Smoothness Characterization and Stability of Nonlinear and Non-Separable Multi-Scale Representations

Basarab Matei\textsuperscript{a}, Sylvain Meignen\textsuperscript{b}, Anastasia Zakharova\textsuperscript{c}

\textsuperscript{a}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

\textsuperscript{b}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

\textsuperscript{c}LJK Laboratory, University of Grenoble, France
Tel:0033-4-76-51-43-95
FAX:0033-4-76-63-12-63
E-mail: anastasia.zakharova@imag.fr

Abstract

The aim of the paper is the construction and the analysis of nonlinear and non-separable multi-scale representations for multivariate functions. The proposed multi-scale representation is associated with a non-diagonal dilation matrix $M$. We show that the smoothness of a function can be characterized by the rate of decay of its multi-scale coefficients. We also study the stability of these representations, a key issue in the designing of adaptive algorithms.

Key words: Nonlinear Multiscale approximation, Besov Spaces, Stability.

2000 MSC: 41A46, MSC 41A60, MSC 41A63

1. Introduction

A multi-scale representation of an abstract object $v$ (e.g. a function representing the grey level of an image) is defined as $\mathcal{M}v := (v^0, d^0, d^1, d^2, \cdots)$,
where $v^0$ is the coarsest approximation of $v$ in some sense and $d^j$, with $j \geq 0$, are additional detail coefficients representing the fluctuations between two successive levels.

Several strategies exist to build such representations: wavelet basis, lifting schemes and also the discrete framework of Harten [8]. Using a wavelet basis, we compute $(v^0, d^0, d^1, d^2, \cdots)$ through linear filtering and thus the multi-scale representation corresponds to a change of basis. Although wavelet bases are optimal for one-dimensional functions, this is no longer the case for multivariate objects such as images where the presence of singularities requires special treatments. The approximation property of wavelet bases and their use in image processing are now well understood (see [5] and [12] for details).

Overcoming this ”curse of dimensionality” for wavelet basis was in the past decade the subject of active research. We mention here several strategies developed from the wavelets theory: the curvelets transforms [3], the directionlets transforms [6] and the bandelets transform [11]. Another approach proposed in [13] and studied in [2] uses the discrete framework of Harten, which allows a better treatment of singularities and consequently better approximation results.

The applications of all these methods to image processing are numerous: let us mention some of these works in [2], [1] and [4]. In [2], the extension of univariate methods using tensor product representations is studied. Although this extension is natural and simple, the results are not optimal.

We propose in the present paper a nonlinear multi-scale representation based on the general framework of A. Harten (see [8] and [9]). Our represen-
tation is non-separable and is associated with a non-diagonal dilation matrix $M$. Since the details are computed adaptively, the multi-scale transform is completely nonlinear and is no more equivalent to a change of basis. To study these representations, we develop some new analysis tools and we prove that these representations give the same approximation order as for wavelet basis. This strategy is fruitful in applications since it allows to cope up with the deficiencies of wavelet bases without loosing the approximation order.

The outline of the paper is the following. After having introduced nonlinear and non-separable multi-scale representations, we give an illustration of potential interest of using non-diagonal dilation matrix for image compression (section 3). Extending the results in [14], we characterize the smoothness of a function $v$ belonging to some Besov spaces by means of the decay of the detail coefficients of its nonlinear and non-separable multiscale representation (section 4 and 5). In section 6, we study the stability of the underlying subdivision scheme and we then switch on to the stability of the multi-scale representation in section 7 (for similar, one-dimensional results see [17] and [14]). To conclude the paper, we give an illustration of a covering multi-scale representation in two dimensions (see section 8).

2. Multi-scale Representations

For the reader convenience, we recall the construction of linear multi-scale representations based on multiresolution analysis (MRA) of some $d$-dimensional Hilbert space $V$. To this end, let $M$ be a $d \times d$ dilation matrix.

**Definition 1.** A multiresolution analysis of $V$ is a sequence $(V_j)_{j \in \mathbb{Z}^d}$ of closed subspaces of $V$ satisfying the following properties:
1. The subspaces are embedded: \( V_j \subset V_{j+1} \);
2. \( f \in V_j \) if and only if \( f(M_\cdot) \in V_{j+1} \);
3. \( \bigcup_{j \in \mathbb{Z}} V_j = V \);
4. \( \bigcap_{j \in \mathbb{Z}} V_j = \{0\} \);
5. There exists a compactly supported function \( \varphi \in V_0 \) such that the family \( \{\varphi(\cdot - k)\}_{k \in \mathbb{Z}^d} \) forms a Riesz basis of \( V_0 \).

The function \( \varphi \) is called the scaling function. Since \( V_0 \subset V_1 \), \( \varphi \) satisfies the following equation:

\[
\varphi = \sum_{k \in \mathbb{Z}^d} g_k \varphi(M \cdot - k), \text{ with } \sum_k g_k = m. \tag{1}
\]

To get the approximation of a given function \( v \) at the level \( j \), we assume the existence of a compactly supported function \( \tilde{\varphi} \) dual to \( \varphi \) (i.e. for all \( k, n \in \mathbb{Z}^d \), \( \langle \tilde{\varphi}(\cdot - n), \varphi(\cdot - k) \rangle = \delta_{n,k} \), where \( \delta_{n,k} \) denotes the Kronecker symbol and \( \langle \cdot, \cdot \rangle \) the inner product on \( V \)), which satisfies a so-called scaling equation

\[
\tilde{\varphi} = \sum_{n \in \mathbb{Z}^d : \|n\|_\infty \leq P} \tilde{h}_n \tilde{\varphi}(M \cdot - n), \text{ with } \sum_k \tilde{h}_k = m. \tag{2}
\]

The approximation \( v_j \) of \( v \) we consider is then obtained by projection of \( v \) on \( V_j \) as follows:

\[
v_j = \sum_{n \in \mathbb{Z}^d} v^j_n \varphi(M^j \cdot - n). \tag{3}
\]

where

\[
v^j_n = \int v(x) m^j \tilde{\varphi}(M^j x - n) dx, \quad n \in \mathbb{Z}^d. \tag{4}
\]

Multi-scale representations based on specific choice for \( \tilde{\varphi} \) are commonly used in image processing and numerical analysis. We mention two of them: the first one is the point value case obtained when \( \tilde{\varphi} \) is the Dirac distribution.
and the second one is the cell average case obtained when \( \tilde{\varphi} \) is the indicator function of some domain on \( \mathbb{R}^d \). In the theoretical study that follows, we assume that the data are obtained through a projection of a functional \( v \) as in (4).

A strategy which allows to build nonlinear multi-scale representations based on such projection can be done in terms of a very general discrete framework based on the concept of inter-scale operators introduced by A. Harten in [8], which we now recall. Assume that we have two inter-scale discrete operators associated to this sequence: the projection operator \( P_{j-1}^j \) and the prediction operator \( P_{j}^{j-1} \). The projection operator \( P_{j-1}^j \) acts from fine to coarse level, that is, \( v^{j-1} = P_{j-1}^j v^j \). This operator is always assumed to be linear. The prediction operator \( P_{j}^{j-1} \) acts from coarse to fine level. It computes the ‘approximation’ \( \hat{v}^j \) of \( v^j \) from the vector \((v_{k-1}^{j-1})_{k \in \mathbb{Z}^d} \) which is associated to \( v_{j-1} \in V_{j-1} \):

\[
\hat{v}^j = P_{j}^{j-1}v^{j-1}.
\]

This operator may be nonlinear. Besides, we assume that these operators satisfy the consistency property:

\[
P_{j-1}^j P_j^{j-1} = I, \tag{5}
\]

i.e., the projection of \( \hat{v}^j \) coincides with \( v^{j-1} \). Having defined the prediction error \( e^j := v^j - \hat{v}^j \), we obtain a redundant representation of vector \( v^j \):

\[
v^j = \hat{v}^j + e^j. \tag{6}
\]

By the consistency property, one has

\[
P_{j-1}^j e^j = P_{j-1}^j v^j - P_{j-1}^j \hat{v}^j = v^{j-1} - v^{j-1} = 0.
\]

5
Hence, \( e^j \in \text{Ker}(P_{j-1}^j) \). Using a basis of this kernel, we write the error \( e^j \) in a non-redundant way and get the detail vector \( d^{j-1} \). The data \( v^j \) is thus completely equivalent to the data \((v^{j-1}, d^{j-1})\). Iterating this process from the initial data \( v^J \), we obtain its **nonlinear multi-scale representation**

\[
\mathcal{M}v^J = (v^0, d^0, \ldots, d^{J-1}).
\]  

(7)

From here on, we assume the equivalence

\[
\|e^j\|_{\ell^p(\mathbb{Z}^d)} \sim \|d^{j-1}\|_{\ell^p(\mathbb{Z}^d)}.
\]  

(8)

The details are computed adaptively, then the underlying multi-scale transform is nonlinear and no more equivalent to a change of basis. Moreover, the discrete setting used here is not based on the study of scaling equations as for wavelet basis, which implies that the results of wavelet theory cannot be used in our analysis. Note also that the projection operator is completely characterized by the function \( \tilde{\varphi} \). Namely, if we consider the discretization defined by (4) then, in view of (2), we may write the projection operator as follows:

\[
v_{j-1}^k = m^{-1} \sum_{\|n\|_{\infty} \leq P} \tilde{h}_n v_{j}^{Mk+n} = m^{-1} \sum_{\|n-Mk\|_{\infty} \leq P} \tilde{h}_{n-Mk} v_{j}^{n} := (P_{j-1}^j v^j)_k.
\]  

(9)

To describe the prediction operator, for every \( w \in \ell^\infty(\mathbb{Z}^d) \) we consider a linear operator \( S(w) \) defined on \( \ell^\infty(\mathbb{Z}^d) \) by

\[
(S(w)u)_k := \sum_{l \in \mathbb{Z}^d} a_{k-Ml}(w) u_l, \quad k \in \mathbb{Z}^d.
\]  

(10)

Note that the coefficients \( a_k(w) \) depend on \( w \). We assume that \( S \) is local:

\[
\exists K > 0 \text{ such that } a_{k-Ml}(w) = 0 \text{ if } \|k - Ml\|_{\infty} > K.
\]  

(11)
and that \( a_k(w) \) is bounded independently of \( w \):

\[ \exists C > 0 \text{ such that } \forall w \in \ell^\infty(\mathbb{Z}^d) \quad \forall k,l \in \mathbb{Z}^d \quad |a_{k-Ml}(w)| < C. \quad (12) \]

**Remark 2.1.** From (12) it immediately follows that for any \( p \geq 1 \) the norms \( \|S(w)\|_{\ell^p(\mathbb{Z}^d)} \) are bounded independently of \( w \).

