Numerical Algorithms manuscript No.
(will be inserted by the editor)

Analysis of a Class of Nonlinear and Non-Separable Multiscale Representations

Basarab Mateï and Sylvain Meignen

Received: date / Accepted: date

Abstract In this paper, we introduce a particular class of nonlinear and non-separable multiscale representations which embeds most of these representations. After motivating the introduction of such a class on one-dimensional examples, we investigate the multi-dimensional and non-separable case where the scaling factor is given by a non-diagonal dilation matrix M. We also propose new convergence and stability results in L^p and Besov spaces for that class of nonlinear and non-separable multiscale representations. We end the paper with an application of the proposed study to the convergence and the stability of some nonlinear multiscale representations.

Keywords Nonlinear multiscale representation · Convergence · Stability.

Mathematics Subject Classification (2000) MSC 41A46 · MSC 41A60 · MSC 41A63

1 Introduction

Multiscale representations such as wavelet-type pyramid transforms for hierarchical data representation [1] and subdivision methods for computer-aided geometric design [13] have completely changed the domains of data and geometry processing. Linear multiscale representations of functions are now well understood in terms of approximation performance [8]. While in the univariate case the wavelet-type pyramid transforms provide optimal algorithms [8], in the multivariate case almost all algorithms fail in the treatment of nonlinear constraints that are inherent to the analyzed objects, i.e., 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

Basarab Mateï
LAGA Laboratory, Paris XIII University, France,
Tel:0033-1-49-40-35-71
FAX:0033-1-49-40-35-68
E-mail: matei@math.univ-paris13.fr

Sylvain Meignen
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

N-term approximation for cartoon images. Improving this rate through a better representation of images near edges has motivated the study of ridgelets [6], curvelets [7] and bandlets [20]. These are bases or frames allowing for anisotropic refinement close to edges. Nonlinear multiscale representations [9] are another possibility to perform anisotropic refinement by using a nonlinear prediction operator. In a nutshell, the main difference between ridgelets, curvelets or bandlets and nonlinear multiscale representations is that the former are based on a functional point of view while the latter adopts a discrete point of view. To obtain approximation properties for these edge-adapted bases/frames representations, one uses the same kind of techniques as those used in wavelet analysis. On the contrary, the analysis of nonlinear multiscale representations requires a different mathematical framework. The development of nonlinear prediction operators such as quasi-linear prediction operators [22], median-interpolating schemes [25], normal multiresolution for curves and surfaces [11], nonlinear four point schemes [19] or power-P schemes [5] has enabled to design different kinds of nonlinear multiscale representations. The applications of such nonlinear multiscale representations range from geometrical image representations [9][5], non-Gaussian noise removal [12][25], or surfaces and curves compression [11].

In what follows, we analyze a broad class of nonlinear prediction operators that embeds, for instance, quasi-linear prediction operators and some nonlinear four point schemes. This class consists in *bounded nonlinear* prediction operators (BNPO) that are the sum of a linear prediction operator and of a bounded perturbation term (in a sense made clearer later).

After having introduced some notation, the notion of nonlinear multiscale representation and the definition of BNPO (section 2 to section 4), we show that the WENO prediction operator, some nonlinear four point schemes and a modified version of the power-P scheme are particular cases of one-dimensional BNPO (section 5). Examples of non-separable multidimensional BNPO are then given in section 6. The potential interest for such prediction operators lies in the fact that examples exist in image processing where the use of representations built using non-dyadic grids significantly improves the compression performance [10][21][18]. Section 7 establishes some new convergence results, in L^p and Besov spaces, for nonlinear multiscale representations based on BNPO. The stability of the multiscale representations requires to consider a slightly stronger hypothesis on the prediction operator, i.e., the perturbation term has to be a bounded and Lipschitz function. We will call these prediction operators Lipschitz nonlinear prediction operators. Stability theorems are stated in section 8 still in L^p and Besov spaces. The novelty of the proposed approach lies in the fact that the convergence and stability results are valid for non interpolatory and non-separable multiscale representations while the results available so far for non-separable multiscale representations only involved interpolatory schemes [24].

A new aspect is then introduced in section 9, through the notion of bounded (or Lipschitz) nonlinear prediction operators compatible with a set of finite differences, which we call $(\mathscr{A}, I)$ compatible in the present paper. The idea is to remark that for the multiscale representations associated with that kind of prediction operators the convergence or the stability can be proved by studying only a restricted set of finite differences (the directions for the differences being defined by the set $\mathscr{A}$, while the orders of differentiation by the vector I). From a practical point of view, the extension of the convergence and stability results when only a restricted set of finite differences is involved, allows us to show the convergence of some non-separable bidimensional schemes for which theoretical results did not exist (section 10).

