbb{R}^2$ (which is always the case since we consider a convex combination) and that $(x-y)(\frac{\partial\alpha}{\partial x}(x,y), \frac{\partial\alpha}{\partial y}(x,y))$ is bounded on $\mathbb{R}^2$. A typical example of such function is when $\alpha(x,y) = \frac{1}{1+(\frac{y}{x})^\beta}$, where β is some even integer larger than 2. The motivation for such a weight function is that it favors the smoothest prediction operator that is the one based on the least oscillating polynomial: if $\Delta^4 v^{j-1}_{k-1}$ is small compared to $\Delta^4 v^{j-1}_{k-2}$ the weight $\alpha^j_{1,k}$ should be close to 1 and to zero in the opposite case. This model corresponds to a small change in the traditional WENO method and it preserves its main properties as we will see in the application section.

5.5. PPH scheme as Lipschitz-Linear Prediction Operator

We first show that PPH is an example of *Lipschitz-Linear* prediction operator ([3]). The PPH scheme is defined by:

$$\left\{\begin{array}{c} \hat{v}^j_{2k+1} = \frac{v^{j-1}_{k+1}+v^{j-1}_k}{2} - \frac{1}{8}H(\Delta^2 v^{j-1}_{k-1}, \Delta^2 v^{j-1}_k) \\ \hat{v}^j_{2k} = v^{j-1}_k \end{array}\right. \tag{15}$$

where $H(x,y) := 2\left(\frac{xy}{x+y}\right)\chi_{\{xy>0\}}(x,y)$, where χ_X is the characteristic function of X. Since H satisfies $|H(x,y)-H(x',y')| \le 2\max\{|x-x'|, |y-y'|\}$, it is Lipschitz with respect to (x,y) and, since the linear scheme $\frac{v^{j-1}_{k+1}+v^{j-1}_k}{2}$ reproduces polynomials of degree 1, the PPH-scheme is a *Lipschitz-Linear* prediction operator of order 2. The power-P scheme [22], is a generalization of the PPH-scheme replacing H by

$$H_q(x,y) \;=\; \left(\frac{x+y}{2}\left(1-\left|\frac{x-y}{x+y}\right|^q\right)\right)\chi_{\{xy>0\}}(x,y). \tag{16}$$

The main difference between the PPH and the power-P scheme is that $H_q(x,y)$ is not a Lipschitz function but is only piecewise Lipschitz as remarked in [13]. This scheme does not therefore correspond to a *Lipschitz-Linear* prediction operator. However, a careful look at the model shows that

the power-P scheme is very close to a *Lipschitz-Linear* prediction operator. Indeed, consider the following definition for $x \neq y$:

$$\tilde{H}_q(x,y) = \left(\frac{x+y}{2}\left(1-\left|\frac{x-y}{x+y}\right|^q\right)\right) \times (\rho_\varepsilon * \chi_{\{xy>0\}})(x,y), \quad (17)$$

where $\rho_\varepsilon > 0$ is a $C^\infty(\mathbb{R}^2)$ compactly supported function with support embedded in $B(0,\varepsilon)$, the ball with center $(0,0)$ and with radius ε, and such that $\int \rho_\varepsilon = 1$. It is clear that $\tilde{H}_q(x,y) = H_q(x,y)$ as soon as (x,y) does not belong to the set

$$V_\varepsilon = \{|x| \leq \varepsilon\} \bigcup \{|y| \leq \varepsilon\}. \quad (18)$$

Note that $\frac{x+y}{2}\left(1-|\frac{x-y}{x+y}|^q\right)$ is differentiable for $x \neq y$, and that this differential is bounded (see Lemma 3.6 of [13] for the computation). Then $\tilde{H}_q(x,y)$ is Lipschitz when $x \neq y$. By taking into account the definition set for H_q, we deduce that it is Lipschitz on $\mathbb{R}^2 \setminus \{(x,x), |x| \leq \sqrt{2}\varepsilon\}$. Finally, since ε can be chosen arbitrarily small the new model is very close to the original power-P scheme. The two models differ on an arbitrarily small band depending on ε.

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 [9] and [18]). Having defined the grid $\Gamma^j = \left\{M^{-j}k, \quad k \in \mathbb{Z}^d\right\}$ using a dilation matrix M, one considers discrete quantities v_k^j defined on each of these grids. A typical example of this is the bidimensional PPH-scheme, associated to the quincunx dilation matrix, i.e.:

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (19)$$

and where the prediction is defined by:

$$\begin{aligned} \hat{v}^j_{Mk+e_1} &= \frac{v_k^{j-1} + v_{k+Me_1}^{j-1}}{2} - \frac{1}{8}H(\Delta^2_{Me_1}v_k^{j-1}, \Delta^2_{Me_1}v_{k-Me_1}^{j-1}) \\ \hat{v}^j_{Mk} &= v_k^{j-1}. \end{aligned} \quad (20)$$

Note that the linear part of the prediction operator is obtained by considering an affine interpolation polynomial at $v_k^{j-1}, v_{k+e_1}^{j-1}$ and $v_{k+e_1+e_2}^{j-1}$ 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, namely:

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

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, namely $|\Delta_{Me_1}^3 v_{k-Me_1}^{j-1}|$ and $|\Delta_{Me_2}^3 v_{k+e_2-Me_2}^{j-1}|$.