The *quasi-linear* prediction operator is then defined by

\[ \hat{v}^j = P_j^{-1} v^{j-1} = S(v^{j-1}) v^{j-1}. \quad (13) \]

If for all \( k,l \in \mathbb{Z}^d \) and all \( w \in \ell^\infty(\mathbb{Z}^d) \) we put \( a_{k-Ml}(w) = g_{k-Ml} \), where \( g_{k-Ml} \) is defined by the scaling equation (1), we get the so-called *linear* prediction operator. In the general case, the prediction operator \( P_j^{-1} \) could be considered as a perturbation of the *linear* prediction operator because of the *consistency* property, that is why we will call it *quasi-linear* prediction operator. The operator-valued function which associates to any \( w \) an operator \( S(w) \) is called a *quasi-linear* prediction rule.

For what follows, we need to introduce the notion of polynomial reproduction for *quasi-linear* prediction rules. A polynomial \( q \) of total degree \( N \) is defined as a linear combination \( q(x) = \sum_{|n| \leq N} c_n x^n \). Let us denote by \( \Pi \) the linear space of all polynomials, by \( \Pi_N \) the linear space of all polynomials of total degree \( N \). With this in mind, we have the following definition for polynomial reproduction:

**Definition 2.1.** We will say that the *quasi-linear* prediction rule \( S(w) \) reproduces polynomials of total degree \( N \) if for any \( w \in \ell^\infty(\mathbb{Z}^d) \) and any \( u \in \ell^\infty(\mathbb{Z}^d) \) such that \( u_k = p(k) \ \forall k \in \mathbb{Z}^d \) for some \( p \in \Pi_N \), we have:

\[ (S(w)u)_k = p(M^{-1}k) + q(k), \]
where \( \deg(q) < \deg(p) \). If \( q = 0 \), we say that the quasi-linear prediction rule \( S \) exactly reproduces polynomials of total degree \( N \).

Note that the property is required for any data \( w \), and not only for \( w = u \).

3. On the Interest of Using Non-Diagonal Matrix for Image Compression

In the section, we shed light on the interest of using non-diagonal matrix for image compression. The dilation matrix we use is the quincunx matrix defined by:

\[
M = \begin{pmatrix}
-1 & 1 \\
1 & 1
\end{pmatrix},
\]

whose coset vectors are \( \epsilon_0 = (0,0)^T \) and \( \epsilon_1 = (0,1)^T \). We consider an interpolatory multi-scale representation which implies that \( v^K_j = v(M^{-j}k) \) (i.e. \( \hat{\varphi} \) is the Dirac function). By construction \( v^{j}_{Mk} = v^{j-1}_{k} \), so we only need to predict \( v(M^{-j}k + \epsilon_1) \). To do so, we define four polynomials of degree 1 (i.e. \( a + bx + cy \)) interpolating \( v \) on the following stencils:

\[
\begin{align*}
V^{j,1}_k &= M^{-j+1}\{k, k + e_1, k + e_2\} \\
V^{j,2}_k &= M^{-j+1}\{k, k + e_1, k + e_1 + e_2\} \\
V^{j,3}_k &= M^{-j+1}\{k + e_1, k + e_2, k + e_1 + e_2\} \\
V^{j,4}_k &= M^{-j+1}\{k, k + e_2, k + e_1 + e_2\},
\end{align*}
\]

which in turn entails the following two predictions for \( v(M^{-j}k + \epsilon_1) \):

\[
\hat{v}_{Mk+\epsilon_1}^{j,1} = \frac{1}{2}(v_{k}^{j-1} + v_{k+e_1+e_2}^{j-1})
\]

(14)
\[ \nu_{Mk+e_1}^{j/2} = \frac{1}{2}(v_{k+e_1}^{j-1} + v_{k+e_2}^{j-1}). \]  

(15)

We now show that to choose between the two predictions (14) and (15) appropriately improves the compression performance on natural images. The cost function we use to make the choice of stencil is as follows:

\[ C^j_k = \min(|v_{k+e_1}^{j-1} - v_{k+e_2}^{j-1}|, |v_k^{j-1} - v_{k+e_1+e_2}^{j-1}|). \]

When the minimum of \( C^j_k \) corresponds to the first (resp. second) argument the prediction (15) (resp. (14)) is used. The motivation for the choice of such a cost function is the following: when an edge intersect the cell \( Q_{j-1}^k \) defined by the vertices \( \{k, k+e_1, k+e_2, k+e_1+e_2\} \), several cases may happen:

1. either the edge intersect \([M^{-j+1}k, M^{-j+1}(k + e_1)] \) and \([M^{-j+1}(k + e_1), M^{-j+1}(k + e_2)] \) in which case no direction is favored.
2. or the edge intersect \([M^{-j+1}k, M^{-j+1}(k + e_2)] \) or \([M^{-j+1}(k + e_1), M^{-j+1}(k + e_2)] \), in which case the prediction operator favors the direction which is not intersected by the edge (this will lead to a better prediction).

When \( Q_{j-1}^k \) is not intersected by an edge, the gain between choosing one direction or the other is negligible and, in that case, we will apply predictions (14) and (15) successively. It thus remains to determine when a cell is intersected by an edge. In our procedure, we will test two different conditions
[((C_1) \text{ and } (C_2)) \text{ to determine these cells. The first one is:}

\[
\argmin_{k' = k,k+e_1,k-e_1} \left( |v_{k'}^{j-1} - v_{k+e_1}^{j-1}| + |v_{k'}^{j-1} - v_{k+e_2}^{j-1}| + |v_{k'}^{j-1} - v_{k+e_1+e_2}^{j-1}| + |v_{k'+e_1}^{j-1} - v_{k'+e_2}^{j-1}| + |v_{k'+e_1+e_2}^{j-1}| + |v_{k'}^{j-1} - v_{k'+e_1+e_2}^{j-1}| \right) = 1
\]

\text{or}

\[
\argmin_{k' = k,k+e_2,k-e_2} \left( |v_{k'}^{j-1} - v_{k+e_1}^{j-1}| + |v_{k'}^{j-1} - v_{k+e_2}^{j-1}| + |v_{k'}^{j-1} - v_{k'+e_1+e_2}^{j-1}| + |v_{k'+e_1}^{j-1} - v_{k'+e_2}^{j-1}| + |v_{k'+e_1+e_2}^{j-1}| + |v_{k'}^{j-1} - v_{k'+e_1+e_2}^{j-1}| \right) = 1 \quad (C_1),
\]

which corresponds to the case where the average first order differences are locally maximum in either vertical or horizontal direction. The second is:

\[
\argmin_{k' = k,k+e_1,k-e_2} |v_{k'}^{j-1} - v_{k+e_1+e_2}^{j-1}| = 1 \text{ or }
\argmin_{k' = k,k+e_1,k-e_2} |v_{k'}^{j-1} - v_{k'+e_2}^{j-1}| = 1 \quad (C_2),
\]

which corresponds to the case where the first order differences are locally maximum is the direction of prediction. Then, to encode the representation,

\text{we use an adapted version to our context of the EZW (Embedded Zero-tree}

Figure 1: (A): a 256 × 256 Lena image, (B): a 256 × 256 peppers image

we use an adapted version to our context of the EZW (Embedded Zero-tree
Wavelet) [18]. To simplify, consider a $N \times N$ image with $N = 2^J$, with $J$ even, then $d^j$ (defined in (7)) is associated to a $2^j \times 2^{j+1}$ matrix of coefficients when $j$ is odd and $d^j$ is associated to a $2^j \times 2^j$ matrix of coefficients when $j$ is even. We denote by $T^j_1$ the number of lines of the matrix associated to $d^j$. We display the compression results for the $256 \times 256$ images of Figure 1 on Figure 2 (A) and (B). To detect "edge-cells" where we apply nonlinear prediction we have either used conditions ($C_1$) or ($C_2$). Furthermore, we use nonlinear prediction to compute $d^j$ only when $T^j_1 \geq T_1$. It means that for instance, if we take $T_1 = 64$ and $N = 256$, we only predict nonlinearly the last finest four detail coefficients subspaces. A typical compression result is labeled by $ENO, C_1, T_1 = 64$ where ENO is for "Essentially non Oscillatory", $C_1$ means we use conditions ($C_1$) to detect the edges and $T_1 = 64$ means only the last four detail coefficients are predicted with the nonlinear procedure.

In any case, using nonlinear predictions leads to better compression results than the linear one (based on a diagonal dilation matrix). These encouraging results motivate a deeper study of nonlinear multi-scale representations associated with non-diagonal dilation matrix which we tackle in the following sections.

