Orthonormal bases for α-modulation spaces

Nielsen, Morten

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Nielsen, M. (2008). Orthonormal bases for α-modulation spaces. Department of Mathematical Sciences, Aalborg University. Research Report Series No. R-2008-03

AALBORG UNIVERSITY

Orthonormal bases for α-modulation spaces

by

Morten Nielsen

R-2008-03 March 2008

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 99 40 80 80 ▪ Telefax: +45 98 15 81 29
URL: http://www.math.aau.dk

ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

ORTHONORMAL BASES FOR α-MODULATION SPACES

MORTEN NIELSEN

ABSTRACT. We construct an orthonormal basis for the family of bi-variate α-modulation spaces. The construction is based on local trigonometric bases, and the basis elements are closely related to so-called brushlets. As an application, we show that m-term nonlinear approximation with the system in an α-modulation space can be completely characterized.

1. INTRODUCTION

The α-modulation spaces $M^{s,\alpha}_{p,q}(\mathbb{R}^d)$, $\alpha \in [0,1]$, form a parameterized family of smoothness spaces defined on $\mathbb{R}^d$ that include the Besov and modulation spaces as special cases, corresponding to $\alpha = 1$ and $\alpha = 0$, respectively. The spaces are built from the same type of scheme arising from different segmentations of the frequency space. The α-parameter determines the nature of the segmentation. For example, the Besov spaces ($\alpha = 1$) correspond to a dyadic segmentation of the frequency space, while the modulation spaces ($\alpha = 0$) correspond to a uniform covering. The intermediate cases correspond to "polynomial type" segmentations of the frequency space.

The α-modulation spaces were introduced by Gröbner [17] , and it was pointed out by Feichtinger and Gröbner [12,13] that Besov and modulation spaces are special cases of an abstract construction, the so-called decomposition type Banach spaces. The coverings giving rise to α-modulation spaces have also been considered by Päivärinta and Somersalo in [25] as a tool to study pseudo-differential operators.

In this paper the main focus is on discrete representation of functions in α-modulation spaces relative to a basis. The connection between abstract notions of smoothness and properties of discrete representations is very important in applicable harmonic analysis. The guiding principle is often that smoothness should be characterized by (or at least imply) some decay or sparseness of an associated discrete expansion. A well-known example is provided by wavelets, for which it is known (see [23]) that suitable sparseness of a wavelet expansion is equivalent to smoothness measured in a Besov space. Sparse representations of functions are useful for e.g. compression purposes. For example, wavelets have been successfully applied to compress sound signals and images contained in a suitable Besov space, see [9,10].

The main contribution of the present paper is to offer a construction of an orthonormal basis for bi-variate α-modulation spaces. Our construction is actually the first example of a non-redundant representation system for multivariate α-modulation spaces. The orthonormal basis is constructed using a carefully calibrated tensor product approach based on so-called univariate brushlet systems. Brushlets are the image of a

2000 *Mathematics Subject Classification.* 41A17, 42B35 & 42C15.

Key words and phrases. α-modulation space, smoothness space, brushlets, local trigonometric bases, nonlinear approximation.

local trigonometric basis under the Fourier transform, and such systems were introduced by Laeng [21]. Later Coifman and Meyer [22] used brushlets as a tool for image compression. In [5], Borup and Nielsen used the freedom to choose the frequency localization of a brushlet system to construct (orthonormal) unconditional brushlet bases for the univariate α-modulation spaces. Using the orthonormal basis for bi-variate α-modulation spaces, we give a characterization of the bi-variate α-modulation spaces in terms a sparseness condition on the expansion coefficients, and we also identify the α-modulation spaces as approximation spaces associated with nonlinear m-term approximation.

Redundant frames for α-modulation spaces have been considered in a number of papers. Fornasier has studied Gabor-type Banach frames for univariate α-modulation spaces in [14]. Banach frames for univariate α-modulation spaces in the context of (generalized) coorbit theory has been studied by Dahlke et al. in [8]. Non-tight frames for multivariate α-modulation spaces were considered by Borup and Nielsen in [3], and more recently, an improved construction yielding tight frames was introduced in [6] where tight frames for more general decomposition type smoothness spaces are also considered.

The structure of the paper is as follows. In Section 2 we give the precise definition of univariate brushlets. These functions will serve as building blocks for the bi-variate construction. The main contribution of this paper can be found in Section 3 where our construction of bi-variate brushlet bases with flexible frequency properties is presented. We use a tensor product construction based on subsystems extracted from a sequence of univariate brushlet bases. This somewhat complicated approach is needed in order to obtain the right shape of the system in the frequency plane while retaining orthonormality. In Section 4, we give the precise definition of the α-modulation spaces and show how to adapt the bases from Section 3 to form unconditional bases for the α-modulation spaces. We also give a discrete characterization of the norm in the α-modulation spaces using the orthonormal basis. Section 5 contains an application to nonlinear approximation. We prove that the orthonormal basis (properly normalized) forms a so-called greedy basis for certain α-modulation spaces, and it is also show that the α-modulation spaces can be *completely* characterized in terms of nonlinear m-term approximation with the basis. Finally, there is an appendix where two of the more technical proof can be found.

2. ORTHONORMAL BRUSHLET SYSTEMS

In this section we introduce the brushlet orthonormal bases for $L_2(\mathbb{R})$. The univariate systems will then be used to carefully construct bi-variate brushlet bases that form unconditional bases for the family of α-modulation spaces.

Throughout the paper, we let $\hat{f}(\xi) = \mathcal{F}(f)(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x)e^{-ix\cdot\xi}\,dx$, $f \in L_2(\mathbb{R}^d)$, denote the Fourier transform.

2.1. **Brushlet systems.**

Let us introduce the univariate brushlets. Each brushlet basis is associated with a partition of the frequency axis. The partition can be chosen with almost no restrictions, but in order to have good properties of the associated basis we need to impose some growth conditions on the partition. We have the following definition.

Definition 2.1. A family $\mathbb{I}$ of intervals is called a *disjoint covering* of $\mathbb{R}$ if it consists of a countable set of pairwise disjoint half-open intervals $I = [\alpha_I, \alpha'_I)$, $\alpha_I < \alpha'_I$, such that $\cup_{I\in\mathbb{I}} I = \mathbb{R}$. If, furthermore, each interval in $\mathbb{I}$ has a unique adjacent interval in $\mathbb{I}$ to the left and to the right, and there exists a constant $A > 1$ such that

$$A^{-1} \leq \frac{|I|}{|I'|} \leq A, \qquad \text{for all adjacent } I, I' \in \mathbb{I}, \tag{2.1}$$

we call $\mathbb{I}$ a *moderate disjoint covering* of $\mathbb{R}$.