Since

$$\frac{v_k^{j-1}+v_{k+Me_1}^{j-1}}{2} = \frac{v_{k+e_1}^{j-1}+v_{k+e_2}^{j-1}}{2} + \frac{1}{2}\Delta_{e_2}\Delta_{e_1}v_k^{j-1},$$

both prediction operators are perturbation of the linear prediction $\frac{v_{k+e_1}^{j-1}+v_{k+e_2}^{j-1}}{2}$, the perturbation being a function of the second order differences $\Delta_{e_1}^2, \Delta_{e_2}^2$ and $\Delta_{e_1}\Delta_{e_2}$. Therefore we can rewrite the model as follows:

$$\hat{v}_{Mk+e_1}^{j} = \frac{v_{k+e_1}^{j-1}+v_{k+e_2}^{j-1}}{2} + \Phi_1((\Delta_{e_1}^2 v_{k+p}^{j-1}, \Delta_{e_2}^2 v_{k+p}^{j-1}, \Delta_{e_1}\Delta_{e_2} v_{k+p}^{j-1})_{p\in V})$$

where V is a fixed neighborhood of $(0,0)$ in $\mathbb{Z}^2$ and where Φ_1 is defined as follows: if $|\Delta_{Me_1}^3 v_{k-Me_1}^{j-1}| > |\Delta_{Me_2}^3 v_{k+e_2-Me_2}^{j-1}|$ then

$$\Phi_1((\Delta_{e_1}^2 v_{k+p}^{j-1}, \Delta_{e_2}^2 v_{k+p}^{j-1}, \Delta_{e_1}\Delta_{e_2} v_{k+p}^{j-1})_{p\in V}) = \frac{1}{2}\Delta_{e_1}\Delta_{e_2}v_k^{j-1} - \frac{1}{8}H(\Delta_{Me_1}^2 v_k^{j-1}, \Delta_{Me_1}^2 v_{k-Me_1}^{j-1}) \quad (22)$$

otherwise

$$\Phi_1((\Delta_{e_1}^2 v_{k+p}^{j-1}, \Delta_{e_2}^2 v_{k+p}^{j-1}, \Delta_{e_1}\Delta_{e_2} v_{k+p}^{j-1})_{p\in V}) = -\frac{1}{8}H(\Delta_{Me_2}^2 v_{k+e_2}^{j-1}, \Delta_{Me_2}^2 v_{k+e_2-Me_2}^{j-1}). \quad (23)$$

We prove that the model defined in (38) is not *Lipschitz-Linear* since Φ_1 is discontinuous and therefore cannot be Lipschitz. Indeed, consider v^{j-1}

such that $\Delta^2_{Me_1} v_k^{j-1} = \Delta^2_{Me_1} v_{k-Me_1}^{j-1} = \Delta^2_{Me_2} v_{k+e_2}^{j-1} = \Delta^2_{Me_2} v_{k+e_2-Me_2}^{j-1}$, and assume that $\Delta_{e_1}\Delta_{e_2} v_k^{j-1}$ is non-zero. In such a case, the prediction (23) is used. Now, consider an ϵ-perturbation of v^{j-1} such that the prediction (22) is used. Since H is a Lipschitz function, thus continuous, and by considering the just defined ϵ-perturbation, the function Φ_1 jumps by $\frac{1}{2}\Delta_{e_1}\Delta_{e_2} v_k^{j-1}$. For that reason, the function Φ_1 is discontinuous and also not Lipschitz. The proposed method, called quincunx-PPH, was profitably used in image processing in [2], and the stability was studied through a so-called *error-control* strategy. This strategy consists in considering that the perturbation of the multiscale representation only arises from quantization error at the encoding of the representation. Controlling the level of quantization makes the representation stable. However, perturbation in the multiscale representation may come from noise that have nothing to do with quantization therefore more general notion of stability is used for instance in [13] and [8]. This is the latter point of view that we will adopt in our study of stability of multiscale representations based on *Lipschitz-Linear* prediction operators.

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 [20]. 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 $v^j = Sv^{j-1} + e^j$, 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 (where $r_k^d = \#\{\alpha \in \mathbb{Z}^d, |\alpha| = k\}$ and where $\#X$ stands for the cardinal of X), for $k \leq N+1$ is given by*

$$\begin{aligned} \rho_p(S^{(k)}) &:= \inf_{j>0} \|(S^{(k)})^j\|^{1/j}_{(\ell^p(\mathbb{Z}^d))^{r_k^d} \to (\ell^p(\mathbb{Z}^d))^{r_k^d}} \qquad (24) \\ &= \inf_{j>0}\{\rho, \exists j \quad \|\Delta^k S^j v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \leq \rho^j \|\Delta^k v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}, \forall v \in \ell^p(\mathbb{Z}^d)\}. \end{aligned}$$

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 (for more details see (5)). We first need to establish some extensions to the non-separable case of results obtained in [19]:

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

$$\|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \;\lesssim\; m^{-j/p}\left(\|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} + \|e^{j+1}\|_{\ell^p(\mathbb{Z}^d)}\right). \qquad (25)$$

Moreover, if $\rho_p(S^{(k)}) < m^{1/p}$, then for any ρ such that $\rho_p(S^{(k)}) < \rho < m^{1/p}$, there exists an n such that:

$$m^{-j/p}\|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \;\lesssim\; \delta^j \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{s-1} \delta^{nr} \sum_{l=j-(r+1)n+1}^{l=j-rn} m^{-l/p}\|e^l\|_{\ell^p(\mathbb{Z}^d)} \qquad (26)$$

where $\delta = \rho m^{-1/p}$ and $s = \lfloor j/n \rfloor$.