4. Notations and Generalities

We start by introducing some notations that will be used throughout the paper. Let us consider a multi-index $\mu = (\mu_1, \mu_2, \ldots, \mu_d) \in \mathbb{N}^d$ and a vector $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. We define $|\mu| = \sum_{i=1}^{d} \mu_i$ and $x^\mu = \prod_{i=1}^{d} x_i^\mu$. For two
Figure 2: (A): linear prediction (solid line) and ENO prediction for varying $T_1$ using either $C_1$ or $C_2$ to compute cells containing an edge for the image of Lena, (B): idem but for the image of peppers.

Multi-indices $m, \mu \in \mathbb{N}^d$ we define

\[
\begin{pmatrix}
\mu \\
m
\end{pmatrix} = \begin{pmatrix}
\mu_1 \\
m_1
\end{pmatrix} \cdots \begin{pmatrix}
\mu_d \\
m_d
\end{pmatrix}.
\]

For a fixed integer $N \in \mathbb{N}$, we define

\[ q_N = \# \{ \mu, |\mu| = N \} \]  \hspace{1cm} (16)

where $\#Q$ stands for the cardinal of the set $Q$. Let $\ell(\mathbb{Z}^d)$ be the space of all sequences indexed by $\mathbb{Z}^d$. The subspace of bounded sequences is denoted by $\ell^\infty(\mathbb{Z}^d)$ and $\|u\|_{\ell^\infty(\mathbb{Z}^d)}$ is the supremum of $\{|u_k| : k \in \mathbb{Z}^d\}$. As usual, let $\ell^p(\mathbb{Z}^d)$ be the Banach space of sequences $u$ on $\mathbb{Z}^d$ such that $\|u\|_{\ell^p(\mathbb{Z}^d)} < \infty$, where

\[ \|u\|_{\ell^p(\mathbb{Z}^d)} := \left( \sum_{k \in \mathbb{Z}^d} |u_k|^p \right)^{1/p} \]  for $1 \leq p < \infty.$
We denote by $L^p(\mathbb{R}^d)$, the space of all measurable functions $v$ such that

$$\|v\|_{L^p(\mathbb{R}^d)} < \infty,$$

where

$$\|v\|_{L^p(\mathbb{R}^d)} := \left( \int_{\mathbb{R}^d} |v(x)|^p \, dx \right)^{\frac{1}{p}} \text{ for } 1 \leq p < \infty,$$

$$\|v\|_{L^\infty(\mathbb{R}^d)} := \text{ess sup}_{x \in \mathbb{R}^d} |v(x)|.$$

Throughout the paper, the symbol $\| \cdot \|_{\infty}$ is the sup norm in $\mathbb{Z}^d$ when applied either to a vector or a matrix. Let us recall that, for a function $v$, the finite difference of order $N \in \mathbb{N}$, in the direction $h \in \mathbb{R}^d$ is defined by:

$$\nabla_h^N v(x) := \sum_{k=0}^{N} (-1)^k \binom{N}{k} v(x + kh).$$

and the mixed finite difference of order $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$ in the direction $h = (h_1, \ldots, h_d) \in \mathbb{R}^d$ by:

$$\nabla_h^n v(x) := \nabla_{h_{1} e_1} \cdots \nabla_{h_d e_d} v(x) = \sum_{k_1, \ldots, k_d = 0}^{\max(n_1, \ldots, n_d)} (-1)^{|n|} \binom{n}{k} v(x + k \cdot h),$$

where $k \cdot h := \sum_{i=1}^{d} k_i h_i$ is the usual inner product while $(e_1, \ldots, e_d)$ is the canonical basis on $\mathbb{Z}^d$. For any invertible matrix $B$ we put

$$\nabla_B^n v(x) := \nabla_{Be_1} \cdots \nabla_{Be_d} v(x).$$

Similarly, we define $D^nu(x) = D_1^{n_1} \cdots D_d^{n_d} v(x)$, where $D_j$ is the differential operator with respect to the $j$th coordinate of the canonical basis. For a sequence $(u_p)_{p \in \mathbb{Z}^d}$, we will use the mixed finite differences of order $N$ defined by the formulas

$$\nabla^n u := \nabla_{e_1}^{n_1} \nabla_{e_2}^{n_2} \cdots \nabla_{e_d}^{n_d} u.$$
where $\nabla_{e_i}^n$ is defined recursively by

$$\nabla_{e_i}^n u_k = \nabla_{e_i}^{n-1} u_{k+e_i} - \nabla_{e_i}^{n-1} u_k.$$ 

Then we put:

$$\Delta^N u : = \{\nabla^n u, |n| = N, n \in \mathbb{N}^d\}.$$ 

In the following, we will consider dilation matrices to define inter-scale operators. A dilation matrix is an invertible integer-valued matrix $M$ satisfying

$$\lim_{n \to \infty} M^{-n} = 0,$$

and $m := |\det(M)|$. Besides, we will use the definition of isotropic matrices:

**Definition 4.1.** A matrix $M$ is called isotropic if it is similar to the diagonal matrix $\text{diag}(\sigma_1, \ldots, \sigma_d)$, i.e. there exists an invertible matrix $\Lambda$ such that

$$M = \Lambda^{-1} \text{diag}(\sigma_1, \ldots, \sigma_d) \Lambda,$$

with $\sigma_1, \ldots, \sigma_d$ being the eigenvalues of matrix $M$, $\sigma := |\sigma_1| = \ldots = |\sigma_d| = m^{\frac{1}{d}}$.

Moreover, for any given norm in $\mathbb{R}^d$ there exist constants $C_1, C_2 > 0$ such that for any integer $n$ and for any $v \in \mathbb{R}^d$

$$C_1 m^{\frac{n}{d}} \|v\| \leq \|M^n v\| \leq C_2 m^{\frac{n}{d}} \|v\|.$$

We end this section with the following remark on notations: 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 $A \leq C B$. Also $A \sim B$ means $A \lesssim B$ and $B \lesssim A$. 

14
4.1. Besov Spaces

Let us recall the definition of Besov spaces. Let \( p, q \geq 1 \), \( s \) be a positive real number and \( N \) be any integer such that \( N > s \). The Besov space \( B^s_{p,q}(\mathbb{R}^d) \) consists of those functions \( v \in L^p(\mathbb{R}^d) \) satisfying

\[
(2^{js}\omega_N(v,2^{-j}L^p))_{j \geq 0} \in \ell^q(\mathbb{Z}^d),
\]

where \( \omega_N(v,t) \) is the modulus of smoothness of \( v \) of order \( N \in \mathbb{N} \setminus \{0\} \) in \( L^p(\mathbb{R}^d) \):

\[
\omega_N(v,t)_L^p = \sup_{h \in \mathbb{R}^d \atop \|h\|_2 \leq t} \|\nabla^N_h v\|_{L^p(\mathbb{R}^d)}, \quad t \geq 0,
\]

where \( \| \cdot \|_2 \) is the Euclidean norm. The norm in \( B^s_{p,q}(\mathbb{R}^d) \) is then given by

\[
\|v\|_{B^s_{p,q}(\mathbb{R}^d)} := \|v\|_{L^p(\mathbb{R}^d)} + \|(2^{js}\omega_N(v,2^{-j}L^p))_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)}.
\]

Let us now introduce a new modulus of smoothness \( \tilde{\omega}_N \) that uses mixed finite differences of order \( N \):

\[
\tilde{\omega}_N(v,t)_L^p = \sup_{n \in \mathbb{N}^d \atop |n| = N} \sup_{h \in \mathbb{R}^d \atop \|h\|_2 \leq t} \|\nabla^N_h v\|_{L^p(\mathbb{R}^d)}, \quad t > 0.
\]

It is easy to see that for any \( v \) in \( L^p(\mathbb{R}^d) \), \( \|\nabla^N_h v\|_{L^p(\mathbb{R}^d)} \lesssim \sum_{|n| = N} \|\nabla^n_h v\|_{L^p(\mathbb{R}^d)} \), thus \( \omega_N(v,t)_L^p \lesssim \tilde{\omega}_N(v,t)_L^p \). The inverse inequality \( \tilde{\omega}_N(v,t)_L^p \lesssim \omega_N(v,t)_L^p \) immediately follows from Lemma 4 of [17]. It implies that:

\[
\|v\|_{B^s_{p,q}(\mathbb{R}^d)} \sim \|v\|_{L^p(\mathbb{R}^d)} + \|(2^{js}\omega_N(v,2^{-j}L^p))_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)}.
\]

Going further, there exists a family of equivalent norms on \( B^s_{p,q}(\mathbb{R}^d) \).

**Lemma 4.1.** For all \( \sigma > 1 \), \( \|v\|_{B^s_{p,q}(\mathbb{R}^d)} \sim \|v\|_{L^p(\mathbb{R}^d)} + \|(\sigma^{js}\omega_N(v,\sigma^{-j}L^p))_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)} \).
Proof: Since \( \sigma > 1 \), for any \( j > 0 \) there exists \( j' > 0 \) such that \( 2^{j'} \leq \sigma^j \leq 2^{j'+1} \). According to this, we have the inequalities

\[
2^{j'} s \tilde{\omega}_N(v, 2^{-j'-1})_{L^p} \leq \sigma^j s \tilde{\omega}_N(v, \sigma^{-j})_{L^p} \leq 2^{(j'+1)} s \tilde{\omega}_N(v, 2^{-j'})_{L^p},
\]

from which the norm equivalence follows.