Given a moderate disjoint covering $\mathbb{I}$ of $\mathbb{R}$, assign to each interval $I \in \mathbb{I}$ a cutoff radius $\varepsilon_I > 0$ at the left endpoint and a cutoff radius $\varepsilon'_I > 0$ at the right endpoint, satisfying

$$\begin{cases} \text{(i)} & \varepsilon'_I = \varepsilon_{I'} \text{ whenever } \alpha'_I = \alpha_{I'} \\ \text{(ii)} & \varepsilon_I + \varepsilon'_I \leq |I| \\ \text{(iii)} & \varepsilon_I \geq c|I|, \end{cases} \tag{2.2}$$

with $c > 0$ independent of I.

We are now ready to define the brushlet system. For each $I \in \mathbb{I}$, we will construct a smooth bell function localized in a neighborhood of this interval. Take a non-negative ramp function $\rho \in C^\infty(\mathbb{R})$ satisfying

$$\rho(\xi) = \begin{cases} 0 & \text{for } \xi \leq -1, \\ 1 & \text{for } \xi \geq 1, \end{cases} \tag{2.3}$$

with the property that

$$\rho(\xi)^2 + \rho(-\xi)^2 = 1 \qquad \text{for all } \xi \in \mathbb{R}. \tag{2.4}$$

Define for each $I = [\alpha_I, \alpha'_I) \in \mathbb{I}$ the *bell function*

$$b_I(\xi) := \rho\left(\frac{\xi - \alpha_I}{\varepsilon_I}\right)\rho\left(\frac{\alpha'_I - \xi}{\varepsilon'_I}\right). \tag{2.5}$$

Notice that $\operatorname{supp}(b_I) \subset [\alpha_I - \varepsilon_I, \alpha'_I + \varepsilon'_I]$ and $b_I(\xi) = 1$ for $\xi \in [\alpha_I + \varepsilon_I, \alpha'_I - \varepsilon'_I]$. Now the set of local cosine functions

$$\hat{w}_{n,I}(\xi) = \sqrt{\frac{2}{|I|}} b_I(\xi) \cos\left(\pi\left(n + \tfrac{1}{2}\right)\frac{\xi - \alpha_I}{|I|}\right), \quad n \in \mathbb{N}_0, \quad I \in \mathbb{I}, \tag{2.6}$$

constitute an orthonormal basis for $L_2(\mathbb{R})$, see e.g. [1]. We call the collection $\{w_{n,I} \colon I \in \mathbb{I}, n \in \mathbb{N}_0\}$ a *brushlet system*. The brushlets also have an explicit representation in the time domain. Define the set of *central bell functions* $\{g_I\}_{I\in\mathbb{I}}$ by

$$\hat{g}_I(\xi) := \rho\left(\frac{|I|}{\varepsilon_I}\xi\right)\rho\left(\frac{|I|}{\varepsilon'_I}(1 - \xi)\right), \tag{2.7}$$

such that $b_I(\xi) = \hat{g}_I\big(|I|^{-1}(\xi - \alpha_I)\big)$, and let for notational convenience

$$e_{n,I} := \frac{\pi\left(n + \frac{1}{2}\right)}{|I|}, \qquad I \in \mathbb{I},\ n \in \mathbb{N}_0.$$

Then,

$$w_{n,I}(x) = \sqrt{\frac{|I|}{2}} e^{i\alpha_I x}\big\{g_I\big(|I|(x+e_{n,I})\big) + g_I\big(|I|(x-e_{n,I})\big)\big\}. \tag{2.8}$$

By a straight forward calculation it can be verified (see [2]) that for $r \geq 1$ there exists a constant $C := C(r) < \infty$, independent of $I \in \mathbb{I}$, such that

$$|g_I(x)| \leq C(1+|x|)^{-r}. \tag{2.9}$$

Thus a brushlet $w_{n,I}$ essentially consists of two well localized humps at the points $\pm e_{n,I}$.

Given a bell function b_I, define an operator $\mathcal{P}_I : L_2(\mathbb{R}) \to L_2(\mathbb{R})$ by

$$\widehat{\mathcal{P}_I f}(\xi) := b_I(\xi)\big[b_I(\xi)\hat{f}(\xi) + b_I(2\alpha_I-\xi)\hat{f}(2\alpha_I-\xi) - b_I(2\alpha_I'-\xi)\hat{f}(2\alpha_I'-\xi)\big]. \tag{2.10}$$

It can be verified that $\mathcal{P}_I$ is an orthogonal projection, mapping $L_2(\mathbb{R})$ onto $\overline{\text{span}}\{w_{n,I} : n \in \mathbb{N}_0\}$. In Section 3, we will need some of the finer properties of the operator given by (2.10). Let us list properties here, and refer the reader to [18, Chap. 1] for a more detailed discussion of the properties of local trigonometric bases.

Suppose $I = [\alpha_I, \alpha_I')$ and $J = [\alpha_J, \alpha_J')$ are two adjacent compatible intervals (i.e., $\alpha_I' = \alpha_J$ and $\varepsilon_I' = \varepsilon_J$). Then it holds true that

$$\widehat{\mathcal{P}_I f}(\xi) + \widehat{\mathcal{P}_J f}(\xi) = \hat{f}(\xi), \qquad \xi \in [\alpha_I+\varepsilon_I, \alpha_J'-\varepsilon_J'], \qquad f \in L_2(\mathbb{R}). \tag{2.11}$$

We can verify (2.11) using the fact that $b_I \equiv 1$ on $[\alpha_I+\varepsilon_I, \alpha_I'-\varepsilon_I']$ and that $b_J \equiv 1$ on $[\alpha_J+\varepsilon_J, \alpha_J'-\varepsilon_J']$, together with the fact that

$$\operatorname{supp}\big(b_I(\cdot)b_I(2\alpha_I-\cdot)\big) \subseteq [\alpha_I-\varepsilon_I, \alpha_I+\varepsilon_I]$$

and

$$\operatorname{supp}\big(b_I(\cdot)b_I(2\alpha_I'-\cdot)\big) \subseteq [\alpha_I'-\varepsilon_I', \alpha_I'+\varepsilon_I'].$$

For $\xi \in [\alpha_I'-\varepsilon_I', \alpha_J+\varepsilon_J]$ we notice that

$$\begin{aligned}\widehat{\mathcal{P}_I f}(\xi) + \widehat{\mathcal{P}_J f}(\xi) &= [b_I^2(\xi) + b_J^2(\xi)(\xi)]\hat{f}(\xi)\\ &\quad + b_J(\xi)b_J(2\alpha'-\xi)\hat{f}(2\alpha'-\xi) - b_I(\xi)b_I(2\alpha'-\xi)\hat{f}(2\alpha'-\xi).\end{aligned} \tag{2.12}$$

We can then conclude that (2.11) holds true using the following facts (see [18, Chap. 1])

$$b_I(\xi) = b_J(2\alpha_I'-\xi), \qquad b_J(\xi) = b_I(2\alpha_J'-\xi), \qquad \text{for } \xi \in [\alpha_I'-\varepsilon_I', \alpha_J+\varepsilon_J],$$

and

$$b_I^2(\xi) + b_J^2(\xi) = 1, \qquad \text{for } \xi \in [\alpha_I+\varepsilon_I, \alpha_J'-\varepsilon_J'].$$

Moreover, $\mathcal{P}_I + \mathcal{P}_J = \mathcal{P}_{I\cup J}$ with the ε-values ε_I and ε_J' for $I \cup J$.