2 Notation

Before we start, we need to introduce some standard multi-index notation. For some $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}^d$, 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 $r_N^d = \binom{N+d-1}{N}$ 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 $\mathbb{Z}^d$. 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 $\Delta_a^{\alpha_i} v_k$, for any vector a in $\mathbb{Z}^d$ is defined recursively by:

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

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

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

3 Multiscale Representations

We assume that the data $(v_k^j)_{k\in\mathbb{Z}^d}$ are associated to the locations $\Gamma^j := \{M^{-j}k, k \in \mathbb{Z}^d\}$, $j \geq 0$, where M is a dilation matrix, i.e., a $d \times d$ invertible matrix defined on $\mathbb{Z}$ satisfying $\lim_{n\to+\infty} M^{-n} = 0$. 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*

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

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

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

The scaling equation (4) gives rise to the definition of a local and linear prediction operator S_l as follows:

$$S_l v_k = \sum_{l \in \mathbb{Z}^d} g_{k-Ml} v_l. \tag{5}$$

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 $(v^0,(e^j)_{j>0})$ is convergent in that space.

4 Bounded and Lipschitz Nonlinear Prediction Operators

In this paper, we study a particular type of nonlinear prediction operators which is the sum of S_l and of a perturbation term. For that sake, we will need the notion of polynomial reproduction for prediction operator adapted to the non-separable context:

Definition 1 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 S exactly reproduces polynomials.

Let us consider $Q := \mathbb{Z}^d/M\mathbb{Z}^d$, which is made of m equivalence classes called *cosets* associated with the matrix M. We define the set of representatives of the cosets $C(M)$ by $MU \bigcap \mathbb{Z}^d$ where $U = [0,1[^d$. With this in mind, we first introduce the definition *bounded nonlinear* prediction operators:

Definition 2 Assume that S_l reproduces polynomials of degree N, then a nonlinear prediction operator S is bounded of order $N+1$, if it can be written under the following form:

$$Sv_{Mk+i} = S_l v_{Mk+i} + \Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}), \quad \forall i \in C(M), \tag{6}$$

where $\{p_1,\cdots,p_q\}$ is a fixed set and where Φ_i is bounded in the following sense:

$$|\Phi_i(x_{p_1},\cdots,x_{p_q})| \lesssim \max_{i \in p_1,\cdots,p_q} \|x_i\| \tag{7}$$

where $\|.\|$ denotes any norm on $\mathbb{R}^{r_d^{N+1}}$.

Remark 1 In the definition above we ask for the perturbation term Φ_i to be bounded in the sense of (7). The boundedness property for Φ_i will be useful to prove the convergence of the multiscale representations based on prediction operators satisfying (8).

To prove the stability of such multiscale representations, we will need that the function Φ_i be bounded and Lipschitz, therefore we introduce the following definition of *Lipschitz nonlinear* prediction operators:

Definition 3 Assume that S_l reproduces polynomials of degree N, then a nonlinear prediction operator S is Lipschitz of order $N+1$, if it can be written under the following form:

$$Sv_{Mk+i} = S_l v_{Mk+i} + \Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}), \quad \forall i \in C(M), \tag{8}$$

where $\{p_1,\cdots,p_q\}$ is a fixed set and where Φ_i is a Lipschitz function satisfying $\Phi_i(0) = 0$.

Remark 2 Note that if Φ_i is a Lipschitz function satisfying $\Phi_i(0) = 0$, then it entails that Φ_i is bounded

5 One-Dimensional Lipschitz Nonlinear Prediction Operators

We first recall how to define prediction operators in the point-value or cell-average settings and then show how the WENO prediction operator [16] can be viewed as a Lipschitz nonlinear prediction operator. Another illustration is given by a nonlinear four-point scheme often called the PPH-scheme in the literature [4].

5.1 Preliminaries

We start by considering the one-dimensional case with $M = 2$. Given a set of embedded grids $\Gamma^j = \left\{2^{-j}k,\ k \in \mathbb{Z}\right\}$, 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, i.e., $v_k^j = v(2^{-j}k)$ and (ii) cell-average, where v_k^j is 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 representations.

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 and assuming the x_i are equi-spaced, we write the difference between the two polynomials as:

$$p_{N,1}(x) - p_N(x) = \Delta^{N+1} v_0 \frac{1}{N! h^N} \prod_{i=1}^{N} (x - x_i), \tag{9}$$

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 Prediction Operators in the Point-Value Setting

Here, we use identity (9) 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 $v_k^{j-1} = v(2^{-j+1}k),\ k \in \mathbb{Z}$. In that framework, since $v_{2k}^j = v_k^{j-1}$, 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 $\hat{v}_{2k+1}^j$ through the so-called *centered* prediction as follows:

$$\hat{v}_{2k+1}^j = p_{2N+1}(2^{-j}(2k+1)). \tag{10}$$

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 literature (see [14]). 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:

$$\hat{v}^j_{2k+1,1} := p_{3,1}(2^{-j}(2k+1)) = \frac{5}{16}v^{j-1}_k + \frac{15}{16}v^{j-1}_{k+1} - \frac{5}{16}v^{j-1}_{k+2} + \frac{1}{16}v^{j-1}_{k+3}. \quad (11)$$

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

$$\hat{v}^j_{2k+1,1} - \hat{v}^j_{2k+1} = \frac{1}{16}\Delta^4 v^{j-1}_{k-1}, \quad (12)$$

which corresponds to (9), 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 by 2^{-j+1} to the left. We can generalize the above formula to any N through the following proposition:

Proposition 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:*

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

Proof Let us put $x_0 = 2^{-j+1}(k-N),\cdots,x_{2N+1} = 2^{-j+1}(k+N+1)$. Then, using (9) 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 3 Note that we can define other polynomials $p_{2N+1,q}$ for $-N \le q \le N$, 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 this is 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 1.