5. Smoothness of Nonlinear Multi-scale Representations

In this section, we prove the equivalence between the norm of a function \( v \) belonging to \( B^s_{p,q}(\mathbb{R}^d) \) and the discrete quantity computed using its nonlinear detail coefficients \( d^j \). Lower estimates of the Besov norm are associated to a so-called direct theorem while upper estimates are associated to a so-called inverse theorem. Note that a similar technique was applied in [5] in a wavelet setting.

5.1. Direct Theorem

Let \( v \) be a function in some Besov space \( B^s_{p,q}(\mathbb{R}^d) \) with \( p, q \geq 1 \) and \( s > 0 \), \((v^0, (d^j))_{j \geq 0}\) be its nonlinear multi-scale representation. We now show under what conditions we are able to get a lower estimate of \( \| v \|_{B^s_{p,q}(\mathbb{R}^d)} \) using \((v^0, (d^j))_{j \geq 0}\). To prove such a result, we need to have first an estimate of the norm of the prediction error:

**Lemma 5.1.** Assume that the quasi-linear prediction rule \( S(w) \) exactly reproduces polynomials up to degree \( N - 1 \) then the following estimation holds

\[
\| e^j \|_{L^p(\mathbb{Z}^d)} \lesssim \sum_{|n|=N} \| \nabla^n v^j \|_{L^p(\mathbb{Z}^d)},
\]

(18)
Proof: Let us compute

\[ e_j^k(w) := v_k^j - \sum_{\|k-Ml\|_\infty \leq K} a_{k-Ml}(w)v_l^{j-1}. \]

Using (9), we can write it down as

\[
e_j^k(w) = v_k^j - m^{-1} \sum_{l \in \mathbb{Z}^d} a_{k-Ml}(w) \sum_{n \in \mathbb{Z}^d} \|n-Ml\|_\infty \leq P \|k-n\|_\infty \leq K K + P \phi(n) = b_k(n) v_n^j, \]

where \( b_k,n(w) = \sum_{l \in \mathbb{Z}^d} a_{k-Ml}(w)h_{n-Ml} \), and \( F(k) = \{ n \in \mathbb{Z}^d : \|n-k\|_\infty \leq P + K \} \) is a finite set for any given \( k \). For any \( k \in \mathbb{Z}^d \) let us define a vector \( b_k(w) := (b_k,n(w))_{n \in F(k)} \). By hypothesis, \( e_j^k(w) = 0 \) if there exists \( p \in \Pi_{N'} \), \( 0 \leq N' < N \) such that \( v_k = p(k) \). Consequently, for any \( q \in \mathbb{Z}^d, |q| < N \), \( b_k(w) \) is orthogonal to any polynomial sequence associated to the polynomial \( l^q = l_1^q \cdots l_d^q \), thus it can be written in terms of a basis of the space orthogonal to the space spanned by these vectors. According to [10], Theorem 4.3, we can take \( \{ \nabla^\mu \delta_{-l}, |\mu| = N, l \in \mathbb{Z}^d \} \) as a basis of this space. By denoting \( e_l^\mu(w) \) the coordinates of \( b_k(w) \) in this basis, we may write:

\[
b_{k,n}(w) = \sum_{|\mu|=N} \sum_{l \in \mathbb{Z}^d} e_l^\mu(w) \nabla^\mu \delta_{n-l} \]

and taking \( w = v^{j-1} \) we get

\[
e_k^j := e_k^j(v^{j-1}) = \sum_{n \in F(k)} \sum_{|\mu|=N} \sum_{l \in \mathbb{Z}^d} e_l^\mu(v^{j-1}) \nabla^\mu \delta_{n-l} v_n^j = \sum_{n \in F(k)} \sum_{|\mu|=N} e_n^\mu(v^{j-1}) \nabla^\mu v_n^j. \]

(19)
Finally, we use (12) to conclude that the coefficients \( b_{k,n}(v^{j-1}) \) and \( c^j_t(v^{j-1}) \) are bounded independently of \( k, n \) and \( w \), and (18) follows from (19).

The preceding lemma enables us to compute the lower estimate:

**Theorem 5.1.** If for \( p, q \geq 1 \) and some positive \( s \), \( v \) belongs to \( B_{p,q}^s(\mathbb{R}^d) \), if the quasi-linear prediction rule exactly reproduces polynomials up to degree \( N - 1 \) with \( N > s \), if the matrix \( M \) is isotropic and if the equivalence (8) is satisfied, then

\[
\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j})(d^j_k)_{k \in \mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \geq 0 \|e_j(\mathbb{Z}^d) \lesssim \|v\|_{B_{p,q}^s(\mathbb{R}^d)}. \tag{20}
\]

**Proof:** Using the Hölder inequality and the fact that \( \tilde{\varphi} \) is compactly supported, we first obtain

\[
\|v^0\|_{\ell^p(\mathbb{Z}^d)} = \|(v, \tilde{\varphi}(\cdot - k))\|_{\ell^p(\mathbb{Z}^d)} \lesssim \|(v)_{L^p(\text{Supp}(\tilde{\varphi}(\cdot - k)))}\|_{\ell^p(\mathbb{Z}^d)} \lesssim \|v\|_{L^p(\mathbb{R}^d)}.
\]

Let us then consider a quasi-linear prediction rule which exactly reproduces polynomials up to degree \( N - 1 \). Since \( \|e_j\|_{\ell^p(\mathbb{Z}^d)} \sim \|d^{j-1}\|_{\ell^p(\mathbb{Z}^d)} \), by Lemma 5.1 we get

\[
\|(m^{(s/d-1/p)j})(d^j_k)_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} \lesssim \|(m^{(s/d-1/p)j}) \sum_{|n|=N} \|\nabla^n v_j\|_{\ell^p(\mathbb{Z}^d)} \|_{\ell^p(\mathbb{Z}^d)}.
\]

We have successively

\[
\sum_{|n|=N} \|\nabla^n v_j\|_{\ell^p(\mathbb{Z}^d)} = \sum_{|n|=N} \|\nabla^n (v, M^j \cdot - k)\|_{\ell^p(\mathbb{Z}^d)}
\]

\[
= \sum_{|n|=N} \|\nabla^n_{M^j \cdot - k} v\|_{\ell^p(\text{Supp}(\tilde{\varphi}(\cdot - k)))}_{\ell^p(\mathbb{Z}^d)}
\]

\[
\lesssim m^{j/p} \sum_{|n|=N} \|\nabla^n_{M^j \cdot - k} v\|_{L^p(\mathbb{R}^d)}
\]

\[
\lesssim m^{j/p} \omega_N(v, C_2 m^{-j/d}_{L^p}),
\]

18
since $M$ is isotropic. Furthermore, for any integer $C > 0$ and any $t > 0$, $\tilde{\omega}_N(v, Ct)_{L^p} \leq C \tilde{\omega}_N(v, t)_{L^p}$. Thus,

$$\sum_{|n|=N} \| \nabla^n v^j \|_{L^p(\mathbb{Z}^d)} \lesssim m^{j/p} \tilde{\omega}_N(v, m^{-j/d})_{L^p},$$

which implies (20). \hfill \Box

5.2. Inverse Theorem

We consider the sequence $(v^0, (d^j)_{j \geq 0})$ and we study the convergence of the reconstruction process:

$$v^j = S(v^{j-1})v^{j-1} + Ed^{j-1},$$

throughout the study of the limit of the sequence of functions

$$v_j(x) = \sum_{k \in \mathbb{Z}^d} v^j_k \varphi(M^j x - k), \quad \text{(21)}$$

where $\varphi$ is defined in (1). More precisely, we show that under certain conditions on the sequence $(v^0, (d^j)_{j \geq 0})$ and on $\varphi$, $v_j$ converges to some function $v$ belonging to a Besov space.

For that purpose, we establish that if the quasi-linear prediction rule $S(w)$ reproduces polynomials up to total degree $N - 1$ then all the mixed finite differences of order lesser than $N$ can be defined using local and bounded difference operators:

**Proposition 5.1.** Let $S(w)$ be a quasi-linear prediction rule reproducing polynomials up to total degree $N - 1$. Then for any $l$, $0 < l \leq N$ there exists a local bounded difference operator $S_l$ such that:

$$\Delta^l S(w)u := S_l(w)\Delta^l u,$$

for all $u, w \in \ell^\infty(\mathbb{Z}^d)$. 19
The proof is available in [15], Proposition 1. In contrast to the tensor product case studied in [14], the operator \( S_l(w) \) is multi-dimensional and is defined from \((\ell^\infty(\mathbb{Z}^d))^q\) onto \((\ell^\infty(\mathbb{Z}^d))^q\), \( q_l = \#\{\mu, |\mu| = l\} \), and cannot be reduced to a set of difference operators in some given directions.