Finally, for a rectangle $Q = I \times J \subset \mathbb{R}^2$ with $I = [\alpha_I, \alpha_I')$ and $J = [\alpha_J, \alpha_J')$, we define $P_Q = \mathcal{P}_I \otimes \mathcal{P}_J$. Clearly, P_Q is a projection operator $P_Q \colon L_2(\mathbb{R}^2) \to \overline{\text{span}}\{w_{i,I} \otimes w_{j,J} \colon i,j \in \mathbb{N}_0\}$. Notice that,

$$P_Q = S_Q\Big[(\mathrm{Id} + R_{\alpha_I} - R_{\alpha_I'}) \otimes (\mathrm{Id} + R_{\alpha_J} - R_{\alpha_J'})\Big]S_Q, \tag{2.13}$$

where $\widehat{S_Q f} := b_Q \hat{f}$, with $b_Q := b_I \otimes b_J$, and $R_a f(x) := e^{i2a} f(-x)$, $x, a \in \mathbb{R}$.

3. Bi-variate brushlet bases

The *naive* way to construct bi-variate brushlet bases from univariate brushlets is to use a simple tensor product approach. This will indeed give us a basis for $L_2(\mathbb{R}^2)$, but the time-frequency properties of the resulting system are not desirable for our purposes. In the frequency plane we end up with long "skinny" elements that are not compatible with the structure of isotropic smoothness spaces such as α-modulation spaces. The same problem occurs when we use tensor products of orthonormal wavelets. We end up with so-called hyperbolic bi-variate wavelet systems that do no offer a characterizations of isotropic smoothness spaces.

Here we follow a different path. We still use a tensor product construction, but we modify the construction by carefully keeping track of the shape of the system in the frequency plane by extracting subsystems from a sequence of brushlet bases.

To keep the notation manageable, we consider only the bi-variate case, but the reader should notice that the basic idea behind the construction can be adapted to the general multivariate case.

3.1. **A sequence of brushlets.**

To get the bi-variate construction off the ground we begin by construction a sequence of univariate brushlet systems. The univariate systems are not complete, but the idea is to carefully "glue" them together later to create the wanted bi-variate system.

Let $n \geq 2$ and fix $\beta \geq 1$. We create a brushlet system living on the frequency interval $[-(n+1)^\beta, (n+1)^\beta]$ subdivided into $n+3$ intervals. Since our goal is to fit the systems together later for different values of n, we divide $[-(n+1)^\beta, (n+1)^\beta]$ carefully by putting knots the points $\pm(n-1)^\beta$ and $\pm n^\beta$. The "inner" interval

$$[-(n-1)^\beta, (n-1)^\beta]$$

is then divided into $n-1$ intervals of equal length, see Figure 1.

FIGURE 1. Brushlet partition at level n.

Notice that the length of each of the $n+3$ subintervals is of the order $n^{\beta-1}$ uniformly in n. We denote by I_k^n, $k = 1, \ldots, n+3$, the $n+3$ subintervals intervals, taken sequentially from the left to the right. Next we define the ε-values at the breakpoints needed for the brushlet system. We fix $\gamma := \gamma(\beta)$ sufficiently small (see below for details). Define the corresponding ε-values at the breakpoints as follows: at $\pm n^\beta$ we use $\varepsilon = \gamma n^{\beta-1}$ and at $\pm(n+1)^\beta$ we use $\varepsilon = \gamma(n+1)^{\beta-1}$. At the remaining breakpoints we use $\varepsilon = \gamma(n-1)^{\beta-1}$. The basic estimate

$$3^{-\beta} \leq \left(\frac{n-1}{n+1}\right)^\beta \leq 1, \qquad n \geq 2,$$

can be used to verify that $\gamma = \frac{\beta}{16}3^{-\beta}$ is sufficiently small to work for all n. We now create the associated orthonormal brushlet system

$$\mathcal{W}^n := \bigcup_{i=1}^{n+3} \{w_{k,I_i^n}\}_{k\in\mathbb{N}_0},$$

numbered from the left to the right so that w^n_{1,I_1^n} is associated with the frequency interval $I_1^n = [-(n+1)^\beta, -n^\beta]$ and $w_{k,I^n_{n+3}}$ is associated with the frequency interval $I^n_{n+3} = [n^\beta, (n+1)^\beta]$.

Remark 3.1. Notice that the construction ensures that the functions on level n associated with the extremal frequency intervals I_1^n and I_{n+3}^n are all contained in the orthonormal system on level $n+1$. So, in particular, the functions $w_{j,I_1^{n+1}}$ and w_{k,I_1^n}, $j,k \in \mathbb{N}_0$, from adjacent levels are orthogonal. The same holds true for $w_{j,I^{n+1}_{n+4}}$ and $w_{k,I^n_{n+3}}$, $j,k\in\mathbb{N}_0$.

3.2. **Bi-variate brushlets.** We now define a orthonormal bivariate system living on the rectangular frequency "annulus" A_n defined by the eight points $\pm((n+1)^\beta, \pm(n+1)^\beta)$ and $\pm(n^\beta, \pm n^\beta)$. We use the the system $\mathcal{W}^n = \cup_{i=1}^{n+3}\{w_{k,I_i^n}\}_{k\in\mathbb{N}_0}$ for the construction. Let

$$\mathcal{A}^n := E_1^n \cup E_2^n \cup E_3^n \cup E_4^n, \tag{3.1}$$

where

$$\begin{aligned}
E_1^n &:= \{\omega^{I\times J}_{(j,k)}(x,y) := w_{j,I}(x)w_{k,J}(y)\}_{(j,k)\in\mathbb{N}_0^2, I\times J\in\mathbb{P}_1^n}, & \mathbb{P}_1^n &:= \{I_1^n \times I_i^n\}_{i=1}^{n+3},\\
E_2^n &:= \{\omega^{I\times J}_{(j,k)}(x,y) := w_{j,I}(x)w_{k,J}(y)\}_{(j,k)\in\mathbb{N}_0^2, I\times J\in\mathbb{P}_2^n}, & \mathbb{P}_2^n &:= \{I_i^n \times I_{n+3}^n\}_{i=2}^{n+3},\\
E_3^n &:= \{\omega^{I\times J}_{(j,k)}(x,y) := w_{j,I}(x)w_{k,J}(y)\}_{(j,k)\in\mathbb{N}_0^2, I\times J\in\mathbb{P}_3^n}, & \mathbb{P}_3^n &:= \{I_{n+3}^n \times I_i^n\}_{i=1}^{n+2},\\
E_4^n &:= \{\omega^{I\times J}_{(j,k)}(x,y) := w_{j,I}(x)w_{k,J}(y)\}_{(j,k)\in\mathbb{N}_0^2, I\times J\in\mathbb{P}_4^n}, & \mathbb{P}_4^n &:= \{I_i^n \times I_1^n\}_{i=2}^{n+2}.
\end{aligned}$$

The following figure illustrates the frequency content of each family E_i^n.