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:

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

The proof of (25) is thus complete. Note that we have used

$$\|\Phi_i(\Delta^{N+1}v^j_{\cdot+p_1}, \cdots, \Delta^{N+1}v^j_{\cdot+p_q})\|_{\ell^p(\mathbb{Z}^d)} = \|\bar{\Phi}_i(\Delta^{k}v^j_{\cdot+\bar{p}_1}, \cdots, \Delta^{k}v^j_{\cdot+\bar{p}_{\bar{q}}})\|_{\ell^p(\mathbb{Z}^d)}, \tag{27}$$

where $\bar{\Phi}_i$ is a Lipschitz function.Indeed, higher order finite differences can be expressed as linear combinations of lower order finite differences. To prove (26), we note that for any $\rho_p(S^{(k)}) < \rho < m^{1/p}$, there exists an n such that for all v:

$$\|(S^{(k)})^n v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \leq \rho^n \|v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}. \tag{28}$$

Using the boundedness of the operator $S^{(k)}$, we obtain:

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

Then for any j, define $s := \lfloor j/n \rfloor$, after s iterations of the above inequality, we get:

$$\|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \leq \rho^{ns} \|\Delta^k v^{j-ns}\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} + D \sum_{r=0}^{s-1} \rho^{nr} \sum_{l=rn}^{(r+1)n-1} \|e^{j-l}\|_{\ell^p(\mathbb{Z}^d)}$$

Then putting as in [13], $\delta = \rho m^{-1/p}$, and $A_j = m^{-j/p} \|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}$, we get:

$$A_j \leq \delta^{ns} A_{j-ns} + D \sum_{r=0}^{s-1} \delta^{nr} \sum_{l=nr}^{l=(r+1)n-1} m^{-(j-l)/p} \|e^{j-l}\|_{\ell^p(\mathbb{Z}^d)}$$

Then, we may write, due the boundedness of $S^{(k)}$, for $j' < n$:

$$A_{j'} \lesssim \|v_0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=0}^{j'} m^{-l/p} \|e^l\|_{\ell^p(\mathbb{Z}^d)}$$

which finally leads to:

$$m^{-j/p} \|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \lesssim \delta^j \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{s-1} \delta^{nr} \sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p} \|e^l\|_{\ell^p(\mathbb{Z}^d)}$$

□

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

$$\|v\|_{L^p(\mathbb{R}^d)} \leq \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p} \|e^j\|_{\ell^p(\mathbb{Z}^d)} \tag{29}$$

Proof. From estimates (25) and (26) one has, in particular

$$\|v_{j+1}-v_j\|_{L^p(\mathbb{R}^d)} \lesssim \delta^j\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{s}\delta^{nr}\sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p}\|e^l\|_{\ell^p(\mathbb{Z}^d)} \quad (30)$$

Considering that $\rho_p(S^{(k)}) < m^{1/p}$ and then $\rho_p(S^{(k)}) < \rho < m^{1/p}$, and then using (30), we get:

$$\begin{aligned}
\|v\|_{L^p(\mathbb{R}^d)} &\le \|v_0\|_{L^p(\mathbb{R}^d)} + \sum_{j\ge 0}\|v_{j+1}-v_j\|_{L^p(\mathbb{R}^d)} \\
&\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j\ge 0}\delta^j\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{s}\delta^{nr}\sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p}\|e^l\|_{\ell^p(\mathbb{Z}^d)} \\
&\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{s=0}^{\infty}\sum_{q=0}^{n-1}\sum_{r'=0}^{s}\delta^{n(s-r')}\sum_{l=r'n-n+q+1}^{l=r'n+q} m^{-l/p}\|e^l\|_{\ell^p(\mathbb{Z}^d)} \\
&\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r'=0}^{\infty}\sum_{s>r'}\delta^{n(s-r')}\sum_{q=0}^{n-1}\sum_{l=r'n-n+q+1}^{l=r'n+q} m^{-l/p}\|e^l\|_{\ell^p(\mathbb{Z}^d)} \\
&\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l}^{\infty} m^{-l/p}\|e^l\|_{\ell^p(\mathbb{Z}^d)}
\end{aligned}$$

The last equality being obtained remarking that $\sum_{s>r'}\delta^{n(s-r')} = \frac{1}{1-\delta^n}$. □

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

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 $s \ge N$ and some $k \le N+1$, and also that $(v^0, e^1, e^2, \ldots)$ satisfies*

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

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

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

The proof of (31) is similar to that of Theorem 5.3 of [20], 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 (because of Lemma 7.1).

8. Stability in L^p and Besov spaces

In applications, the multiscale 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 multiscale 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 there exist an $n \in \mathbb{N}$ and a $\rho < m^{1/p}$ such that:*

$$\|(S^{(k)})^n v - (S^{(k)})^n w\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \leq \rho^n \|v - w\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \quad \forall v, w \in (\ell^p(\mathbb{Z}^d))^{r_k^d},$$

for some $k \leq N+1$. Assume also that v_j and $\tilde{v}_j$ converge to v and $\tilde{v}$ in $L^p(\mathbb{R}^d)$ respectively. Then, we have:

$$\|v - \tilde{v}\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{L^p(\mathbb{R}^d)} + \sum_{l=1}^{j} m^{-l/p} \|e^l - \tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)} \tag{32}$$

Proof. We note that for all v and using the boundedness of the operator $S^{(k)}$, we obtain:

$$\begin{aligned}
\|\Delta^k(v^n-\tilde{v}^n)\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} &\leq \|S^{(k)}\Delta^k v^{n-1}-S^{(k)}\Delta^k\tilde{v}^{n-1}\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}+\|\Delta^k(e^n-\tilde{e}^n)\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}\\
&\leq \|(S^{(k)})^n\Delta^k v^0-(S^{(k)})^n\Delta^k v^0\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}+D\sum_{l=1}^{n}\|e^l-\tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}\\
&\leq \rho^n\|\Delta^k v^0-\Delta^k\tilde{v}^0\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}+D\sum_{l=1}^{n}\|e^l-\tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}
\end{aligned}$$