The inverse theorem proved in this section is based on the contractivity of the difference operators. This is done by studying the joint spectral radius, which we now define:

**Definition 5.1.** Let us consider a set of local and bounded difference operators \((S_l)_{l\geq 0}\), defined in Proposition 5.1 with \( S_0 := S \). The joint spectral radius in \((\ell^p(\mathbb{Z}^d))^q\) of \( S_l \) is given by

\[
\rho_p(S_l) := \inf_{j \geq 0} \sup_{(w^0, \cdots, w^{j-1}) \in (\ell^p(\mathbb{Z}^d))^q} \|S_l(w^{j-1}) \cdots S_l(w^0)\|^{1/j}_{(\ell^p(\mathbb{Z}^d))^q \rightarrow (\ell^p(\mathbb{Z}^d))^q} = \inf\{\rho, \|S_l(w^{j-1}) \cdots S_l(w^0)\| \leq \rho^j \|\Delta^1 v\|_{(\ell^p(\mathbb{Z}^d))^q}, \forall v \in \ell^p(\mathbb{Z}^d)\}.
\]

**Remark 5.1.** When \( v^j = S(v^{j-1})v^{j-1} \), for all \( j > 0 \) we may write:

\[
\Delta^1 S(v^j)v^j = S_l(S(v^{j-1})v^{j-1})\Delta^1 v^j = S_l(S(v^{j-1})v^{j-1})S_l(v^{j-1})\Delta^1 v^{j-1} = \cdots = (S_l)^j v^0.
\]

This naturally leads to another definition of the joint spectral radius by putting \( w^j = S_l v^0 \) in the above definition. In [16], the following definition was introduced to study the convergence and stability of one-dimensional power-P scheme: let us consider a set of local and bounded difference operators \((S_l)_{l\geq 0}\), defined as in Proposition 5.1 where \( S_0 := S \). The joint spectral radius in \((\ell^p(\mathbb{Z}^d))^q\) of \( S_l \) is given by

\[
\tilde{\rho}_p(S_l) := \inf_{j \geq 0} \|(S_l)^j\|_{(\ell^p(\mathbb{Z}^d))^q \rightarrow (\ell^p(\mathbb{Z}^d))^q}^{1/j} = \inf\{\rho, \|\Delta^1 S^j v\| \leq \rho^j \|\Delta^1 v\|_{(\ell^p(\mathbb{Z}^d))^q}, \forall v \in \ell^p(\mathbb{Z}^d)\}.
\]
Since our prediction operator is quasi-linear, the definition (5.1) is more appropriate.

Before we prove the inverse theorem, we need to establish some extensions to the non-separable case of results obtained in [14]:

**Lemma 5.2.** Let $S(w)$ be a quasi-linear prediction rule that exactly reproduces polynomials of total degree $N - 1$. Then,

$$
\|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p} \left( \|\Delta^N v^j\|_{(L^p(\mathbb{Z}^d))^d} + \|d^j\|_{L^p(\mathbb{Z}^d)} \right). \quad (22)
$$

Moreover, if $\rho > \rho_0(S_N)$,

$$
\|\Delta^N v^j\|_{(L^p(\mathbb{Z}^d))^d} \lesssim \rho^j \left( \|v^0\|_{L^p(\mathbb{Z}^d)} + \sum_{l=0}^{j-1} \rho^{-l} \|d^j\|_{L^p(\mathbb{Z}^d)} \right). \quad (23)
$$

**Proof:** Using the definition of functions $v_j(x)$ and the scaling equation (1), we get

$$
v_{j+1}(x) - v_j(x) = \sum_k v_{j+1}^k \varphi(M^{j+1}x - k) - \sum_k v_j^k \varphi(M^jx - k)
$$

$$
= \sum_k \left( ((S(v^j)v^j)_k + d_k^j) \varphi(M^{j+1}x - k) - \sum_k v_j^k \sum_l g_{k-M^j} \varphi(M^{j+1}x - l) \right)
$$

$$
= \sum_k \left( ((S(v^j)v^j)_k - \sum_l g_{k-M^j}v_l^j) \varphi(M^{j+1}x - k) + \sum_k d_k^j \varphi(M^{j+1}x - k) \right).
$$

If $S$ exactly reproduces polynomials up to total degree $N - 1$ then having in mind that $\sum_l g_{k-M^j}v_l^j$ is a linear prediction that exactly reproduces polynomials up to total degree $N - 1$ and using the same arguments as in Lemma 5.1, we get

$$
\| \sum_k \left( ((S(v^j)v^j)_k - \sum_l g_{k-M^j}v_l^j) \varphi(M^{j+1}x - k) \right) \|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p} \|\Delta^N v^j\|_{(L^p(\mathbb{Z}^d))^d}. \quad (21)
$$
The proof of (22) is thus complete. To prove (23), we note that:

\[
\| \Delta^N v^j \|_{(p(Z^d))^N} \lesssim \| S_N(v^{j-1}) \Delta^N v^{j-1} \|_{(p(Z^d))^N} + \| \Delta^N d^{j-1} \|_{(p(Z^d))^N} \\
\lesssim \rho \| \Delta^N v^{j-1} \|_{(p(Z^d))^N} + \| \Delta^N d^{j-1} \|_{(p(Z^d))^N} \\
\lesssim \rho^j \left( \| v^0 \|_{(p(Z^d))} + \sum_{l=0}^{j-1} \rho^{-l} \| d^l \|_{(p(Z^d))} \right).
\]

By abusing a little bit terminology, we say that \( \varphi \) exactly reproduces polynomials if the underlying subdivision scheme does. With this in mind, we are ready to state the inverse theorems: the first one deals with \( L^p \) convergence under the main hypothesis \( \rho_p(S_1) < m^{1/p} \), while the second deals with the convergence in \( B^s_{p,q}(\mathbb{R}^d) \) under the main hypothesis \( \rho_p(S_N) < m^{1/p-s/d} \) for some \( N > 1 \) and \( N - 1 \leq s < N \).

**Theorem 5.2.** Let \( S(w) \) be a quasi-linear prediction rule reproducing the constants. Assume that \( \rho_p(S_1) < m^{1/p} \), if

\[
\| v^0 \|_{(p(Z^d))} + \sum_{j \geq 0} m^{-j/p} \| d^j \|_{(p(Z^d))} < \infty,
\]

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

\[
\| v \|_{L^p(\mathbb{R}^d)} \leq \| v^0 \|_{(p(Z^d))} + \sum_{j \geq 0} m^{-j/p} \| d^j \|_{(p(Z^d))}.
\]

**Proof:** From estimates (22) and (23) for \( N = 1 \) one has, in particular

\[
\| v_{j+1} - v_j \|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p} \rho^j \left( \| v^0 \|_{(p(Z^d))} + \sum_{l=0}^{j} \rho^{-l} \| d^l \|_{(p(Z^d))} \right).
\]
for $\rho > \rho_p(S_1)$. We then choose $\rho$ such that $\rho_p(S_1) < \rho < m^{1/p}$ to obtain:

$$\|v\|_{L^p(\mathbb{R}^d)} \leq \|v_0\|_{L^p(\mathbb{R}^d)} + \sum_{j \geq 0} \|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)}$$

$$\lesssim \left(25\right) \|v_0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j \geq 0} \|v_j\|_{\ell^p(\mathbb{Z}^d)} \rho_p(S_1)^j + \sum_{l \geq 0} \|d_l\|_{\ell^p(\mathbb{Z}^d)} \rho^{-l} \sum_{j \geq l} (m^{1/p} \rho)^j$$

$$\lesssim \|v_0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l \geq 0} \|d_l\|_{\ell^p(\mathbb{Z}^d)} m^{-l/p}.$$  

This proves (24).

Now we extend this result to the case of Besov spaces:

**Theorem 5.3.** Let $S(w)$ be a quasi-linear prediction rule exactly reproducing polynomials up to total degree $N - 1$ and let $\varphi$ exactly reproduce polynomials up to degree $N - 1$. Assume that $\rho_p(S_N) < m^{1/p-s/d}$ for some $N > s \geq N - 1$. If $(v_0, d^0, d^1, \ldots)$ are such that

$$\|v_0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)} j) \|((d^j_k)_{k \in \mathbb{Z}^d})_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} < \infty,$$

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

$$\|v\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v_0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)} j) \|((d^j_k)_{k \in \mathbb{Z}^d})_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} < \infty.$$  

(26)

**Proof:** First, by Hölder inequality for any $q, q' > 0, \frac{1}{q} + \frac{1}{q'} = 1$, it holds that

$$\sum_{l \geq 0} \|d_l\|_{\ell^{p}(\mathbb{Z}^d)} m^{-l/p} \leq \|(m^{(s/d-1/p)} j) \|((d^j_k)_{k \in \mathbb{Z}^d})_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} \|((m^{-j s/d})_j)_{j \geq 0}\|_{\ell^{q'}(\mathbb{R}^d)}$$

$$\lesssim \|(m^{(s/d-1/p)} j) \|((d^j_k)_{k \in \mathbb{Z}^d})_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} \|((m^{-j s/d})_j)_{j \geq 0}\|_{\ell^{q'}(\mathbb{R}^d)},$$

23
and finally,

$$\|v\|_{L^p(\mathbb{R}^d)} \lesssim \|v_0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d - 1/p)j})(d_k^j)_{k \in \mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \|e_j\|_{\ell^p(\mathbb{Z}^d)}.$$

It remains to evaluate the semi-norm $\|v\|_{B_{p,q}(\mathbb{R}^d)} := \|(m^{js/d}N(v, m^{-j/d})L^p)_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)}$.