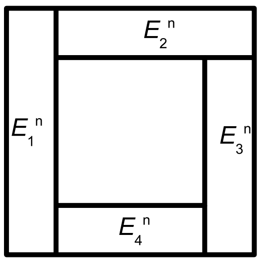

FIGURE 2. The frequency content of E_i^n, $i = 1,\dots,4$.

We observe that the system $\mathcal{A}^n$ splits naturally into $4n+8$ subsystems each associated with a distinct frequency support rectangle from $\mathbb{P}^n = \mathbb{P}_1^n \cup \cdots \mathbb{P}_4^n$, providing a natural partition of the annulus A_n. We will use the condensed notation $\mathcal{A}^n = \{\omega_k^Q\}_{Q\in\mathbb{P}^n, k\in\mathbb{N}_0^2}$ extensively. We also observe that $\mathcal{A}^n$ is an orthonormal system in $L_2(\mathbb{R}^2)$, which follows by direct inspection using that $\mathcal{W}^n$ is an orthonormal system in $L_2(\mathbb{R})$.

To complete the picture, we denote by $\{w_{k,[-1,1)}\}_{k\in\mathbb{N}_0}$ the "low-pass" brushlet system associated with $[-1,1)$ with ε values γ, and we define

$$\mathcal{A}^1 := \bigcup_{j,k\in\mathbb{N}_0} \left\{\omega_{(j,k)}^{[-1,1)\times[-1,1]} := w_{j,[-1,1)}(x)w_{k,[-1,1)}(y)\right\}.$$

We are now ready to define the full bivariate system by taking the union of all the sets $\mathcal{A}^n$. The full system is defined by

$$\mathcal{W}(\beta) := \bigcup_{n=1}^{\infty} \mathcal{A}^n. \tag{3.2}$$

We let

$$\mathbb{P} = [-1,1]^2 \cup \left(\bigcup_{n=2}^{\infty} \mathbb{P}^n\right)$$

denote the collection of all elementary frequency rectangles associated with $\mathcal{W}(\beta)$. Hence, we can write $\mathcal{W}(\beta) = \bigcup_{Q\in\mathbb{P}}\{\omega_k^Q\}_{k\in\mathbb{N}_0^2}$. For $Q\in\mathbb{P}$, we let P_Q denote the orthogonal projection onto $\overline{\operatorname{span}}(\{\omega_k^Q\}_{k\in\mathbb{N}_0^2})$ given by (2.13), i.e.,

$$P_Q f = \sum_{k\in\mathbb{N}_0^2} \langle f, \omega_k^Q\rangle \omega_k^Q. \tag{3.3}$$

Notice how the frequency content of the sets $\mathcal{A}^n$ cover the whole frequency plane indicating that the construction is reasonable and will give us a complete orthonormal system. Let us state and prove the most fundamental property of $\mathcal{W}(\beta)$.

Proposition 3.2. *The system $\mathcal{W}(\beta)$ defined by* (3.2) *is an orthonormal basis for $L_2(\mathbb{R}^2)$.*

Proof. Let us first verify that the system is orthonormal. Recall that each $\mathcal{A}^n$ is an orthonormal system. It follows from Remark 3.1 that the system $\mathcal{A}^n$ is orthogonal to $\mathcal{A}^{n+1}$, $n\geq 2$. Moreover, $\mathcal{A}^m$ and $\mathcal{A}^n$ have disjoint frequency support whenever $|m-n|\geq 2$. Finally, we notice that $\mathcal{A}^2$ is orthogonal to $\{\omega_k^{[0,1)^2}\}_{k\in\mathbb{N}_0^2}$ due to their construction (the univariate brushlets have the same ε-value at the breakpoints ± 1). We conclude that $\mathcal{W}(\beta)$ is orthonormal.

The slightly more involved part is to verify that $\mathcal{W}(\beta)$ is complete in $L_2(\mathbb{R}^2)$. Observe that the frequency properties of the system $\mathcal{W}(\beta)$ ensure that for $f\in L_2(\mathbb{R}^2)$ and $\xi\in\mathbb{R}^2$, $\sum_{Q\in F}\widehat{P_Q f}(\xi)$ contains at most four non-zero terms, which can easily be used to deduce that $\sum_Q P_Q$ converges strongly to a bounded operator on $L_2(\mathbb{R}^2)$. Hence, it suffices to prove that

$$\sum_{Q\in\mathbb{P}} \widehat{P_Q s} = \hat{s} \tag{3.4}$$

for functions s from a suitable dense subset of $L_2(\mathbb{R}^2)$. Since finite linear combinations of separable functions are dense in $L_2(\mathbb{R}^2)$, we just need to verify (3.4) for a separable function $s(x,y)=g(x)h(y)$ with $g,h\in L_2(\mathbb{R})$.

The sum (3.4) is local in nature, and first we consider it pointwise on a rectangle of the form

$$R_V^n := [-(1-\gamma)(n+1)^\beta, -(1-\gamma)n^\beta]\times[-(1-\gamma)(n+1)^\beta, (1-\gamma)(n+1)^\beta],$$

with $n \geq 3$ fixed. This will restrict the analysis to a vertical frequency strip associated with the systems $\mathcal{A}^{n-1}$ and $\mathcal{A}^n$, see Figure 3.

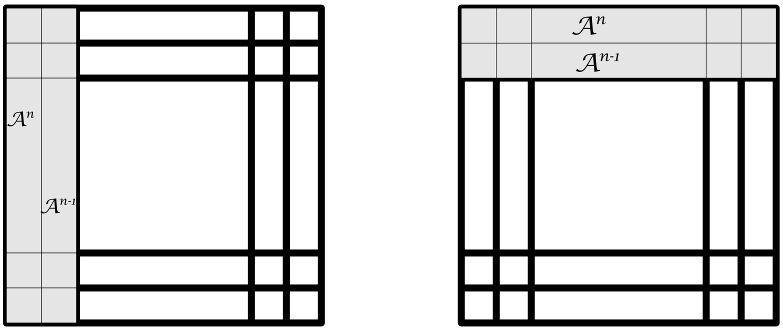

FIGURE 3. A local interpretation of the sum (3.4). It is advantageous to factor the terms so we can do the summation “vertically” or “horizontally”.