5.3 Prediction Operators in the Cell-Average Setting

We now show how Proposition 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:

$$v^j_k = 2^j\int_{I_{j,k}} v(t)dt. \quad (13)$$

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

$$v_k^{j-1} = \frac{1}{2}(v_{2k}^j + v_{2k+1}^j). \tag{14}$$

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

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

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:

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

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

$$\hat{v}_{2k+1,1}^j - \hat{v}_{2k+1}^j = -\frac{1}{8}\Delta^3 v_{k-1}^{j-1} \text{ and } \hat{v}_{2k,1}^j - \hat{v}_{2k}^j = \frac{1}{8}\Delta^3 v_{k-1}^{j-1}. \tag{15}$$

Similarly, we can define a prediction using the set of intervals shifted to the left and obtain the same kind of result. The equality (15) can then be generalized to any N:

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

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

The proof is available in Appendix A. As in the point-value setting, we can define $p_{2N,q}$, for any q, by shifting the computation intervals and then predict using this polynomial to obtain $\hat{v}_{k,q}^j$.

5.4 WENO-Prediction Operator as a Lipschitz Prediction Operator

Given a type a multiscale representation (i.e. either point-value or cell- average), the WENO prediction operator is based on a convex combination of the potential prediction rules $\hat{v}^j_{k,q}$, that is we write: $\hat{v}^j_{k,w} := \sum_{q=-N,q\neq 0}^{N} \alpha^j_{k,q}\hat{v}^j_{k,q}$ with $\alpha^j_{k,q} \geq 0$ and $\sum_{q=-N,q\neq 0}^{N} \alpha^j_{k,q} = 1$ and where the index w stands for WENO. The weights depends on v^{j-1} and on the corresponding rule q.

As an illustration, let us consider the point-value setting when $N = 1$, for which we have:

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

If one defines S_l the prediction operator associated to the centered prediction and S the prediction operator associated to the WENO prediction, it is clear that the WENO prediction can be written in the form (8) where Φ_i is a bounded function (using the fact that we consider a convex combination). It follows that the just defined WENO prediction operator is a bounded nonlinear prediction operator.

Note that in this case the perturbation term is not a Lipschitz function. To obtain a Lipschitz perturbation term, we can consider that $\alpha^j_{1,k}$ is a given function $\alpha(\Delta^4 v^{j-1}_{k-2}, \Delta^4 v^{j-1}_{k-1})$. Since $\alpha^j_{-1,k}$ equals $1 - \alpha^j_{1,k}$, we determine α such that $\alpha(x,y)(x-y)$ is a Lipschitz function which is true when 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 will be shown in the Applications section.

5.5 PPH-scheme as a Lipschitz Nonlinear Prediction Operator

We now show that the PPH-scheme defined by:

$$\begin{cases} \hat{v}^j_{2k+1} = \frac{v^{j-1}_{k+1}+v^{j-1}_k}{2} - \frac{1}{8}H(\Delta^2 v^{j-1}_{k-1}, \Delta^2 v^{j-1}_k) \\ \hat{v}^j_{2k} = v^{j-1}_k \end{cases} \tag{17}$$

where $H(x,y) := 2\left(\frac{xy}{x+y}\right)\chi_{\{xy>0\}}(x,y)$, and where χ_X is the characteristic function of X, is an example of Lipschitz nonlinear prediction operator. Note that since H satisfies $|H(x,y) - H(x',y')| \leq 2\max\{|x-x'|, |y-y'|\}$, the boundedness in the sense of Definition 2 follows. 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 bounded nonlinear prediction operator of order 2. Moreover $H(x,y)$ is Lipschitz with respect to (x,y), which implies that this scheme is also a Lipschitz nonlinear prediction operator.

A related example is the power-P scheme [26]. This scheme, is a generalization of the PPH-scheme replacing H by

$$H_q(x,y) = \left(\frac{x+y}{2}\left(1 - \left|\frac{x-y}{x+y}\right|^q\right)\right)\chi_{\{xy>0\}}(x,y). \tag{18}$$

Since $H_q(x,y)$ is bounded with respect to (x,y) as in Definition 2, the power-P scheme defines a bounded nonlinear prediction operator of order 2.

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 [15]. Nevertheless, a more careful look shows that the power-P scheme is very close to a Lipschitz nonlinear prediction operator. Indeed, consider the following definition for $x \neq y$:

$$\tilde{H}_q(x,y) = \left(\frac{x+y}{2}\left(1-\left|\frac{x-y}{x+y}\right|^q\right)\right) \times (\rho_\varepsilon * \chi_{\{xy>0\}})(x,y), \tag{19}$$

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

$$V_\varepsilon = \{|x| \leq \varepsilon\} \bigcup \{|y| \leq \varepsilon\}. \tag{20}$$

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 [15] for the computation). Then $\tilde{H}_q(x,y)$ is Lipschitz when $x \neq y$. By taking into account the definition set for $\tilde{H}_q$, 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 two models differ on a small band depending on ε.