For each $j \geq 0$, we have

$$\hat{\omega}_N(v, m^{-j/d})_{L^p} \leq \hat{\omega}_N(v - v_j, m^{-j/d})_{L^p} + \hat{\omega}_N(v_j, m^{-j/d})_{L^p}. \quad (27)$$

For the first term on the right hand side of (27), one has using Lemma 5.2

$$\hat{\omega}_N(v - v_j, m^{-j/d})_{L^p} \lesssim \sum_{l \geq j} \|v_{l+1} - v_l\|_{L^p(\mathbb{R}^d)}$$

$$\lesssim \sum_{l \geq j} m^{-l/p} (\|\Delta^N v_l\|_{\ell^p(\mathbb{Z}^d)} + \|d^l\|_{\ell^p(\mathbb{Z}^d)})$$

$$\lesssim \sum_{l \geq j} m^{-l/p} \left( \rho^l \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k=0}^l \rho^{l-k} \|d^k\|_{\ell^p(\mathbb{Z}^d)} \right).$$

For the first term, choosing $\rho_p(S_N) < \rho < m^{1/p}$, we have

$$\sum_{l \geq j} m^{-l/p} \rho^l \|v^0\|_{\ell^p(\mathbb{Z}^d)} \sim m^{-j/p} \rho^j \|v^0\|_{\ell^p(\mathbb{Z}^d)}.$$
For the second term, we get
\[
\sum_{l \geq j} m^{-l/p} \sum_{k=0}^l \rho^{l-k} \|d^k\|_{L^p(\mathbb{Z}^d)}
\]
\[
= m^{-j/p} \sum_{k=0}^j \rho^{j-k} \|d^k\|_{L^p(\mathbb{Z}^d)} + \sum_{l > j} m^{-l/p} \sum_{k=0}^l \rho^{l-k} \|d^k\|_{L^p(\mathbb{Z}^d)}
\]
\[
\lesssim m^{-j/p} \sum_{k=0}^j \rho^{j-k} \|d^k\|_{L^p(\mathbb{Z}^d)} + \sum_{k \geq 0} \sum_{l \geq \max(k, j)} m^{-l/p} \rho^{j-k} \|d^k\|_{L^p(\mathbb{Z}^d)}
\]
\[
\lesssim m^{-j/p} \sum_{k=0}^j \rho^{j-k} \|d^k\|_{L^p(\mathbb{Z}^d)} + \sum_{k \geq 0} \|d^k\|_{L^p(\mathbb{Z}^d)} \rho^{-k} \sum_{l \geq \max(k, j)} \rho^{l-m^{-l/p}}
\]
\[
\lesssim m^{-j/p} \sum_{k=0}^j \rho^{j-k} \|d^k\|_{L^p(\mathbb{Z}^d)} + \sum_{k > j} m^{-k/p} \|d^k\|_{L^p(\mathbb{Z}^d)}.
\]

Similarly, for the second term on the right hand side of (27), one has
\[
\tilde{\omega}_N(v_j, m^{-j/d})_{L^p} \lesssim \|v_j\|_{L^p(\mathbb{R}^d)} \lesssim \|v\|_{L^p(\mathbb{R}^d)}
\]

The estimate of the semi-norm $|v|_{B^*_{p,q}}$ is then reduced to the estimates of $\|(m^{j/d}a_j)_{j \geq 0}\|_{L^p(\mathbb{Z}^d)}$, $\|(m^{j/d}b_j)_{j \geq 0}\|_{L^p(\mathbb{Z}^d)}$ and $\|(m^{j/d}c_j)_{j \geq 0}\|_{L^p(\mathbb{Z}^d)}$, with
\[
a_j := m^{-j/p} \rho^j \|v^0\|_{L^p(\mathbb{Z}^d)},
\]
\[
b_j := m^{-j/p} \rho^j \sum_{l=0}^j \rho^{-l} \|d^l\|_{L^p(\mathbb{Z}^d)},
\]
\[
c_j := \sum_{l > j} m^{-l/p} \|d^l\|_{L^p(\mathbb{Z}^d)}.
\]

The main point here is that $\rho_p(S_N)(m^{s/d-1/p}) < 1$, thus, we can choose $\rho > \rho_p(S_N)$ such that
\[
\alpha = m^{s/d-1/p} \rho < 1.
\]
Therefore
\[ \| (\sigma^j a_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} = \| v^0 \|_{\ell^q(\mathbb{Z}^d)} \| (\alpha_j^j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \lesssim \| v^0 \|_{\ell^q(\mathbb{Z}^d)}. \] (28)

In order to estimate \( \| (m^{js/d} b_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \), we rewrite it in the following form:
\[
\| (m^{js/d} b_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} = \| (m^{(s/d-1/p)} \rho^j \sum_{l=0}^j \rho^{-l} \| d^l \|_{\ell^q(\mathbb{Z}^d)})_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)}.
\]

We now make use of the following discrete Hardy inequality: if \( 0 < \alpha < 1 \), then
\[
\| (\alpha^j \sum_{l=0}^j \alpha^{-l} x_l)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \lesssim \| (x_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)}.
\]
Applying it to \( x_l = m^{(s/d-1/p)} \| d^l \|_{\ell^q(\mathbb{Z}^d)} \) yields
\[
\| (m^{js/d} b_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \lesssim \| (m^{(s/d-1/p)} \| (d^l)_{k \in \mathbb{Z}^d} \|_{\ell^q(\mathbb{Z}^d)})_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)}. \] (29)

To estimate \( \| (m^{js/d} c_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \), we rewrite it as follows
\[
\| (m^{js/d} c_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} = \| m^{js/d} \sum_{l \geq j} m^{-ls/d} (m^{(s/d-1/p)} \| (d^l)_{k \in \mathbb{Z}^d} \|_{\ell^q(\mathbb{Z}^d)}) \|_{\ell^q(\mathbb{Z}^d)}
\]
and make use of another discrete Hardy inequality: if \( \beta > 1 \), then
\[
\| (\beta^j \sum_{l \geq j} \beta^{-l} y_l)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \lesssim \| (y_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)}.
\]
Taking \( y_l = m^{(s/d-1/p)} \| d^l \|_{\ell^q(\mathbb{Z}^d)} \), we obtain, since \( s > N - 1 \),
\[
\| (m^{js/d} c_j)_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)} \lesssim \| (m^{js/d-1/p}) \|_{\ell^q(\mathbb{Z}^d)})_{j \geq 0} \|_{\ell^q(\mathbb{Z}^d)}. \] (30)

Then (26) follows by combining (28), (29) and (30). \( \square \)
6. Stability of Nonlinear Subdivision Schemes

In this section, we focus on the stability of the iteration \( v^j = S(v^{j-1})v^{j-1} \) in Sobolev spaces \( W^p_N(\mathbb{R}^d) \). We forget the detail term and we obtain a subdivision scheme. We recall that the elements of \( W^p_N(\mathbb{R}^d) \) are those functions of \( L^p(\mathbb{R}^d) \) having their differential up to order \( N \) in \( L^p(\mathbb{R}^d) \), the norm on \( W^p_N(\mathbb{R}^d) \) being defined as follows:

\[
\|v\|_{W^p_N(\mathbb{R}^d)} = \|v\|_{L^p} + \sum_{|\mu| \leq N} \|D^\mu v\|_{L^p(\mathbb{R}^d)}.
\]

More precisely, we use the following definition for the stability of subdivision scheme in Sobolev spaces:

**Definition 6.1.** Let \( S(w) \) be a quasi-linear prediction rule, the subdivision scheme \( v^j = S(v^{j-1})v^{j-1} \) is said to be stable in \( W^p_N(\mathbb{R}^d) \) if for all \( v^0 \) and \( \tilde{v}^0 \) in \( \ell^\infty(\mathbb{Z}^d) \), we have:

\[
\|D^\mu v_j - D^\mu \tilde{v}_j\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} \quad \forall |\mu| \leq N
\]

where \( v_j = \sum_{k \in \mathbb{Z}^d} v^j_k \varphi(M^j x - k) \), \( \varphi \) belonging to \( W^p_N(\mathbb{R}^d) \).

Before proving the theorem on stability, we need to show the following lemma:

**Lemma 6.1.** Assume that \( S(w) \) is a quasi-linear prediction rule exactly reproducing polynomials up to total degree \( N \), that \( \varphi \) exactly reproduces polynomials up to total degree \( N \) and that \( M \) is isotropic then

\[
\|v^j - \tilde{v}^j\|_{\ell^p(\mathbb{Z}^d)} \leq m^{1/p - N/d} \|v^{j-1} - \tilde{v}^{j-1}\|_{\ell^p(\mathbb{Z}^d)} + C \|\Delta^{N+1}(v^{j-1} - \tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^{N+1}}
\]

(31)
Proof: Since \( \varphi \) exactly reproduces polynomials up to total degree \( N \), it defines a subdivision scheme \( \tilde{S} \) that exactly reproduces polynomials up to total degree \( N \). Then, let us define for any \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \)

\[
r_i(x) = \sum_{l=1}^{d} \lambda_{i,l} x_l, \quad i = 1, \ldots, d,
\]

where the matrix \( \Lambda = (\lambda_{i,l})_{i,l=1}^{d} \) is the same as in (17). For a multi-index \( \mu = (\mu_1, \ldots, \mu_d) \in \mathbb{N}^d \) let

\[
r_\mu(x) := r_{\mu_1}(x) \cdot \ldots \cdot r_{\mu_d}(x)
\]

and let us consider the differential operator

\[
r_\mu(D) := r_{\mu_1}(D) \cdot \ldots \cdot r_{\mu_d}(D), \quad \text{where } r_i(D) = \sum_{l=1}^{d} \lambda_{i,l} D_{e_l}, \quad i = 1, \ldots, d.
\]

Since \( \Lambda \) is invertible, the set \( \{r_\mu : |\mu| = N\} \) forms a basis of the space of all polynomials of exact degree \( N \), which proves that

\[
\sum_{|\mu| = N} \|D_{e_1} \ldots D_{e_d}(v - v_j)\|_{L^p(\mathbb{R}^d)} \sim \sum_{|\mu| = N} \sum_{l \geq j} \|r_\mu(D)(v_{l+1} - v_l)\|_{L^p(\mathbb{R}^d)}.
\]

Now, since \( M \) is isotropic, then \( r_\mu(D)(f(M^l x)) = \sigma^{l \mu/d}(r_\mu(D) f)(M^l x) \) ([10]), where \( \sigma^\mu = \prod_{i=1}^{d} \sigma_i^{\mu_i} \). With this in mind, we may write:

\[
\|r_\mu(D)(v_j - v_{j-1})\|_{L^p(\mathbb{R}^d)} \sim m^{l(-1/p+N/d)} \|v_j - \tilde{S}v_{j-1}\|_{L^p(\mathbb{Z}^d)}.
\]

Using the fact that \( \tilde{S} \) exactly reproduces polynomials up to total degree \( N \), we deduce

\[
\|r_\mu(D)(v_j - \tilde{v}_j)\|_{L^p(\mathbb{R}^d)} \leq \|r_\mu(D)(v_{j-1} - \tilde{v}_{j-1})\|_{L^p(\mathbb{R}^d)} + C m^{l(-1/p+N/d)} \|\Delta^{N+1}(v_j - \tilde{v}_j)\|_{(L^p(\mathbb{Z}^d))^{N+1}}
\]

28
and we finally get:

$$\|v^j - \bar{v}^j\|_{L^p(\mathbb{Z}^d)} \leq m^{1/p-N/d} \|v^{j-1} - \bar{v}^{j-1}\|_{L^p(\mathbb{Z}^d)} + C \|\Delta^{N+1}(v^j - \bar{v}^j)\|_{L^p(\mathbb{Z}^d)^{\otimes N+1}}$$

We now state the theorem on the stability of the subdivision scheme in $W^p_N(\mathbb{R}^d)$.