Then for any j, define $s:=\lfloor j/n\rfloor$, after s iterations of the above inequality, we get:

$$\|\Delta^k(v^j-\tilde{v}^j)\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}\leq\rho^{ns}\|\Delta^k(v^{j-ns}-v^{j-ns})\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}+D\sum_{r=0}^{s-1}\rho^{nr}\sum_{l=rn}^{(r+1)n-1}\|e^{j-l}-\tilde{e}^{j-l}\|_{\ell^p(\mathbb{Z}^d)}$$

Then by using the same reasoning as in the proof of (26), we get:

$$\begin{aligned}
m^{-j/p}\|\Delta^k(v^j-\tilde{v}^j)\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} &\lesssim \delta^j\|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}\\
&+\sum_{r=0}^{s}\delta^{nr}\sum_{l=j-(r+1)n+1}^{l=j-rn}m^{-l/p}\|e^l-\tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}
\end{aligned}\tag{33}$$

Now, note that:

$$\begin{aligned}
\|v\ -\tilde{v}\|_{L^p(\mathbb{R}^d)} &\leq \|v_0-\tilde{v}_0\|_{L^p(\mathbb{R}^d)}+\sum_{j>0}\|v_j-\tilde{v}_j-v_{j-1}+\tilde{v}_{j-1}\|_{L^p(\mathbb{R}^d)}\\
&\leq \|v_0-\tilde{v}_0\|_{L^p(\mathbb{R}^d)}+\sum_{j>0}\|Sv^{j-1}-S\tilde{v}_j+e^j-\tilde{e}^j-S_lv_{j-1}+S_l\tilde{v}_{j-1}\|_{L^p(\mathbb{R}^d)}\\
&\leq \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\\
&\quad\sum_{j>0,i\in\ \mathrm{coset}(M)}\|\bar{\Phi}_i(\Delta^k v^{j-1}_{\cdot+\bar{p}_1},\cdots,\Delta^k v^{j-1}_{\cdot+\bar{p}_{\bar{q}}})-\bar{\Phi}_i(\Delta^k\tilde{v}^{j-1}_{\cdot+\bar{p}_1},\cdots,\Delta^k\tilde{v}^{j-1}_{\cdot+\bar{p}_{\bar{q}}})+e^j-\tilde{e}^j\|_{L^p(\mathbb{R}^d)}\\
&\lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\sum_{j>0}m^{-j/p}\left(\|\Delta^k(v^{j-1}-\tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}+\|e^j-\tilde{e}^j\|_{\ell^p(\mathbb{Z}^d)}\right)\\
&\lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\sum_{j>0}m^{-j/p}\|e^j-\tilde{e}^j\|_{\ell^p(\mathbb{Z}^d)},
\end{aligned}$$

the last inequality being obtained using (33) and then making the same computation as in Theorem 7.2. □

With our formalism, the convergence and the stability of the nonlinear multiscale decomposition is based on the study of $S^{(k)}$ for some k. On the contrary, in [13], the study is carried out in L^∞ and the stability and the convergence are proved through the study of two different spectral radii. More precisely, the convergence of the multiscale representation is based on the study of the joint spectral radius of $S^{(k)}$ while the stability is based on the study of the joint spectral radius of the differential of $S^{(k)}$ (noted $DS^{(k)}$). Such a differential may sometimes be hard to compute. The formalism we propose enables simpler proofs for the stability. However, we are aware that the more complex mathematical framework developed in [13] aims at dealing with a wider class of prediction operators (for instance, the median interpolating scheme studied in [13] is not *Lipschitz-Linear*).

8.2. Stability in Besov spaces

In view of the inverse inequality (31), to show the stability, it seems natural to seek an inequality of type:

$$\|v-\tilde{v}\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\|(m^{(s/d-1/p)j}\|e^j_\cdot-\tilde{e}^j_\cdot\|_{\ell^p(\mathbb{Z}^d)})_{j>0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (34)$$

We now state without a 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 there exist an n in $\mathbb{N}$ and a $\rho \leq m^{1/p-s/d}$ for some $s > N$ such that:*

$$\|(S^{(k)})^n v - (S^{(k)})^n w\|_{(\ell^p(\mathbb{Z}^d))^{r^d_k}} \leq \rho^n \|v-w\|_{(\ell^p(\mathbb{Z}^d))^{r^d_k}} \quad \forall v, w \in (\ell^p(\mathbb{Z}^d))^{r^d_k},$$

for some $k \leq N+1$. Also assume that v_j and $\tilde{v}_j$ converge to v and $\tilde{v}$ in $B^s_{p,q}(\mathbb{R}^d)$ respectively. Then, we have:

$$\|v-\tilde{v}\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\|(m^{j(s/d-1/p)}\|(e^j_k-\tilde{e}^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j>0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (35)$$

The proof is the same as that of Theorem 6.2 of [20], except that we do not require the exact polynomial reproduction property.

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 fixed set, Φ_i are Lipschitz functions and if there exists an operator $S_l^{\mathcal{A}_I}$ satisfying:

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

From the definition of S, there exists 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. multiscale 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 to compute the joint spectral radius. Note also that the compatibility notion is not linked to polynomial reproduction for prediction operators, which makes it a new tool for analysis. From a 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 multiscale representation.

10. Applications

10.1. Convergence and Stability of One-Dimensional Multiscale Representation: the PPH 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 and stability

theorems (7.1 and 8.1 respectively). The novelty of the proposed approach is two-fold. First, it enables to characterize the stability in L^p not only in L^∞ as in [13] (Theorem 2.3) or in [4] (Proposition 1, for the PPH scheme). Second, the convergence and the stability of the multiscale representation is based on the study of $S^{(k)}$ for some $k \leq N+1$), while in [13] the convergence in L^∞ is related to the study of $\rho_\infty(S^{(k)})$ for some k and the stability is related to $\rho_\infty(DS^{(k)})$ where D stands for the Frêchet differential. This latter joint spectral radius is harder to study than $\rho_\infty(S^{(k)})$ and requires that $S^{(k)}$ is indeed differentiable. However, we must confess that the class of prediction operators studied in [13] is wider therefore the proofs for the stability are different.