6 Multi-Dimensional Lipschitz Nonlinear Prediction Operators on Non-Dyadic Grids

To illustrate the notion of Lipschitz nonlinear 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 [10] and [21]). 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.,

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{21}$$

and where the prediction is defined by:

$$\begin{aligned} \hat{v}^j_{Mk+e_1} &= \frac{v_k^{j-1} + v_{k+Me_1}^{j-1}}{2} - \frac{1}{8}H(\Delta^2_{Me_1}v_k^{j-1}, \Delta^2_{Me_1}v_{k-Me_1}^{j-1}) \\ \hat{v}^j_{Mk} &= v_k^{j-1}. \end{aligned} \tag{22}$$

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, and since H is a bounded function with respect to its argument, this multi-dimensional prediction operator is a Lipschitz nonlinear prediction operator of order 2. In a recent paper [3], another version of the multi-dimensional PPH-scheme was introduced. In the proposed framework, the scale j is associated to the location 2^{-j}, and the prediction operator is as follows:

$$
\begin{aligned}
\hat{v}^{j}_{2k+e_1} &= \frac{v^{j-1}_k + v^{j-1}_{k+e_1}}{2} - \frac{1}{8}H(\Delta^2_{e_1} v^{j-1}_k, \Delta^2_{e_1} v^{j-1}_{k-e_1}) \\
\hat{v}^{j,2}_{2k+e_2} &= \frac{v^{j-1}_k + v^{j-1}_{k+e_2}}{2} - \frac{1}{8}H(\Delta^2_{e_2} v^{j-1}_k, \Delta^2_{Me_2} v^{j-1}_{k-e_2}), \\
\hat{v}^{j,2}_{2k+e_1+e_2} &= \frac{9}{16}(v^{j-1}_k + v^{j-1}_{k+Me_1}) - \frac{1}{16}(v^{j-1}_{k-Me_1} + v^{j-1}_{k+2Me_1}) \\
\hat{v}^{j}_{2k} &= v^{j-1}_k.
\end{aligned} \tag{23}
$$

It is clear that this scheme is a particular example of Lipschitz and bounded nonlinear prediction operator since we can write:

$$
\hat{v}^{j,2}_{2k+e_1+e_2} = \frac{v^{j-1}_k + v^{j-1}_{k+Me_1}}{2} - \frac{1}{16}(\Delta^2_{Me_1} v^{j-1}_k + \Delta^2_{Me_1} v^{j-1}_{k-Me_1}).
$$

and then define S_l as the prediction operator associated with the linear part.

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 with bounded nonlinear prediction operators. The existence of such difference operators is ensured by the following theorem:

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

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

Proof Since S reproduces polynomials of degree N, the existence of $S^{(N+1)}$ was already proved in [23]. What is particular here is the form for the differences of order $N+1$:

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

From which, we deduce (putting $\Delta^{N+1} v = w$):

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

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 4 Let us consider a *bounded nonlinear* 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

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

In all the theorems that follow $v_j(x) = \sum_{k\in\mathbb{Z}^d} v_k^j \varphi_{j,k}(x)$, where φ satisfies (4) with g associated to the linear prediction operator S_l (for more details see (5)). Before we state a convergence (also called inverse) theorem for the multiscale respresentation, we need to establish some extensions to the non-separable case of a lemma proved in [22]:

Lemma 1 *Let S be a bounded nonlinear prediction operator of order $N+1$. Then, for any $k \leq N+1$*

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

Moreover, if $\rho_p(S^{(k)}) < m^{1/p}$, then for any ρ such that $\rho_p(S^{(k)}) < \rho < m^{1/p}$, we have

$$m^{-j/p} \|\Delta^k v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \lesssim \delta^j \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=0}^{j} \delta^{j-l} m^{-l/p} \|e^l\|_{\ell^p(\mathbb{Z}^d)} \tag{26}$$

where $\delta = \rho m^{-1/p}$.

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

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

Since S is a *bounded nonlinear* 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 bounded function in the sense of (7). Indeed, higher order finite differences can be expressed as linear combinations of lower order finite differences. To prove (26), by using the definition of the joint spectral radius, we note that for any $\rho_p(S^{(k)}) < \rho < m^{1/p}$, for all n and all v we have:

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

It follows that

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

Then putting as in [15], $\delta = \rho m^{-1/p}$, this finally leads to:

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

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

Theorem 2 *Let S be a bounded nonlinear 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

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

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:

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

The last equality being obtained remarking that $\sum_{s\geq 0} \delta^s = \frac{1}{1-\delta}$.

Remark 4 Usually, the convergence in L^p is associated with the condition $\rho_p(S^{(1)}) < m^{1/p}$. With a *bounded nonlinear* 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 $k \leq N+1$. This remark is of interest since there is no relation between $\rho_p(S^{(k)})$ and $\rho_p(S^{(k+1)})$. Thus, 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 exactly reproduce polynomials (because of Lemma 1).