For $\xi = (\xi_1, \xi_2) \in R_V^n$, we have

$$\sum_Q \widehat{P_Q s}(\xi) = \sum_{Q \in F : Q \cap R_V^n \neq \emptyset} \widehat{P_Q s}(\xi).$$

We let $\mathcal{P}_{I_i^n}$ denote the $L_2(\mathbb{R})$-orthogonal projection onto $\overline{\text{span}}(\{w_{j,I_i^n}\}_{j \in \mathbb{N}_0})$. We now use (2.12) repeatedly, together with the fact that by construction, $\mathcal{P}_{I_2^n} = \mathcal{P}_{I_1^{n-1}}$ and $\mathcal{P}_{I_{n+2}^n} = \mathcal{P}_{I_{n+2}^{n-1}}$, to obtain

$$\begin{aligned}
\sum_{Q \in F : Q \cap R_V^n \neq \emptyset} \widehat{P_Q s}(\xi) &= \widehat{\mathcal{P}_{I_1^n} g}(\xi_1) \sum_{i=1}^{n+3} \widehat{\mathcal{P}_{I_i^n} h}(\xi_2) + \widehat{\mathcal{P}_{I_1^{n-1}} g}(\xi_1) \times \\
&\qquad \left(\sum_{i=1}^{n+2} \widehat{\mathcal{P}_{I_i^{n-1}} h}(\xi_2) + \widehat{\mathcal{P}_{I_1^n} h}(\xi_2) + \widehat{\mathcal{P}_{I_{n+3}^n} h}(\xi_2) \right) \\
&= \widehat{\mathcal{P}_{I_1^n} g}(\xi_1) \hat{h}(\xi_2) + \widehat{\mathcal{P}_{I_1^{n-1}} g}(\xi_1) \hat{h}(\xi_2) \\
&= \hat{g}(\xi_1) \hat{h}(\xi_2).
\end{aligned}$$

The same argument can be used to verify (3.4) on the horizontal rectangle

$$R_H^n := [-(1-\gamma)(n+1)^\beta, (1-\gamma)(n+1)^\beta] \times [(1+\gamma)n^\beta, (1-\gamma)(n+1)^\beta],$$

and on the “mirrored” rectangle

$$\widetilde{R_V^n} := [(1+\gamma)n^\beta, (1-\gamma)(n+1)^\beta] \times [-(1-\gamma)(n+1)^\beta, (1-\gamma)(n+1)^\beta],$$

together with

$$\widetilde{R_H^n} := [-(1-\gamma)(n+1)^\beta, (1-\gamma)(n+1)^\beta] \times [-(1-\gamma)(n+1)^\beta, -(1+\gamma)n^\beta].$$

We now let n vary to cover the whole plane, where we make the remark that a similar argument as above works where the low-pass system $\mathcal{A}^1$ meets $\mathcal{A}^2$. Consequently, we conclude that (3.4) holds true everywhere for $s = g \cdot h$. Hence, (3.4) holds true on a dense subset of $L_2(\mathbb{R}^2)$ and we deduce that $\mathcal{W}(\beta)$ is indeed complete in $L_2(\mathbb{R}^2)$. $\square$

4. Orthonormal bases for α-modulation spaces

Now we have a family $\mathcal{W}(\beta)$ of orthonormal bases for $L_2(\mathbb{R}^2)$ with time-frequency properties that can be "tuned" using the parameter $\beta \geq 1$. Our claim is that these bases are well-suited to analyze the family of α-modulation spaces. In this section, we give the precise definition of the α-modulation spaces, and we show that by selecting the appropriate β value relative to α, $\mathcal{W}(\beta)$ will form an unconditional basis for the full range of α-modulation spaces defined on $\mathbb{R}^2$.

The α-modulation spaces are defined by a parameter α, belonging to the interval $[0,1]$. This parameter determines a segmentation of the frequency plane from which the spaces are built. Thus, we need to define "nice" partitions of the frequency space. We should also point out that α-modulation can be considered on any $\mathbb{R}^d$, and the discussion below is restricted to $d = 2$ only because our basis from Section 3.2 is bivariate.

Let $B(c,r) \subset \mathbb{R}^2$ denote the open disc with center c and radius r, and let $\langle x \rangle := (1+|x|^2)^{1/2}$.

Definition 4.1. A countable set $\mathcal{Q}$ of subsets $Q \subset \mathbb{R}^2$ is called an admissible covering if $\mathbb{R}^2 = \cup_{Q\in\mathcal{Q}} Q$ and there exists $n_0 < \infty$ such that $\#\{Q' \in \mathcal{Q} : Q \cap Q' \neq \emptyset\} \leq n_0$ for all $Q \in \mathcal{Q}$. Let

$$r_Q = \sup\{r \in \mathbb{R}\colon B(c_r, r) \subset Q \text{ for some } c_r \in \mathbb{R}^2\},$$
$$R_Q = \inf\{R \in \mathbb{R}\colon Q \subset B(c_R, R) \text{ for some } c_R \in \mathbb{R}^2\}$$

denote respectively the radius of the inscribed and circumscribed disc of $Q \in \mathcal{Q}$. An admissible covering is called an α-covering, $0 \leq \alpha \leq 1$, of $\mathbb{R}^2$ if $|Q| \asymp \langle x \rangle^{\alpha d}$ (uniformly) for all $x \in Q$ and for all $Q \in \mathcal{Q}$, and there exists a constant $K \geq 1$ such that $R_Q/r_Q \leq K$ for all $Q \in \mathcal{Q}$.

We also need partitions of unity compatible with the covers from Definition 4.1.

Definition 4.2. Given $p \in (0,\infty]$ and an α-covering $\mathcal{Q}$ of $\mathbb{R}^2$. A corresponding bounded admissible partition of unity of order p (p-BAPU) $\{\psi_Q\}_{Q\in\mathcal{Q}}$ is a family of functions satisfying

- $\operatorname{supp}(\psi_Q) \subset Q$
- $\sum_{Q\in\mathcal{Q}} \psi_Q(\xi) = 1$
- $\sup_Q |Q|^{1/\tilde{p}-1} \|\mathcal{F}^{-1}\psi_Q\|_{L_{\tilde{p}}} < \infty$, $\tilde{p} := \min(1,p)$.

Remark 4.3. It is proved in [6] that an α-covering with a corresponding p-BAPU actually exist for every $\alpha \in [0,1]$ and $p > 0$.

For $Q \in \mathcal{Q}$ define the multiplier $\psi_Q(D)f := \mathcal{F}^{-1}(\psi_Q \mathcal{F} f)$, $f \in L_2(\mathbb{R}^2)$. By [26, Proposition 1.5.1], the conditions in Definition 4.2 ensure that $\psi_Q(D)$ extends to a bounded operator on the bandlimited functions in $L_p(\mathbb{R}^2)$, $0 < p \leq \infty$, uniformly in $Q \in \mathcal{Q}$.

We have the following definition of the α-modulation spaces.

Definition 4.4. Given $0 < p, q \leq \infty$, $s \in \mathbb{R}$, and $0 \leq \alpha \leq 1$, let $\mathcal{Q}$ be an α-covering of $\mathbb{R}^2$ for which there exists a p-BAPU Ψ. Then we define the *α-modulation space*, $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ as the set of distributions $f \in S'(\mathbb{R}^2)$ satisfying

$$\|f\|_{M^{s,\alpha}_{p,q}} := \Big(\sum_{Q\in\mathcal{Q}} \langle \xi_Q\rangle^{qs}\|\psi_Q(D)f\|^q_{L_p}\Big)^{1/q} < \infty, \tag{4.1}$$

with $\{\xi_Q\}_{Q\in\mathcal{Q}}$ a sequence satisfying $\xi_Q \in Q$. For $q=\infty$ we have the usual change of the sum to sup over $Q\in\mathcal{Q}$.