Now, let us give an illustration of how Theorems 7.1 and 8.1 apply to the PPH *Lipschitz-Linear* prediction operator (we will then see how the proof of convergence extends to the slightly modified power-P scheme introduced in (17)). Since the PPH prediction operator is *Lipschitz-Linear* of order 2, the convergence in L^p occurs when $\rho_p(S^{(k)}) < 2^{1/p}$ for $k = 1$ or $k = 2$.

Here, we study the convergence of the PPH scheme by finding an upper bound for $\rho_p(S^{(2)})$, whose expression is particularly simple since:

$$\begin{aligned} (S^{(2)}w)_{2i} &= \frac{1}{4}H(w_{i-1}, w_i) \\ (S^{(2)}w)_{2i+1} &= \frac{w_i}{2} - \frac{1}{8}\left(H(w_{i-1}, w_i) + H(w_i, w_{i+1})\right). \end{aligned} \tag{36}$$

Remarking that $|H(x,y)| \leq |\max(x,y)|$, we immediately get: $\rho_\infty(S^{(2)}) < \frac{3}{4}$, which leads the convergence of the multiscale representation in L^∞ according to the Theorem 7.1. Now as far as the stability is concerned, is was proved in [4], Proposition 1 that:

$$\|(S^{(2)})^2 w - (S^{(2)})^2 v\|_\infty \leq \frac{3}{4}\|v - w\|_\infty \quad \forall v, w \in l^\infty(\mathbb{Z}^2),$$

which proves the stability of the scheme in L^∞ using the Theorem 8.1 (a different proof of that result is given by the Theorem 1 of [4]).

Based on the simple expression of $S^{(2)}$, we propose a new proof for the convergence of the PPH-scheme in L^p. Indeed, we may write assuming that

$p \geq 1$:

$$\begin{aligned}
|(S^{(2)}w)_{2i}|^p &\leq \frac{1}{4^p}\max(|w_{i-1}|,|w_i|)^p \\
|(S^{(2)}w)_{2i+1}|^p &\leq \left(\frac{1}{2}|w_i| + \frac{1}{8}\max(|w_{i-1}|,|w_i|) + \frac{1}{8}\max(|w_i|,|w_{i+1}|)\right)^p \\
&\leq \left(\frac{1}{2}|w_i| + \frac{1}{4}(\frac{1}{2}\max(|w_{i-1}|,|w_i|) + \frac{1}{2}\max(|w_i|,|w_{i+1}|))\right)^p \\
&\leq \frac{1}{2}|w_i|^p + \frac{1}{4}\frac{1}{2^p}\max(|w_{i-1}|,|w_i|)^p + \frac{1}{4}\frac{1}{2^p}\max(|w_i|,|w_{i+1}|)^p.
\end{aligned} \tag{37}$$

The last inequality being obtained because we have a convex combination. Now, to obtain an upper bound for $\rho_p(S^{(2)})$, we note:

$$\begin{aligned}
\|S^{(2)}w\|^p_{\ell^p(\mathbb{Z}^d)} &\leq \sum_{i\in\mathbb{Z}} \frac{1}{4^p}\max(|w_{i-1}|,|w_i|)^p + \frac{1}{2}|w_i|^p + \\
&\quad \frac{1}{4}\frac{1}{2^p}\max(|w_{i-1}|,|w_i|)^p + \frac{1}{4}\frac{1}{2^p}\max(|w_i|,|w_{i+1}|)^p.
\end{aligned}$$

The largest coefficient in front of $|w_i|^p$ in the above sum is obtained when $|w_i|$ is larger than $|w_{i-1}|$ and $|w_{i+1}|$. In such a case, one can easily check that the coefficient in front of $|w_i|^p$ is then $\frac{2}{4^p} + \frac{1}{2} + \frac{1}{2^p}$, which means that $\|S^{(2)}\|_{\ell^p(\mathbb{Z}^d)\to\ell^p(\mathbb{Z}^d)} \leq (\frac{1}{2} + \frac{1}{2^p} + \frac{2}{4^p})^{\frac{1}{p}}$. This in turn implies that the multiscale representation is convergent in L^p provided that $\frac{1}{2} + \frac{1}{2^p} + \frac{2}{4^p} < 2$, which is true for any $p \geq 1$.

As far as the stability of the scheme in L^p is concerned we may write (assuming $p \geq 1$),

$$\begin{aligned}
|(S^{(2)}w)_{2i} - (S^{(2)}v)_{2i}|^p &\leq \frac{1}{2^p}\max(|w_{i-1}-v_{i-1}|,|w_i-v_i|)^p \\
|(S^{(2)}w)_{2i+1} - (S^{(2)}v)_{2i+1}|^p &\leq \frac{1}{2}|w_i-v_i|^p + \frac{1}{2}\max(|w_{i-1}-v_{i-1}|,|w_{i+1}-v_{i+1}|)^p,
\end{aligned}$$

the last inequality being a consequence of Lemma 2 of [4]. Now, as in the study of the convergence, we write:

$$\begin{aligned}
\|S^{(2)}w - S^{(2)}v\|^p_{\ell^p(\mathbb{Z}^d)} &\leq \sum_{i\in\mathbb{Z}} \frac{1}{2^p}\max(|w_{i-1}-v_{i-1}|,|w_i-v_i|)^p + \frac{1}{2}|w_i-v_i|^p + \\
&\quad \frac{1}{2}\max(|w_{i-1}-v_{i-1}|,|w_{i+1}-v_{i+1}|)^p
\end{aligned}$$

The largest coefficient in front of $|w_i - v_i|^p$ in the above sum is obtained when $|w_i - v_i|$ is larger than $|w_{i+r} - v_{i+r}|$ $r = -2, -1, 2$. In such a case, one can check that the coefficient in front of $|w_i - v_i|^p$ in the right term of the above inequality is $\frac{3}{2} + \frac{1}{2^p}$, so that we may deduce:

$$\|S^{(2)}w - S^{(2)}v\|_{\ell^p(\mathbb{Z}^d)} \leq (\frac{3}{2} + \frac{1}{2^p})^{1/p}\|w - v\|_{\ell^p(\mathbb{Z}^d)}$$

which proves that the scheme is stable whenever $p > 1$ ($i.e. \frac{3}{2} + \frac{1}{2^p} < 2$, since $m = 2$ in that case), using Theorem 8.1.