Similarly, an inverse theorem can be written in Besov spaces:

Theorem 3 *Let S be a bounded nonlinear 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

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

The proof of (31) is similar to that of Theorem 5.4 of [24], so we will not expand on this here.

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 $v_j(x) = \sum_{k\in\mathbb{Z}^d} v^j_k \varphi_{j,k}(x)$, with $\varphi_{j,k}(x) = \varphi(M^j x - k)$ (and similarly for $\tilde{v}$).

In this section we assume that S obeys Definition 3, that is Φ_i is a bounded Lipschitz function.

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 4 *Let S be a Lipschitz and bounded nonlinear prediction operator of order $N+1$, and suppose that there exist a $\rho < m^{1/p}$ and an $n \in \mathbb{N}$ 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 $k \leq N+1$. 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:

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

Proof We note that for all v:

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

After $j=ns$ iterations of the above inequality, we get:

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

Then, for any j we may write:

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

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

$$m^{-j/p}\|\Delta^k(v^j-\tilde{v}^j)\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}\lesssim \delta^j\|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\sum_{l=1}^{j}\delta^{j-l}m^{-l/p}\|e^l-\tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}\quad(33)$$

Now, note that:

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

At this stage, we use the Lipschitz property of $\bar{\Phi}_i$ to get

$$\begin{aligned}
\|v-\tilde{v}\|_{L^p(\mathbb{R}^d)}&\lesssim\|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\sum_{j>0}m^{-j/p}\left(\|\Delta^k(v^{j-1}-\tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}+\|e^j-\tilde{e}^j\|_{\ell^p(\mathbb{Z}^d)}\right)\\
&\lesssim\|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\sum_{j>0}m^{-j/p}\|e^j-\tilde{e}^j\|_{\ell^p(\mathbb{Z}^d)},
\end{aligned}$$

the last inequality being obtained using (33) and then making the same computation as in Theorem 2.

The convergence and the stability of the nonlinear multiscale decomposition is thus based on the study of $S^{(k)}$ for some k. On the contrary, in [15], 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. To remark that the nonlinear prediction operator in Lipschitz and bounded may simplifies the proofs for convergence and for stability. However, we are aware that the more complex mathematical framework developed in [15] aims at dealing with a wider class of prediction operators (for instance, the median interpolating scheme studied in [15] is not a Lipschitz and bounded prediction operator).

Furthermore, we should also mention that a stability theorem was given in [3] for the scheme defined in (23) using the same kind of argument. However, the above theorem is more general in the sense that Lipschitz nonlinear prediction operators are not necessarily associated with an interpolatory multiscale representation.

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:

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

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

Theorem 5 *Let us assume that S is a Lipschitz nonlinear prediction operator of order $N+1$ such that there exist an n in $\mathbb{N}$ and a $\rho \le m^{1/p-s/d}$ for some $s>N$ such that:*

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

for some $k \le N+1$. 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:

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

The proof is the same as that of Theorem 6.2 of [24], except that we do not require the exact polynomial reproduction property.

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

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

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

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

Definition 5 A nonlinear prediction operator S is called bounded (resp. Lipschitz) $(\mathscr{A},I)$-compatible if there exists a local and linear prediction operator S_l such that S can be written under the following form:

$$Sv_{Mk+i} = S_l v_{Mk+i} + \Psi_i(\Delta^{\mathscr{A}_I} v_{k+p_1},\cdots,\Delta^{\mathscr{A}_I} v_{k+p_q}) \quad \forall i \in C(M)$$

where $\{p_1,\cdots,p_q\}$ is a fixed set, Ψ_i are bounded functions (resp. Lipschitz functions satisfying $\Psi_i(0)=0$) and if the operator S_l is such that there exists an operator $S_l^{\mathscr{A}_I}$ satisfying:

$$\Delta^{\mathscr{A}_I} S_l v = S_l^{\mathscr{A}_I} \Delta^{\mathscr{A}_I} v. \qquad (36)$$

Remark 5 From the definition of S, it is clear that there exists an operator $S^{\mathscr{A}_I}$. Note also that assuming equation (36) is true, the hypothesis of polynomial reproduction for S_l is no longer necessary. Indeed, we directly assume that the operator $S_l^{\mathscr{A}_I}$ uses the same differences as that used in the perturbation function Φ_i. We, also, remark that bounded (resp. Lipschitz) nonlinear prediction operators of order $N+1$ are bounded (resp. Lipschitz) $(\mathscr{A},I)$-compatible with $I=\{i;|i|=N+1\}$ and $\mathscr{A}=\{e_1,\dots,e_d\}$.