It is easy to see that (4.1) defines a quasi-norm (or a norm if $p,q\geq 1$) on $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ and that two different sequences $\{\xi_Q\}_{Q\in\mathcal{Q}}$ give equivalent norms. It it also known that two different p-BAPUs give rise to equivalent norms, see [3]. The definition is given for the full range of p and q, extending Gröbner's original definition in [17].

It can be proved that $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ is a quasi-Banach space, and that

$$\mathcal{S}(\mathbb{R}^2) \hookrightarrow M^{s,\alpha}_{p,q}(\mathbb{R}^2) \hookrightarrow \mathcal{S}(\mathbb{R}^2)',$$

see [5]. Moreover, if $p,q<\infty$, $\mathcal{S}(\mathbb{R}^2)$ is dense in $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$.

4.1. **Bi-variate brushlet bases for α-modulation spaces.** Now we consider a characterization of $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ using the system $\mathcal{W}(\beta)$. First we have to choose the right β, given $\alpha\in[0,1)$. This is easy, we just notice that any rectangle contained in $\mathbb{P}^n$ has area $\mathcal{O}(n^{2(\beta-1)})$ while for $Q\in\mathbb{P}^n$, and $\xi\in Q$, we have $\langle\xi\rangle \asymp n^\beta$. We have $n^{2(\beta-1)}\asymp n^{2\beta\alpha}$ exactly when $\beta = 1/(1-\alpha)$. It is to be expected that we cannot define β for $\alpha=1$, since $\alpha=1$ corresponds to a dyadic covering of $\mathbb{R}^2$ while the covering associated with $\mathcal{W}(\beta)$ is always of "polynomial" type. The choice $\beta=1/(1-\alpha)$ leads to the following characterization.

Proposition 4.5. *Given $\alpha\in[0,1)$, $0<p,q\leq\infty$, and $s\in\mathbb{R}$. Let $\beta=1/(1-\alpha)$, and consider the associated system $\mathcal{W}(\beta)=\{\omega^Q_k\}_{Q\in\mathbb{P},k}$. For any $f\in M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ we have*

$$\|f\|_{M^{s,\alpha}_{p,q}(\mathbb{R}^2)} \asymp \Big(\sum_{Q\in\mathbb{P}}\langle\xi_Q\rangle^{qs}\|P_Qf\|^q_{L_p}\Big)^{1/q}, \tag{4.2}$$

where P_Q is the projection given by (3.3), *and $\xi_Q\in Q$.*

The proof of Proposition 4.5 is similar to the proof of Theorem 3.1 in [3], and it has been included in Appendix A for the sake of completeness. In the following Lemma we estimate the L_p-norm of the projection P_Qf in terms of the associated Fourier coefficients of f relative to $\{\omega^Q_k\}_k$.

Lemma 4.6. *With the same assumptions as in Proposition 4.5, suppose $f\in L_2(\mathbb{R}^2)$, $Q\in\mathbb{P}$, and $p\in(0,\infty]$. Then $P_Qf\in L_p(\mathbb{R}^2)$ if and only if $\{\langle f,\omega^Q_k\rangle\}_{k\in\mathbb{N}_0^2}\in\ell_p$. In fact, if one of these conditions are satisfied we have*

$$\|P_Qf\|_{L_p(\mathbb{R}^2)} \asymp |Q|^{\frac12-\frac1p}\Big(\sum_{k\in\mathbb{N}_0^2}|\langle f,\omega^Q_k\rangle|^p\Big)^{1/p}, \qquad 0<p<\infty, \tag{4.3}$$

with equivalence independent of Q. When $p=\infty$ the sum in (4.3) *is changed to* sup *over $k\in\mathbb{N}_0^2$.*

The proof Lemma 4.6 can be found in Appendix A. We can now combine Lemma 4.6 and Proposition 4.5 to derive our main result on properties of the system $\mathcal{W}(\beta)$ in $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$.

Theorem 4.7. *Given $0 < p, q \leq \infty$, $s \in \mathbb{R}$, and $0 \leq \alpha < 1$. For the system $\mathcal{W}(\beta) = \{\omega_k^Q\}$ with $\beta = 1/(1-\alpha)$, we have the characterization*

$$\|f\|_{M^{s,\alpha}_{p,q}(\mathbb{R}^2)} \asymp \Big(\sum_{n=1}^{\infty} \sum_{Q\in\mathbb{P}^n} \Big(\sum_{k\in\mathbb{N}_0^2} \big(n^{\frac{1}{1-\alpha}(s+\alpha-\frac{2\alpha}{p})} |\langle f, \omega_k^Q\rangle|\big)^p\Big)^{q/p}\Big)^{1/q}. \tag{4.4}$$

Moreover, for $1 \leq p, q < \infty$, $\mathcal{W}(\beta)$ constitutes an unconditional basis for $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$.

Proof. The norm characterization follows at once by combining Lemma 4.6 and Proposition 4.5, using the facts that for $\xi_Q \in Q$, $\langle \xi_Q\rangle^{2\alpha} \asymp |Q|$ uniformly in $Q \in \mathbb{P}$, and for $Q \in \mathbb{P}_n$, $\langle \xi_Q\rangle \asymp n^\beta = n^{1/(1-\alpha)}$. The claim that the system form an unconditional basis when $p, q \geq 1$ follows easily from the fact that $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ is a Banach space, and that finite expansions in $\mathcal{W}(\beta)$ have uniquely determined coefficients giving us a norm characterization of such expansions using (4.4). □

Theorem 4.7 shows that $\mathcal{W}(\beta)$ induces a natural isomorphism between $M^{s,\alpha}_{p,q}(\mathbb{R}^2)$ and the sequence space $m^{s,\alpha}_{p,q}$ given by

Definition 4.8. Given $0 < p, q \leq \infty$, $s \in \mathbb{R}$, and $0 \leq \alpha < 1$, we define the space $m^{s,\alpha}_{p,q}$ as the set of coefficients $c = \{c_k^Q\}_{k\in\mathbb{N}_0^2, Q\in\mathbb{P}} \subset \mathbb{C}$ satisfying

$$\|c\|_{m^{s,\alpha}_{p,q}} := \Big(\sum_{n=1}^{\infty} \sum_{Q\in\mathbb{P}^n} \Big(\sum_{k\in\mathbb{N}_0^2} \big(n^{\frac{1}{1-\alpha}(s+\alpha-\frac{2\alpha}{p})} |c_k^Q|\big)^p\Big)^{q/p}\Big)^{1/q} < \infty.$$

5. An application to nonlinear approximation

In this final section we consider m-term nonlinear approximation with $\mathcal{W}(\beta)$ in certain α-modulation spaces.