The above proof does not extend to the power-P scheme which is proved to be unstable for $q > 4$. However, if one replaces H_q by $\tilde{H}_q$ defined in (17) and assumes that (w_i, w_{i+1}) belongs to $\mathbb{R}^2 \setminus V_\varepsilon$ (see (18)), then $\tilde{H}_q$ is Lipschitz on that set. Now, remarking that $|\tilde{H}_q(x, y)| \leq \max(|x|, |y|)$ and making the same reasoning as for the PPH-scheme, we obtain that the modified power-P scheme leads to a convergent multiscale representation in L^p for any $p \geq 1$.

10.2. Convergence of One-dimensional Multiscale Representations: the WENO Case

We consider here the model defined in section 5.4. In this case, one can show the following Lemma:

Lemma 10.1. *One has*

$$\sup_{u,w \in l_\infty} \|S^{(1)}(u)S^{(1)}(w)\|_{l_\infty} < 1$$

and therefore $\rho_\infty(S^{(1)}) < 1$

The proof is identical to that of Lemma 4 of [8], therefore we do not expand on it here. The multiscale representation based on the proposed WENO model is therefore convergent in L^∞.

10.3. Convergence and Stability of Bidimensional PPH multiscale Representations

We study the convergence and the stability of bidimensional PPH multiscale representations where the prediction operator is given by:

$$\begin{aligned} \hat{v}^j_{Mk+e_1} &= \frac{v^{j-1}_k + v^{j-1}_{k+Me_1}}{2} - \frac{\omega}{8} H(\Delta^2_{Me_1} v^{j-1}_k, \Delta^2_{Me_1} v^{j-1}_{k-Me_1}) \\ \hat{v}^j_{Mk} &= v^{j-1}_k. \end{aligned} \tag{38}$$

for some $0 < \omega < 1$. To consider $\omega < 1$ instead of $\omega = 1$ as in (20) will appear clearer a bit later. We already noticed that the prediction operator is *Lipschitz-Linear*, we now remark 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 of the multiscale representation, we 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:

$$\begin{aligned}
\Delta^2_{e_1}\hat{v}^j_{Mk} &= \frac{\omega}{4}H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) \\
\Delta^2_{e_1}\hat{v}^j_{Mk+e_1} &= \frac{1}{2}\Delta^2_{Me_1}v^{j-1}_k \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_{k+Me_1}, \Delta^2_{Me_1}v^{j-1}_k) \\
\Delta^2_{Me_1}\hat{v}^j_{Mk} &= \Delta^2_{e_1}v^{j-1}_k \\
\Delta^2_{Me_1}\hat{v}^j_{Mk+e_1} &= \frac{1}{2}(\Delta^2_{e_1}v^{j-1}_k + \Delta^2_{e_1}v^{j-1}_{k+Me_1}) + \frac{\omega}{4}H(\Delta^2_{Me_1}v^{j-1}_{k+e_1}, \Delta^2_{Me_1}v^{j-1}_{k+e_1-Me_1}) \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_{k+2e_1}, \Delta^2_{Me_1}v^{j-1}_{k+2e_1-Me_1}). \qquad (39)
\end{aligned}$$

We now study more in detail $\Delta^2_{e_1}v^j_{Mk+e_1}$, the following cases can appear:

1. $\Delta^2_{Me_1}v^{j-1}_k\Delta^2_{Me_1}v^{j-1}_{k-Me_1} > 0$ and $\Delta^2_{Me_1}v^{j-1}_{k+Me_1}\Delta^2_{Me_1}v^{j-1}_k > 0$ we have

   $$|\Delta^2_{e_1}\hat{v}^j_{Mk+e_1}| \leq$$
   $$\max(\frac{1}{2}|\Delta^2_{Me_1}v^{j-1}_k|, \frac{\omega}{8}|H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) + H(\Delta^2_{Me_1}v^{j-1}_{k+Me_1}, \Delta^2_{Me_1}v^{j-1}_k)|)$$

2. $\Delta^2_{Me_1}v^{j-1}_k\Delta^2_{Me_1}v^{j-1}_{k-Me_1} \leq 0$ and $\Delta^2_{Me_1}v^{j-1}_{k+Me_1}\Delta^2_{Me_1}v^{j-1}_k \leq 0$ we have

   $$|\Delta^2_{e_1}\hat{v}^j_{Mk+e_1}| = \frac{1}{2}\Delta^2_{Me_1}v^{j-1}_k$$

3. $\Delta^2_{Me_1}v^{j-1}_k\Delta^2_{Me_1}v^{j-1}_{k-Me_1} \leq 0$ and $\Delta^2_{Me_1}v^{j-1}_{k+Me_1}\Delta^2_{Me_1}v^{j-1}_k > 0$ we have

   $$|\Delta^2_{e_1}\hat{v}^j_{Mk+e_1}| \leq \max(\frac{1}{2}|\Delta^2_{Me_1}v^{j-1}_k|, \frac{\omega}{8}|H(\Delta^2_{Me_1}v^{j-1}_{k+Me_1}, \Delta^2_{Me_1}v^{j-1}_k)|)$$

   A similar equation is obtained assuming
   $\Delta^2_{Me_1}v^{j-1}_k\Delta^2_{Me_1}v^{j-1}_{k-Me_1} > 0$ and $\Delta^2_{Me_1}v^{j-1}_{k+Me_1}\Delta^2_{Me_1}v^{j-1}_k \leq 0$.