Note that we can then extend all the notions described in the previous sections, i.e., multiscale representation, joint spectral radius of $S^{\mathscr{A}_I}$, convergence and stability theorems, replacing bounded (resp. Lipschitz) nonlinear prediction operators by bounded (resp. Lipschitz) $(\mathscr{A},I)$-compatible nonlinear prediction operators. Indeed, if one computes $\|v_{j+1}-v_j\|$ as in Lemma 1 when v_j is computed using an $(\mathscr{A},I)$-compatible prediction operator, it is clear that the result holds provided that $\Delta^k v^j$ is replaced by $\Delta^{\mathscr{A}_I} v^j$. Then, to prove the convergence, one just needs to study $\rho_p(S^{\mathscr{A}_I})$ instead of $\rho_p(S^{(k)})$.

The interest of using the notion of $(\mathscr{A},I)$-compatibility is to provide proofs of convergence where the classical approach fails, as shown in the next section. The $(\mathscr{A},I)$-compatibility also enables to significantly reduce the number of computed differences to compute the joint spectral radius. From a practical point of view, given a prediction operator we first identify its type (i.e. bounded nonlinear or bounded $(\mathscr{A},I)$-compatible for instance) 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 $(\mathscr{A},I)$-compatibility does not make sense. Our point is to give an illustration of the new convergence and stability theorems (2 and 4 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 [15] (Theorem 2.3) or in [5] (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 $k \le N+1$, while in [15] 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 [15] is wider therefore the proofs for the stability are different.

Now, let us give an illustration of how Theorems 2 and 4 apply to the PPH-scheme (we will then see how the proof of convergence extends to the slightly modified power-P scheme introduced in (19)). Since the PPH-scheme is bounded nonlinear 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 multiscale representation associated to the PPH-scheme by finding an upper bound for $\rho_p(S^{(2)})$, whose expression is particularly simple since:

$$\begin{aligned} S^{(2)} w_{2i} &= \frac{1}{4} H(w_{i-1},w_i) \\ S^{(2)} w_{2i+1} &= \frac{w_i}{2} - \frac{1}{8}\left(H(w_{i-1},w_i)+H(w_i,w_{i+1})\right). \end{aligned} \tag{37}$$

Note that proofs of convergence and stability in L^∞ are available in [5]. Based on the simple expression of $S^{(2)}$, we are able to prove new convergence results for the multiscale representation associated to the PPH-scheme in L^p for $p \geq 1$ and also new stability results for $p > 1$ (see Appendix B for details).

If one considers the modified power-P scheme defined using $\tilde{H}_q$ (see (19)) and assumes that (w_i, w_{i+1}) belongs to $\mathbb{R}^2 \setminus \{(x,x), |x| \leq \sqrt{2}\varepsilon\}$ (see section 5.5), 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 in the proof of the convergence of the multiscale representation associated with the PPH-scheme (see Appendix B), we obtain that the modified power-P scheme leads to a convergent multiscale representation in $L^p(\mathbb{R})$ for any $p \geq 1$.

10.2 Convergence and Stability 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 2 *One has*

$$\sup_{u,w \in \ell^\infty(\mathbb{Z}^d)} \|S^{(1)}(u) S^{(1)}(w)\|_{\ell^\infty(\mathbb{Z}^d)} < 1$$

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

The proof is identical to that of Lemma 4 of [9], therefore we do not expand on it here. The multiscale representation based on the proposed WENO model is therefore convergent in L^∞.

To study the stability of the WENO scheme, we modify it a little into

$$\hat{v}^j_{2k+1,w} - \hat{v}^j_{2k+1} = \frac{\omega}{16}\left(\alpha^j_{k,1}\Delta^4 v^{j-1}_{k-1} + \alpha^j_{k,-1}\Delta^4 v^{j-1}_{k-2}\right). \tag{38}$$

We can then prove the stability of the multiscale representation using the first order difference. Indeed, we have:

$$S^{(1)}\Delta v_{2k} - S^{(1)}\Delta u_{2k} = S^{(1)}_l \Delta v_{2k} - S^{(1)}_l \Delta u_{2k} + \frac{\omega}{16}\left((\Delta^4 v_{k-2} - \Delta^4 u_{k-2}) + \right.$$
$$\left.\left(\alpha(\Delta^4 v_{k-2}, \Delta^4 v_{k-1})(\Delta^4 v_{k-1} - \Delta^4 v_{k-2}) - \alpha(\Delta^4 u_{k-2}, \Delta^4 u_{k-1})(\Delta^4 u_{k-1} - \Delta^4 u_{k-2})\right)\right)$$

Then we use the fact that $\rho_\infty(S^{(1)}_l) \leq \frac{5}{8}$ [9] and that $\alpha(x,y)(x-y)$ is a Lipschitz function when $\alpha = \frac{1}{1+(\frac{y}{x})^\beta}$, the Lipschitz constant being smaller than $1+2\beta$ (to obtain this result it suffices to compute the partial derivative of α with respect to x and y). Writing the fourth order differences in terms of first order differences, we finally get:

$$\|S^{(1)}\Delta v_{2k} - S^{(1)}\Delta u_{2k}\|_{\ell^\infty(\mathbb{Z}^d)} \leq \left(\frac{5}{8} + \frac{\omega}{2} + \omega(1+2\beta)\right)\|\Delta v_k - \Delta u_k\|_{\ell^\infty(\mathbb{Z}^d)}.$$

From this, we deduce that the multiscale representation is stable in L^∞, as soon as $\omega < \frac{3}{4}\frac{1}{3+4\beta}$. How to use such new stable WENO representation is beyond the scope of the present article.