Let us first point out why it is important to have access to a non-redundant representation system when it comes to m-term nonlinear approximation. Approximation with redundant systems in multivariate α-modulation spaces have been studied in a number of papers, see [3, 6]. The ultimate goal is to characterize the approximation class $\mathcal{A}^\gamma$ that (roughly) contain the elements in $M^{s,\alpha}_{p,q}$ for which m-term approximation converges at the rate $\mathcal{O}(m^{-\gamma})$, see Definition 5.2 for the precise statement. However, the earlier results have only been concerned with so-called direct or Jackson estimates. A direct estimate makes it possible to identify a certain set contained in the approximation class, so we obtain a "minimum size" estimate of the approximation class. But it may happen that the class is larger than this one-sided estimate indicates. What has been missing from the picture are so-called inverse or Bernstein estimates providing a "maximum size" estimate of the approximation class. Such estimates have been out of reach for the simple reason that the frames considered are *redundant*. It might well be that a Bernstein estimate exists for the redundant frames from [3,6,8,14], but at present it is not known how to prove such estimates. However, for our non-redundant basis such technical difficulties disappear and below we derive a complete characterization of m-term nonlinear approximation with the $\mathcal{W}(\beta)$.

Recent results [16,20] have shown that to characterize nonlinear m-term approximation with elements from a Schauder basis from a Banach space, it is advantageous to deal with so-called greedy bases. Below we demonstrate that a properly normalized version of $\mathcal{W}(\beta)$ form greedy bases for the α-modulation spaces, and from this fact we deduce several direct and inverse estimates for nonlinear m-term approximation with $\mathcal{W}(\beta)$.

First, let us define a greedy basis as introduced by Konyagin and Temlyakov [20]. A greedy basis in a Banach space is an unconditional basis that also satisfies the so-called *democracy* condition.

Definition 5.1. A system $\{g_k\}_{k\in\mathbb{N}}$ in a (quasi-)Banach space X is called *democratic* if there exists a constant $C > 0$ such that

$$\Big\|\sum_{k\in I} g_k\Big\|_X \leq C\Big\|\sum_{k\in J} g_k\Big\|_X,$$

for any two finite sets of indices I and J with the same cardinality, $\#I = \#J$.

Let us fix $\alpha \in (0,1)$, and as before, we put $\beta = 1/(1-\alpha)$. For $0 < p < \infty$, we consider the system $\mathcal{W}(\beta) = \{\omega_k^Q\}$, and the associated normalized system $\widetilde{\mathcal{W}}(\beta) = \{\widetilde{\omega_k^Q}\}$ defined by

$$\widetilde{\omega_k^Q} = \frac{\omega_k^Q}{\|\omega_k^Q\|_{M^{s,\alpha}_{p,p}(\mathbb{R}^2)}}, \qquad Q \in \mathbb{P}, k \in \mathbb{N}_0^2.$$

Observe that by Theorem 4.7,

$$\|f\|_{M^{s,\alpha}_{p,p}(\mathbb{R}^2)} \asymp \Big(\sum_{Q\in\mathbb{P}}\sum_{k\in\mathbb{N}_0^2} |\langle f, \widetilde{\omega_k^Q}\rangle|^p\Big)^{1/p}. \tag{5.1}$$

Thus, for a finite subset $\Lambda \subset \mathbb{P}\times\mathbb{N}_0^2$, we have the uniform estimate

$$\Big\|\sum_{(Q,k)\in\Lambda} \widetilde{\omega_k^Q}\Big\|_{M^{s,\alpha}_{p,p}(\mathbb{R}^2)} \asymp (\#\Lambda)^{1/p}, \tag{5.2}$$

which shows that $\{\widetilde{\omega_k^Q}\}_{Q,k}$ is a democratic and unconditional (i.e., a greedy) basis for $M^{s,\alpha}_{p,p}(\mathbb{R}^2)$ provided $p, q \geq 1$.

Let us introduce some notation that will be needed to explore nonlinear approximation with brushlet bases. Let $\mathcal{D} = \{g_k\}_{k\in\mathbb{N}}$ be a Schauder basis in a Banach space X. We consider the collection of all possible m-term expansions with elements from $\mathcal{D}$:

$$\Sigma_m(\mathcal{D}) := \Big\{\sum_{i\in\Lambda} c_i g_i \,\Big|\, c_i \in \mathbb{C}, \#\Lambda \leq m\Big\}.$$

The error of the best m-term approximation to an element $f \in X$ is then

$$\sigma_m(f,\mathcal{D})_X := \inf_{f_m\in\Sigma_m(\mathcal{D})} \|f - f_m\|_X.$$

Let us introduce the approximation classes $\mathcal{A}^\gamma_q(X,\mathcal{D})$ associated with $\mathcal{D}$. The approximation classes $\mathcal{A}^\gamma_q(X,\mathcal{D})$ is an important for the study of m-term approximation using $\mathcal{D}$, and it (essentially) consists of the elements f for which $\sigma(f,\mathcal{D}) = \mathcal{O}(n^{-\gamma})$.

Definition 5.2 (Approximation spaces)**.** The approximation space $\mathcal{A}_q^\gamma(X,\mathcal{D})$ is defined by

$$|f|_{\mathcal{A}_q^\gamma(X,\mathcal{D})} := \left(\sum_{m=1}^\infty \left(m^\gamma \sigma_m(f,\mathcal{D})_X\right)^q \frac{1}{m}\right)^{1/q} < \infty,$$

and (quasi)normed by $\|f\|_{\mathcal{A}_q^\gamma(X,\mathcal{D})} = \|f\|_X + |f|_{\mathcal{A}_q^\gamma(X,\mathcal{D})}$ for $0 < q,\gamma < \infty$, with the ℓ_q norm replaced by the sup-norm, when $q = \infty$.

We also need to define smoothness spaces in order to characterize the approximation spaces. We give the definition in an abstract setting, but later in this section (Proposition 5.4) it is proved that the smoothness spaces corresponding to brushlet systems can be identified with certain α-modulation spaces.

For $\tau \in (0,\infty)$ and $s \in (0,\infty]$, we let $\mathcal{K}^\tau_s(\mathcal{D},M)$ denote the set

$$\operatorname{clos}_X\left\{f \in X \mid \exists \Lambda \subset \mathbb{N}, \#\Lambda < \infty, f = \sum_{k\in\Lambda} c_k g_k, \|\{c_k\}\|_{\ell_{\tau,s}} \le M\right\},$$

where $\ell_{\tau,s}$ is the Lorentz sequence norm with primary parameter τ and secondary parameter s, see [11]. Then we define

$$\mathcal{K}^\tau_s(X,\mathcal{D}) := \bigcup_{M>0} \mathcal{K}^\tau_s(\mathcal{D},M), \tag{5.3}$$

with

$$\|f\|_{\mathcal{K}^\tau_s(X,\mathcal{D})} = \inf\{M : f \in \mathcal{K}^\tau_s(\mathcal{D},M)\}.$$

For a democratic basis $\mathcal{D} = \{g_k\}_{k\in\mathbb{N}}$ in X, we define $\varphi(n) := \left\|\sum_{k=1}^n g_k\right\|_X$. The following theorem was proved (independently) in [16] and [19].