**Theorem 6.1.** Let $S(w)$ be a quasi-linear prediction rule that exactly reproduces polynomials up to total degree $N$. Assume that $S_{N+1}$ satisfies $\rho_p(S_{N+1}) < m^{1/p-N/d}$, that $M$ is isotropic, that

$$\|\Delta^{N+1}(v^j - \bar{v}^j)\|_{L^p(\mathbb{Z}^d)^{\otimes N+1}} < \rho \|v^0 - \bar{v}^0\|_{L^p(\mathbb{Z}^d)}$$

for some $\rho_p(S_{N+1}) < \rho < m^{1/p-N/d}$, and also that $\varphi$ is in $W^p_N(\mathbb{R}^d)$ and exactly reproduces polynomial up to total degree $N$ then the subdivision scheme is stable in $W^p_N(\mathbb{R}^d)$.

**Proof:** First, we remark that:

$$\sum_{|\mu| = N} \|D^\mu(v_j - \bar{v}_j)\|_{L^p(\mathbb{R}^d)} \sim \sum_{|\mu| = N} \|r_\mu(D)(v_j - \bar{v}_j)\|_{L^p(\mathbb{R}^d)}.$$

Since the matrix $M$ is isotropic, we may write

$$\|r_\mu(D)(v_j - \bar{v}_j)\|_{L^p(\mathbb{R}^d)} \sim m^{j(-1/p+N/d)} \|v^j - \bar{v}^j\|_{L^p(\mathbb{Z}^d)}.$$

Let $\rho$ be such that $\rho_p(S_{N+1}) < \rho < m^{1/p-N/d}$. We define the sequences $\alpha^j := m^{j(-1/p+N/d)} \|v^j - \bar{v}^j\|_{L^p(\mathbb{Z}^d)}$ and $\beta^j = m^{j(-1/p+N/d)} \|\Delta^{N+1}(v^j - \bar{v}^j)\|_{L^p(\mathbb{Z}^d)^{\otimes N+1}}$.

By (32) and Lemma 6.1, these sequences satisfy the following inequalities:

$$\alpha^j \leq \alpha^{j-1} + D\beta^{j-1}$$

$$\beta^{j-1} \leq C(\rho_m^{(-1/p+N/d)})j^{-1}\alpha^0,$$
the constant $C$ being independent of $j$. From the above inequality we deduce the following estimation:

$$m^{j(-1/p+N/d)}\|v^j - \tilde{v}^j\|_{\ell^p(Z^d)} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(Z^d)}$$

7. Stability of the Multi-Scale Representation

Here, we consider two data sets $(v^0, d^0, d^1, \ldots)$ and $(\tilde{v}^0, \tilde{d}^0, \tilde{d}^1, \ldots)$ corresponding to two reconstruction processes

$$v^j = S(v^{j-1})v^{j-1} + e^j = S(v^{j-1})v^{j-1} + Ed^{j-1}$$

and

$$\tilde{v}^j = S(\tilde{v}^{j-1})\tilde{v}^{j-1} + \tilde{e}^j = S(\tilde{v}^{j-1})\tilde{v}^{j-1} + E\tilde{d}^{j-1}.$$ (34)

where $E$ is the matrix corresponding to the basis of the kernel of the projection operator. In that context, $v$ is the limit of $v_j(x) = \sum_{k \in Z^d} v^j_k \varphi(M^j x - k)$ (and similarly for $\tilde{v}$).

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

**Theorem 7.1.** Let $S(w)$ be a quasi-linear prediction rule that reproduces the constants and suppose that $\rho_p(S_1) < m^{\frac{1}{p}}$. Assume that $v_j$ and $\tilde{v}_j$ converge to $v$ and $\tilde{v}$ in $L^p(\mathbb{R}^d)$ respectively and also that

$$\Delta(v^j - \tilde{v}^j)_{\ell^p(Z^d)^d} \lesssim \rho^j \left( \|v^0 - \tilde{v}^0\|_{\ell^p(Z^d)} + \sum_{l=0}^{j-1} \rho^{-l}\|d^l - \tilde{d}^l\|_{\ell^p(Z^d)} \right).$$ (35)
holds for some $\rho_p(S_1) < \rho < m^{1/p}$, we have:

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

(36)

**Proof:** Using the same technique as in the proof of Lemma 6.1 (but without assuming the isotropy of the matrix $M$), one has:

$$
\|v_j - \tilde{v}_j\|_{L^p(\mathbb{R}^d)} \leq \|v_{j-1} - \tilde{v}_{j-1}\|_{L^p(\mathbb{R}^d)} + D \left( m^{-j/p} \Delta^1(v^j - \tilde{v}^j)\|_{L^p(\mathbb{Z}^d)} + m^{-j/p} \|d^{j-1} - \tilde{d}^{j-1}\|_{L^p(\mathbb{Z}^d)} \right).
$$

From the hypothesis (35), we get:

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

Using the fact that $\rho m^{-1/p} < 1$ we immediately get:

$$
\|v_j - \tilde{v}_j\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{L^p(\mathbb{Z}^d)} + \sum_{l = 0}^{j-1} \rho^{-l} \|d^l - \tilde{d}^l\|_{L^p(\mathbb{Z}^d)},
$$

from which we obtain when $j \to +\infty$:

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

In view of the inverse inequality (26), it seems natural to seek an inequality of type

$$
\|v - \tilde{v}\|_{B_{p,q}^s(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{L^p(\mathbb{Z}^d)} + \sum_{j \geq 0} \rho^j \|d^j - \tilde{d}^j\|_{L^p(\mathbb{Z}^d)}.
$$

(37)

We now prove a stability theorem in Besov space $B_{p,q}^s(\mathbb{R}^d)$ in the following theorem:
Theorem 7.2. Let \( S(w) \) be a quasi-linear prediction rule which exactly reproduces polynomials up to degree \( N - 1 \), that \( S \) is such that \( \rho_p(S_N) < m^{1/p-s/d} \) for some \( s > N - 1 \). Assume that \( v_j \) and \( \tilde{v}_j \) converge to \( v \) and \( \tilde{v} \) in \( B^*_{p,q}(\mathbb{R}^d) \) respectively and that

\[
\|\Delta^N(v^j - \tilde{v}^j)\|_{(\ell^p(\mathbb{Z}^d))^{\otimes N}} \lesssim \rho^j \left(\|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=0}^{j-1} \rho^{-l} \|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}\right)
\]

(38)

for some \( \rho_p(S_N) < \rho < m^{1/p-s/d} \). Then the function \( v - \tilde{v} \) belongs to \( B^*_{p,q}(\mathbb{R}^d) \). Moreover, we obtain for that \( \rho \)

\[
\|v - \tilde{v}\|_{B^*_{p,q}(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \| (\rho^{-j}) \|(d^l_k - \tilde{d}^l_k)_{k \in \mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} j \geq 0 \|_{\ell^q(\mathbb{Z}^d)}. \]

(39)

Remark 7.1. With the hypothesis (38), which is a natural extension of (23) for the stability issue, we have no specific condition on \( s, p, q \) which was not the case with the tensor product method proposed in [14]. Furthermore, we shall note that a similar study was proposed in the one-dimensional case to study the stability of so-called \( r \)-shift invariant subdivision operators [16] with a slightly lighter hypothesis than (38). However, in [16] the stability is obtained only when \( \rho_p(S_N) < 1 \), while with our approach the condition on the joint spectral radius is the same both for the stability and the smoothness.