10.3 Convergence and Stability of Bidimensional Multiscale Representations Based on PPH-scheme

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

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

for some $0 < \omega < 1$. To consider $\omega < 1$ instead of $\omega = 1$ as in (22) will appear clearer a bit later. In Appendix C, we show, using the fact that the nonlinear prediction operator is Lipschitz $(\mathscr{A}, I)$-compatible (with $\mathscr{A}$ and I being defined there), the following new results: the associated multiscale representation is convergent in L^∞ as soon as $\omega < 1$, convergent in L^p for any $p \geq 1$ and $\omega = 1$, while the multiscale representation is proved to be stable for $\omega < 1/2$.

The prediction operator defined in (23) is also a typical example of $(\mathscr{A}, I)$-compatible *bounded nonlinear* prediction operator. In that case, by construction $\mathscr{A} = (e_1, e_2, Me_1)$ and also $I = \{(2,0,0), (0,2,0), (0,0,2)\}$. As a proof of stability in L^∞ was already given in [3], we do not expand on it here.

11 Conclusion

In this paper, we have introduced a new formalism for nonlinear and non-separable multiscale representations. The introduced formalism includes some 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 bounded or 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 bounded or Lipschitz nonlinear prediction operators. After having illustrated the proposed formalism on one and multi-dimensional cases, we stated the convergence theorems in L^p and Besov spaces for multiscale representations based on bounded nonlinear prediction operator. We then stated the stability theorems in these spaces for multiscale representations based on Lipschitz and bounded nonlinear prediction operators. We also introduced the notion of bounded (resp. Lipschitz) $(\mathscr{A}, I)$-compatible prediction operators which behaves like bounded (resp. Lipschitz) nonlinear ones in terms of the convergence and the stability of the associated multiscale representation. We saw in applications that to use the $(\mathscr{A}, I)$-compatibility of the prediction operators enabled to give some new proofs of convergence and stability in L^p of the corresponding nonlinear multiscale representation. In terms of perspectives, we are currently investigating how to apply the model of Lipschitz and bounded nonlinear 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

$y_0 = 0, y_1 = v^{j-1}_{k-N}, y_2 = y_1 + v^{j-1}_{k-N+1}, \cdots, y_{2N+1} = y_{2N} + v^{j-1}_{k+N}$ 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:

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

The *centered* prediction following (13) is:

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

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 $\hat{v}^j_{2k,1}$ and $\hat{v}^j_{2k+1,1}$ are identical to that of $\hat{v}^j_{2k}$ and $\hat{v}^j_{2k+1}$ replacing P_{2N} by $P_{2N,1}$. Then, computing the difference between $P_{2N,1}$ and P_{2N} and applying it at $x = 2^{-j}k$, we get:

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

Appendix B

In this section, we study the convergence and the stability in $L^p(\mathbb{R}^d)$ of the multiscale representation associated with the one-dimensional PPH-scheme. To start with, we may write, assuming that $p \geq 1$:

$$\begin{aligned}|S^{(2)} w_{2i}|^p &\leq \frac{1}{4^p}\max(|w_{i-1}|, |w_i|)^p \\ |S^{(2)} w_{2i+1}|^p &\leq \left(\frac{1}{2}|w_i| + \frac{1}{8}\max(|w_{i-1}|, |w_i|) + \frac{1}{8}\max(|w_i|, |w_{i+1}|)\right)^p \\ &\leq \left(\frac{1}{2}|w_i| + \frac{1}{4}(\frac{1}{2}\max(|w_{i-1}|, |w_i|) + \frac{1}{2}\max(|w_i|, |w_{i+1}|))\right)^p \\ &\leq \frac{1}{2}|w_i|^p + \frac{1}{4}\frac{1}{2^p}\max(|w_{i-1}|, |w_i|)^p + \frac{1}{4}\frac{1}{2^p}\max(|w_i|, |w_{i+1}|)^p.\end{aligned} \tag{40}$$

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

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

The 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 check that the coefficient in front of $|w_i|^p$ is $\frac{2}{4^p}+\frac{1}{2}+\frac{1}{2^p}$, 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 $p \geq 1$),

$$|S^{(2)}w_{2i} - S^{(2)}v_{2i}|^p \leq \frac{1}{2^p}\max(|w_{i-1}-v_{i-1}|,|w_i-v_i|)^p$$
$$|S^{(2)}w_{2i+1} - S^{(2)}v_{2i+1}|^p \leq \frac{1}{2}|w_i-v_i|^p + \frac{1}{2}\max(|w_{i-1}-v_{i-1}|,|w_{i+1}-v_{i+1}|)^p,$$

the last inequality being a consequence of Lemma 2 of [5]. Now, as in the study of the convergence, we write:

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

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.