Now, remarking as previously that $|H(x,y)| \leq \max(|x|,|y|)$, we immediately obtain that

$$\begin{aligned}
\|\Delta^2_{e_1}\hat{v}^j\|_\infty &\leq \frac{1}{2}\|\Delta^2_{Me_1}v^{j-1}\|_\infty \\
\|\Delta^2_{Me_1}\hat{v}^j\|_\infty &\leq \|\Delta^2_{e_1}v^{j-1}\|_\infty + \frac{\omega}{2}\|\Delta^2_{Me_1}v^{j-1}\|_\infty.
\end{aligned}$$

From these inequality we immediately deduce that $\rho_\infty(S^{\mathcal{A}_I}) \leq \sqrt{\frac{1+\omega}{2}} < 1$, which proves that the bidimensional PPH defined by (38) is convergent in L^∞.

For the L^p convergence, we do not need the restriction on ω and we consider the model defined by (20), therefore we study:

$$\begin{aligned}
\|\Delta^{\mathcal{A}_I}\hat{v}^j\|^p_{(\ell^p(\mathbb{Z}^d))^2} &= \|S^{\mathcal{A}_I}\Delta^{\mathcal{A}_I}v^{j-1}\|^p_{(\ell^p(\mathbb{Z}^d))^2} \\
&= \sum_{k\in\mathbb{Z}^2} |\Delta^2_{Me_1}\hat{v}^j_{Mk+e_1}|^p + |\Delta^2_{Me_1}\hat{v}^j_{Mk}|^p + |\Delta^2_{e_1}\hat{v}^j_{Mk}|^p + |\Delta^2_{e_1}\hat{v}^j_{Mk+e_1}|^p.
\end{aligned}$$

As in the one-dimensional study, and assuming $p \geq 1$, we have the following majorations (using the property of convex functions):

$$\begin{aligned}
|\Delta^2_{Me_1}\hat{v}^j_{Mk+e_1}|^p &\leq \frac{1}{4}\frac{1}{2^p}|\Delta^2_{e_1}v^{j-1}_k|^p + \frac{1}{4}\frac{1}{2^p}|\Delta^2_{e_1}v^{j-1}_{k+Me_1}|^p \\
&+ \frac{1}{4}\max(|\Delta^2_{Me_1}v^{j-1}_{k+e_1}|, |\Delta^2_{Me_1}v_{k+e_1-Me_1}|)^p \\
&+ \frac{1}{8}\max(|\Delta^2_{Me_1}v^{j-1}_{k}|, |\Delta^2_{Me_1}v_{k-Me_1}|)^p \\
&+ \frac{1}{8}\max(|\Delta^2_{Me_1}v^{j-1}_{k+2e_1}|, |\Delta^2_{Me_1}v_{k+2e_1-Me_1}|)^p \\
|\Delta^2_{Me_1}\hat{v}^j_{Mk}|^p &\leq |\Delta^2_{e_1}v^{j-1}_k|^p \\
|\Delta^2_{e_1}\hat{v}^j_{Mk}|^p &\leq \frac{1}{4^p}\max(|\Delta^2_{Me_1}v^{j-1}_{k}|, |\Delta^2_{Me_1}v^{j-1}_{k-Me_1}|)^p \\
|\Delta^2_{e_1}\hat{v}^j_{Mk+e_1}|^p &\leq \frac{1}{2}|\Delta^2_{Me_1}v^{j-1}_k|^p \\
&+ \frac{1}{2^p}\frac{1}{4}\max(|\Delta^2_{Me_1}v^{j-1}_{k}|, |\Delta^2_{Me_1}v^{j-1}_{k-Me_1}|)^p \\
&+ \frac{1}{2^p}\frac{1}{4}\max(|\Delta^2_{Me_1}v^{j-1}_{k+Me_1}|, |\Delta^2_{Me_1}v^{j-1}_{k}|)^p
\end{aligned}$$

Now, as in the one-dimensional case, we consider the largest possible coefficients in front of each differences, to obtain:

$$\begin{aligned}\|S^{\mathcal{A}_I}\Delta^{\mathcal{A}_I}v^{j-1}\|^p_{(\ell^p(\mathbb{Z}^d))^2} &\leq \sum_{k\in\mathbb{Z}^2}(1+\frac{1}{2\times 2^p})|\Delta^2_{e_1}v^j_k|^p+(1+\frac{2}{4^p}+\frac{1}{2^p})|\Delta^2_{Me_1}v^j_k|^p\\ &\leq \max(1+\frac{1}{2\times 2^p},1+\frac{1}{2^p}+\frac{2}{4^p})\|\Delta^{\mathcal{A}_I}v^{j-1}\|^p_{(\ell^p(\mathbb{Z}^d))^2}\end{aligned}$$

Recalling that $m=2$, we get the L^p convergence and stability as soon as $\max(1+\frac{1}{2\times 2^p},1+\frac{1}{2^p}+\frac{2}{4^p})<2$, which is always true for $p>1$.