Proof: Using the same technique as in Theorem 7.1, but using hypothesis (38), we immediately get:

\[
\|v - \tilde{v}\|_{L^p} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l \geq 0} \rho^{-l} \|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)},
\]

Since by the constraints on \( s \) we can put \( \rho = m^{1/p+(s-\varepsilon)/d} \) for some positive
This enables us to write
\[
\|v - \tilde{v}\|_{L^p(\mathbb{R}^d)} \leq \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j})\|_{\ell^p(\mathbb{Z}^d)} \|d^j - \tilde{d}^j\|_{\ell^p(\mathbb{Z}^d)} \leq 0 \|\ell^q(\mathbb{Z}^d)\|_{\ell^q(\mathbb{Z}^d)}.
\]

It remains to estimate the semi-norm
\[
|w|_{B^p_{r,q}(\mathbb{R}^d)} := \|(m^{s/d})\omega_N(w, m^{-j/d})\|_{L^p(\mathbb{R}^d)} \leq 0 \|\ell^q(\mathbb{Z}^d)\|_{\ell^q(\mathbb{Z}^d)},
\]
for \(w := v - \tilde{v}\). For every \(j \geq 0\), denoting \(w_j = v_j - \tilde{v}_j\), we have
\[
\omega_N(w, m^{-j/d})_{L^p} \leq \omega_N(w - w_j, m^{-j/d})_{L^p} + \omega_N(w_j, m^{-j/d})_{L^p}. \tag{40}
\]

For the first term, using successively Lemma (5.2) and hypothesis (38), one has
\[
\omega_N(w - w_j, m^{-j/d})_{L^p} \lesssim \sum_{l \geq j} \|w_{l+1} - w_l\|_{L^p(\mathbb{R}^d)} \lesssim \sum_{l \geq j} m^{-l/p} \left(\|\Delta_N(v^l - \tilde{v}^l)\|_{\ell^p(\mathbb{Z}^d)} + \|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}\right) \lesssim \sum_{l \geq j} m^{-l/p} \left(\rho^l \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k=0}^{l} \rho^{l-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)}\right).
\]

Since \(\rho < m^{1/p}\), then for the first term we have
\[
\sum_{l \geq j} m^{-l/p} \rho^l \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} \sim m^{-j/p} \rho^j \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}.
\]
For the second term, we get
\[ \sum_{l \geq j} m^{-l/p} \sum_{k=0}^{l} \rho^{l-k} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)} = \]
\[ = m^{-j/p} \sum_{k=0}^{j} \rho^{j-k} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)} + \sum_{l > j} m^{-l/p} \sum_{k=0}^{l} \rho^{l-k} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)} \]
\[ \lesssim m^{-j/p} \sum_{k=0}^{j} \rho^{j-k} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)} + \sum_{l > j} \sum_{k=0}^{\max(l, j+1)} m^{-l/p} \rho^{l-k} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)} \]
\[ \lesssim m^{-j/p} \sum_{k=0}^{j} \rho^{j-k} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)} + \sum_{k > j} m^{-k/p} \| d^k - \tilde{d}^k \|_{\ell^p(\mathbb{Z}^d)}. \]

The second term in (40) is evaluated as follows:
\[ \omega_N(v_j - \tilde{v}_j, m^{-j/d})_{L^p} \lesssim \| v_j - \tilde{v}_j \|_{L^p(\mathbb{R}^d)} \]
\[ \lesssim m^{-j/p} \rho^j \| v^0 - \tilde{v}^0 \|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=0}^{j} \rho^{j-l} \| d^l - \tilde{d}^l \|_{\ell^p(\mathbb{Z}^d)}. \]

The second term on the right hand side of (40), can be evaluated the same way. We have thus reduced the estimate of \( |w|_{B^{r/p}_{p,q}(\mathbb{R}^d)} \), to the estimates of the discrete norms \( \|(m^{j/d}a_j)_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} \), and \( \|(m^{j/d}b_j)_{j \geq 0}\|_{\ell^p(\mathbb{Z}^d)} \), where the sequences are defined as follows:
\[ a_j := \rho^j m^{-j/p} \| v^0 - \tilde{v}^0 \|_{\ell^p(\mathbb{Z}^d)}, \]
\[ b_j := m^{-j/p} \sum_{l=0}^{j} \rho^{j-l} \| d^l - \tilde{d}^l \|_{\ell^p(\mathbb{Z}^d)}, \]
\[ c_j := \sum_{l > j} m^{-l/p} \| d^l - \tilde{d}^l \|_{\ell^p(\mathbb{Z}^d)}. \]

Note that this quantities are identical to that obtained in the convergence theorem replacing \( v^l \) by \( v^l - \tilde{v}^l \) and \( d^l \) by \( d^l - \tilde{d}^l \), so that the end of the proof is identical. \( \square \)
8. Illustration: Bi-Dimensional Nonlinear Affine Prediction Using the Hexagonal Dilation Matrix

We now focus on the construction of a converging nonlinear multi-scale representations using as dilation matrix the hexagonal matrix

\[ M = \begin{pmatrix} 2 & 1 \\ 0 & -2 \end{pmatrix}. \]

In that context, we make up a partition of \( \mathbb{Z}^2 \) using the coset vectors of the matrix \( M \). For the matrix \( M \), with \( m = |\text{det}M| \), there are \( m \) coset vectors which are the following for the hexagonal matrix: \( \varepsilon_0 = (0,0)^T, \varepsilon_1 = (1,0)^T, \varepsilon_2 = (1,-1)^T, \varepsilon_3 = (2,-1)^T \). These coset vectors satisfy \( \mathbb{Z}^2 = \bigcup_{i=0}^{3} \{ Mk + \varepsilon_i, \ k \in \mathbb{Z}^2 \} \). We now build the prediction \( \hat{v}^j \) for the different coset points \( Mk + \varepsilon_i \) using an affine interpolant defined on the coarse grid \( \Gamma^{j-1} = \{ M^{-j+1}k, \ k \in \mathbb{Z}^2 \} \) corresponding to the location of the values \( (v^j_k)_{k \in \mathbb{Z}^2} \) (see (3)). To build the affine interpolant, we use one of the following four different stencils on the grid \( \Gamma^{j-1} \):

\[ V^{j,1}_k = M^{-j+1}\{k, k + e_1, k + e_2\}, \]
\[ V^{j,2}_k = M^{-j+1}\{k, k + e_2, k + e_1 + e_2\}, \]
\[ W^{j,1}_k = M^{-j+1}\{k + e_1, k + e_2, k + e_1 + e_2\}, \]
\[ W^{j,2}_k = M^{-j+1}\{k, k + e_1, k + e_1 + e_2\}. \]

Each approximated value \( v^j_k \) being associated to the location \( M^{-j}k \), we determine the stencil this point belongs to, and we then define the prediction as the value of the affine interpolant at \( M^{-j}k \) using the selected stencil. In that context, the prediction rules at \( Mk \) and \( Mk + \varepsilon_1 \) are independent of the
choice of the stencil, and we always have:

\[
\hat{v}_M^j = v_k^{j-1} \quad \text{and} \quad \hat{v}_M^{j+\varepsilon_1} = \frac{1}{2} v_k^{j-1} + \frac{1}{2} v_{k+\varepsilon_1}^{j-1}. \tag{41}
\]

For the prediction at \(Mk + \varepsilon_2\), we use stencils \(V_{k}^{j,1}\) or \(V_{k}^{j,2}\) leading respectively to:

\[
\begin{align*}
\hat{v}_{Mk+\varepsilon_2}^{j,1} &= \frac{1}{4} v_{k+\varepsilon_1}^{j-1} + \frac{1}{2} v_{k+\varepsilon_2}^{j-1} + \frac{1}{4} v_k^{j-1} \\
\hat{v}_{Mk+\varepsilon_2}^{j,2} &= \frac{1}{2} v_k^{j-1} + \frac{1}{4} v_{k+\varepsilon_2}^{j-1} + \frac{1}{4} v_{k+\varepsilon_1}^{j-1}.
\end{align*} \tag{42}
\]

For the points \(Mk + \varepsilon_3\), we use \(W_{k}^{j,1}\) or \(W_{k}^{j,2}\) leading respectively to:

\[
\begin{align*}
\hat{v}_{Mk+\varepsilon_3}^{j,1} &= \frac{1}{4} v_{k+\varepsilon_1+\varepsilon_2}^{j-1} + \frac{1}{4} v_k^{j-1} + \frac{1}{2} v_{k+\varepsilon_1}^{j-1} \\
\hat{v}_{Mk+\varepsilon_3}^{j,2} &= \frac{1}{4} v_k^{j-1} + \frac{1}{4} v_{k+\varepsilon_1}^{j-1} + \frac{1}{2} v_{k+\varepsilon_1+\varepsilon_2}^{j-1}.
\end{align*} \tag{43}
\]

when the stencils \(W_{k}^{j,1}\) and \(W_{k}^{j,2}\) are used respectively.

This leads to four different linear prediction rules depending on the choice for the prediction operator for coset vector \(\varepsilon_2\) and \(\varepsilon_3\). The described quasi-linear prediction operator satisfies \(\rho_\infty(S_1) < 1\) since we have:

**Proposition 8.1.** *The prediction defined by (41), (42), (43) satisfies:

\[
\|\Delta^1 \hat{v}_{M,+\varepsilon_i}^j\|_{(l\infty(Z^2))^2} \leq \frac{3}{4} \|\Delta^1 v_{j-1}^i\|_{(l\infty(Z^2))^2}.
\]

The proof of the above proposition is carried out considering all the possibilities for the choice of stencil. This example shows how the tensor product approach developed in [14] can be generalized to a non-diagonal dilation matrix \(M\)."
9. Conclusion

In this paper, we have presented a new kind of nonlinear multi-scale representations based on the use of non-diagonal dilation matrices. We have shown that the non-linear scheme proposed by Harten naturally extends in that context and we have shown convergence and stability in Besov spaces. These results require the existence of difference operators associated with the prediction operator we use. The main difference between our approach and the tensor product approach consists in the fact these difference operators use mixed finite differences and cannot thus be reduced to one-dimensional difference operators. This specificity obliged us to consider a definition of Besov spaces using mixed finite differences. After we have shown these theoretical results, we have ended the paper by giving some applications when the underlying subdivision scheme is interpolatory. Future work should involve the designing of non-interpolatory multi-scale representations and also possible extensions of our method to more general prediction operators than quasi-linear ones.

References