Appendix C

We already noticed that the nonlinear prediction operator defined in (39) is bounded. We now remark that this prediction operator is bounded $(\mathscr{A}, I)$-compatible with $\mathscr{A} = \{e_1, Me_1\}$ 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 $S^{\mathscr{A}_I}$. To this end, we compute the differences of order 2 in the directions $\{e_1, Me_1\}$, which are given by:

$$
\begin{aligned}
\Delta^2_{e_1}\hat{v}^j_{Mk} &= \frac{\omega}{4}H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) \\
\Delta^2_{e_1}\hat{v}^j_{Mk+e_1} &= \frac{1}{2}\Delta^2_{Me_1}v^{j-1}_k \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_{k+Me_1}, \Delta^2_{Me_1}v^{j-1}_k) \\
\Delta^2_{Me_1}\hat{v}^j_{Mk} &= \Delta^2_{e_1}v^{j-1}_k \\
\Delta^2_{Me_1}\hat{v}^j_{Mk+e_1} &= \frac{1}{2}(\Delta^2_{e_1}v^{j-1}_k + \Delta^2_{e_1}v^{j-1}_{k+Me_1}) + \frac{\omega}{4}H(\Delta^2_{Me_1}v^{j-1}_{k+e_1}, \Delta^2_{Me_1}v^{j-1}_{k+e_1-Me_1}) \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_k, \Delta^2_{Me_1}v^{j-1}_{k-Me_1}) \\
&- \frac{\omega}{8}H(\Delta^2_{Me_1}v^{j-1}_{k+2e_1}, \Delta^2_{Me_1}v^{j-1}_{k+2e_1-Me_1}).
\end{aligned} \tag{41}
$$

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 inequalities we immediately deduce that $\rho_\infty(S^{\mathscr{A}_I}) \leq \sqrt{\frac{1+\omega}{2}} < 1$, which proves that the bidimensional PPH defined by (39) is convergent in L^∞.

For the L^p convergence, we do not need the restriction on ω and we consider the model defined by (22), therefore we study:

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

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

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

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

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

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 (39) in L^p. We may indeed write:

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

From this we deduce that:

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

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 (39) leads in that case to a stable multiscale representation.

References

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, K. Dadourian, J. Liandrat, J. Ruiz,J.C. Trillo, A family of Stable Nonlinear Nonseparable multiresolution in 2D, J. Comput. Appl. Math., 234, 1277-1290, (2010).
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 Harmonic Analysis, 18, 198-206, (2006).
6. E. J. Candès, Ridgelets and the representation of mutilated Sobolev functions, SIAM J. Math. Anal., 33, 197-218, (1999)
7. 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).
8. A. Cohen Wavelets in numerical analysis, Handbook of Numerical Analysis, VII, P.G. Ciarlet and J.L. Lions, eds., Elsevier, Amsterdam, (2003).
9. A. Cohen, N. Dyn and B. Mateï, Quasi-linear Subdivision Schemes with Applications to ENO Interpolation, Appl. Comput. Harmon. Anal., 15, 89-116, (2003).
10. V. Chappelier and C. Guillemot, Oriented Wavelet for Image Compression and Denoising, IEEE Transactions on Image Processing, 15, 2892-2903, (2006).
11. I. Daubechies, O. Runborg and W. Sweldens, Normal multiresolution approximation of curves, Constr. Approx., 20, 399-463, (2004).
12. D. L. Donoho and T. P.-Y. Yu, Nonlinear Pyramid Transforms Based on Median-Interpolation, SIAM Journal of Math. Anal., 31, 1030-1061, (2000).
13. 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).
14. 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).
15. S. Harizanov and P. Oswald, Stability of Nonlinear Subdivision and multiscale Transforms, Constr. Approx., 31, 359-393, (2010).
16. A. Harten, Discrete Multiresolution Analysis and Generalized Wavelets, J. Appl. Num. Math., 12,153-193, (1993) .
17. 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).
18. I. Kompatsiaris, D. Tzovaras and M.G. Strintzis, Hierarchical Representation and Coding of Surfaces Using 3D Polygon Meshes, IEEE Transactions on Image Processing, 10, 1133-1151, (2001).
19. F. Kuijt and R. van Damme. Convexity preserving interpolatory subdivision schemes., Constr. Approx., 14, 609-630, (1998).
20. E. Le Pennec and S. Mallat, Sparse Geometric Image Representation with Bandelets, IEEE Trans. on Image Processing, 14, 423-438, (2005).
21. S. Malassiotis and M.G. Strintzis, 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).

22. B. Mateï, Smoothness Characterization and Stability in Nonlinear multiscale Framework: Theoretical Results Asymptotic Analysis, 46, 277-309, (2005).
23. B. Mateï, S. Meignen and A. Zakharova, Smoothness of Nonlinear and Non-Separable Subdivision Schemes, accepted for publication in Asymptotic Analysis (2011).
24. B. Mateï, S. Meignen and A. Zakharova, Smoothness Characterization and Stability of Nonlinear and Non-Separable Multiscale Representation, Journal of Approximation Theory (2011), doi:10.1016/j.jat.2011.06.2009.
25. P. Oswald, Smoothness of nonlinear median-interpolation subdivision, Adv. Comput. Math., 20, 401-423, (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).