To finish with, let us study the stability of the PPH-scheme defined by (38) in L^p. We may indeed write:

$$\begin{aligned}|\Delta^2_{e_1}(\hat{v}^j_{Mk}-\hat{\tilde{v}}^j_{Mk})|^p &\leq \left(\frac{\omega}{2}\right)^p\max(|\Delta^2_{Me_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|,|\Delta^2_{Me_1}(v^{j-1}_{k-Me_1}-\tilde{v}^{j-1}_{k-Me_1})|)^p\\ |\Delta^2_{e_1}(\hat{v}^j_{Mk+e_1}-\hat{\tilde{v}}^j_{Mk+e_1})|^p &\leq \frac{1}{2}|\Delta^2_{Me_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|^p\\ &+ \frac{\omega^p}{4}\big(\max(|\Delta^2_{Me_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|,|\Delta^2_{Me_1}(v^{j-1}_{k-Me_1}-\tilde{v}^{j-1}_{k-Me_1}|)^p\\ &+ \max(|\Delta^2_{Me_1}(v^{j-1}_{k+Me_1}-\tilde{v}^{j-1}_{k+Me_1})|,|\Delta^2_{Me_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|)^p\big)\\ |\Delta^2_{Me_1}(\hat{v}^j_{Mk}-\hat{\tilde{v}}^j_{Mk})|^p &\leq |\Delta^2_{e_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|^p\\ |\Delta^2_{Me_1}(v^j_{Mk+e_1}-\hat{\tilde{v}}^j_{Mk+e_1})|^p &\leq \frac{1}{4(2^p)}(|\Delta^2_{e_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|^p+|\Delta^2_{e_1}(v^{j-1}_{k+Me_1}-\tilde{v}^{j-1}_{k+Me_1})|^p)+\frac{(2\omega)^p}{8}\times\\ &\quad\big(2\max(|\Delta^2_{Me_1}(v^{j-1}_{k+e_1}-\tilde{v}^{j-1}_{k+e_1})|,|\Delta^2_{Me_1}(v^{j-1}_{k+e_1-Me_1}-\tilde{v}^{j-1}_{k+e_1-Me_1})|)^p\\ &+ \max(|\Delta^2_{Me_1}(v^{j-1}_k-\tilde{v}^{j-1}_k)|,|\Delta^2_{Me_1}(v^{j-1}_{k-Me_1}-\tilde{v}^{j-1}_{k-Me_1})|)^p\\ &+ \max(|\Delta^2_{Me_1}(v^{j-1}_{k+2e_1}-\tilde{v}^{j-1}_{k+2e_1})|,|\Delta^2_{Me_1}(v^{j-1}_{k+2e_1-Me_1}-\tilde{v}^{j-1}_{k+2e_1-Me_1})|)^p\end{aligned}$$

From this we deduce that:

$$\begin{aligned}\|S^{\mathcal{A}_I}\Delta^{\mathcal{A}_I}(\hat{v}^j-\hat{\tilde{v}}^j)\|^p_{(\ell^p(\mathbb{Z}^d))^2} &\leq \sum_{k\in\mathbb{Z}^2}(1+\frac{1}{2\times 2^p})|\Delta^2_{e_1}v^j_k-\Delta^2_{e_1}v^j_k|^p+\\ &\quad\left(\omega^p+(2\omega)^p+2\left(\frac{\omega}{2}\right)^p\right)|\Delta^2_{Me_1}v^j_k-\Delta^2_{Me_1}\tilde{v}^j_k|^p\\ &\leq \max\left(1+\frac{1}{2(2^p)},\omega^p(1+2^p+2\frac{1}{2^p})\right)\|\Delta^{\mathcal{A}_I}(v^{j-1}-\tilde{v}^{j-1})\|^p_{(\ell^p(\mathbb{Z}^d))^2}\end{aligned}$$

Since $\max\left(1+\frac{1}{2(2^p)},\omega^p(1+2^p+2\frac{1}{2^p})\right)<2$ for all $p\geq 1$ as soon as $\omega<1/2$, we deduce that the scheme defined by (38) is stable in that case.

11. Conclusion

In this paper, we have introduced a new formalism for nonlinear and non-separable multiscale representations. The introduced formalism includes classical nonlinear multiscale representations such as WENO 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. We called these particular kind of prediction operators *Lipschitz-Linear*. After having illustrated the proposed formalism on one and multi-dimensional cases, we stated the convergence and stability theorems in L^p and Besov spaces. The novelty of the result is that these theorems are based on the study of the same difference operator. 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 terms of perspectives, we are currently investigating how to apply the model of *Lipschitz-Linear* prediction operator to design new convergent and stable multiscale representations with application to image compression.

Appendix A

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 $y_0 = 0, y_1 = v_{k-N}^{j-1}, y_2 = y_1 + v_{k-N+1}^{j-1}, \cdots, y_{2N+1} = y_{2N} + v_{k+N}^{j-1}$ 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 $\tilde{y}_1 = 0, \tilde{y}_2 = v_{k-N+1}^{j-1}, \tilde{y}_3 = \tilde{y}_2 + v_{k-N+2}^{j-1}, \cdots, \tilde{y}_{2N+2} = \tilde{y}_{2N+1} + v_{k+N+1}^{j-1}$ 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:

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

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 (12) 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:

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

[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] 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).

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

[5] E. J. Candès, Ridgelets and the representation of mutilated Sobolev functions, SIAM J. Math. Anal, 33, 197-218, (1999)

[6] J. L. Starck, E. J. Candès and D. L. Donoho, The curvelet transform for image denoising, IEEE Transactions on Image Processing, 11, 670-684, (2000).

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

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

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

[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] S. Harizanov and P. Oswald, Stability of Nonlinear Subdivision and multiscale Transforms, Constr. Approx. 31, 359-393, (2010).

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

[17] Le Pennec E., Mallat, S., Sparse Geometric Image Representation with Bandelets, IEEE Trans. on Image Processing, 14, 423-438, (2005).

[18] 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).

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

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

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

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