Theorem 5.3. *Assume $\mathcal{D}$ is a greedy basis for a Banach space X with $\varphi(n) \asymp n^{1/p}$. Then*

$$\mathcal{A}_q^\gamma(X,\mathcal{D}) = \mathcal{K}^\tau_q(X,\mathcal{D}), \qquad \tau^{-1} = \gamma + p^{-1}, \gamma > 0,$$

with equivalent norms.

One problem with Theorem 5.3 is the we are restricted to the Banach space case, which in our setup means that $p,q \ge 1$. However, Theorem 5.3 has recently been extended to the quasi-Banach space case by Garrigós Hernández [15] under the assumption that the dictionary $\mathcal{D}$ induces an isomorphism with a suitable discrete Triebel-Lizorkin sequence space. This is indeed the case here where (5.1) shows that $\mathcal{W}(\beta)$ induces an isomorphism between $M^{s,\alpha}_{p,p}(\mathbb{R}^2)$ and ℓ_p.

We can thus apply Theorem 5.3 and its extension in [15] to $\mathcal{W}(\beta)$, using the estimate (5.2), to obtain the first complete characterization of nonlinear approximation classes associated with a bi-variate α-modulation space.

Theorem 5.4. *For $\alpha \in (0,1)$, $0 < p,q < \infty$ and $s \in \mathbb{R}$, we let $\beta = 1/(1-\alpha)$ and consider the system $\mathcal{D} := \widetilde{\mathcal{W}}(\beta)$ normalized in $M^{s,\alpha}_{p,p}(\mathbb{R}^2)$. We have the characterization*

$$\mathcal{A}_q^\gamma\big(M^{s,\alpha}_{p,p}(\mathbb{R}^2),\mathcal{D}\big) = \mathcal{K}^\tau_q\big(M^{s,\alpha}_{p,p}(\mathbb{R}^2),\mathcal{D}\big), \qquad \tau^{-1} = \gamma + p^{-1}, \gamma > 0, 0 < q \le \infty,$$

with equivalent norms. Moreover, for $\tau > 0$,

$$\mathcal{K}^\tau_\tau\big(M^{s,\alpha}_{p,p}(\mathbb{R}^2),\mathcal{D}\big) = M^{\rho,\alpha}_{\tau,\tau}(\mathbb{R}^2), \qquad \textit{with} \quad \rho = \frac{2\alpha}{\tau} - \frac{2\alpha}{p} + s.$$

Proof. The first part of the Proposition follows at once from Theorem 5.3 and the estimate (5.2). The proof of the second claim reduces to getting the normalization right. We use the fact that for $Q \in \mathbb{P}$, $\|\omega_k^Q\|_{M^{s,\alpha}_{p,p}(\mathbb{R}^2)} \asymp |Q|^{\frac{s}{2\alpha}+1/2-1/p}$, which can easily be deduced from Theorem 4.7. Again using Theorem 4.7, we deduce the following equation for ρ,

$$|Q|^{(\frac{\rho}{2\alpha}+1/2-1/\tau)-(\frac{s}{2\alpha}+1/2-1/p)} = |Q|^0,$$

from which we obtain $\rho = \frac{2\alpha}{\tau} - \frac{2\alpha}{p} + s$. □

Appendix A. Technical proofs

We have postponed two technical proof to this appendix. First we give a proof of Proposition 4.5.

Proof of Proposition 4.5. Let Ψ be a p-BAPU subordinate to an α-covering $\mathcal{Q}$. Take $f \in M^{s,\alpha}_{p,q}(\mathbb{R}^2)$. Then,

$$P_Q f = \sum_{Q' \in A_Q} P_Q(\psi_{Q'}(D)f), \qquad Q \in \mathbb{P},$$

in $\mathcal{S}'(\mathbb{R}^2)$, where A_Q is the collection of $Q' \in \mathcal{Q}$ with $\operatorname{supp}(b_Q) \cap Q' \neq \emptyset$. It is known (see [6])

$$\sup_{Q \in \mathbb{P}} \#A_Q \leq d_A < \infty.$$

Using (2.5) we have that $\|\mathcal{F}^{-1}b_Q\|_{L_p} \leq C_p |Q|^{1-1/p}$, for $0 < p \leq \infty$. We have $\psi_{Q'}(D)f \in L_p$, $0 < p \leq \infty$, so for any $Q' \in \mathcal{Q}$, Proposition 1.5.1 in [26] implies

$$\begin{aligned}\|P_Q f\|_{L_p} &\leq C|Q|^{1/\tilde{p}-1}\|\mathcal{F}^{-1}b_Q\|_{L_{\tilde{p}}} \sum_{Q' \in A_Q} \|\psi_{Q'}(D)f\|_{L_p} \qquad (\tilde{p} := \min(1,p)) \\ &\leq C' \sum_{Q' \in A_Q} \|\psi_{Q'}(D)f\|_{L_p}\end{aligned}$$

with C' independent of Q. Clearly, $\langle \xi_Q \rangle \asymp \langle \xi_{Q'} \rangle$ for any $\xi_Q \in Q$, $\xi_{Q'} \in Q'$, when $Q' \in A_Q$. Hence

$$\langle \xi_Q \rangle^s \|P_Q f\|_{L_p} \leq C' \sum_{Q' \in A_Q} \langle \xi_{Q'} \rangle^s \|\psi_{Q'}(D)f\|_{L_p} \qquad \text{for } 0 < p \leq \infty, \tag{A.1}$$

with C' independent of Q. Suppose $0 < q \leq 1$, then

$$\begin{aligned}\sum_{Q \in \mathbb{P}} \Big(\sum_{Q' \in A_Q} \langle \xi_{Q'} \rangle^s \|\psi_{Q'}(D)f\|_{L_p} \Big)^q &= \sum_{Q \in \mathbb{P}} \Big(\sum_{Q' \in \mathcal{Q}} \mathbf{1}_{A_Q}(Q') \langle \xi_{Q'} \rangle^s \|\psi_{Q'}(D)f\|_{L_p} \Big)^q \\ &\leq \sum_{Q' \in \mathcal{Q}} \sum_{Q \in \mathbb{P}} \big(\mathbf{1}_{A_Q}(Q') \langle \xi_{Q'} \rangle^s \|\psi_{Q'}(D)f\|_{L_p}\big)^q,\end{aligned}$$

where $\mathbf{1}_{A_Q}(Q') = 1$ for $Q' \in A_Q$ and 0 for $Q' \in \mathcal{Q} \setminus A_Q$. Since $\mathbf{1}_{A_Q}(Q') = \mathbf{1}_{A_{Q'}}(Q)$, for any $Q \in \mathbb{P}$ and $Q' \in \mathcal{Q}$, this gives

$$\sum_{Q \in \mathbb{P}} \Big(\sum_{Q' \in A_Q} \langle \xi_{Q'} \rangle^s \|\psi_{Q'}(D)f\|_{L_p} \Big)^q \leq d_A \sum_{Q' \in \mathbb{P}} \langle \xi_{Q'} \rangle^{qs} \|\psi_{Q'}(D)f\|^q_{L_p}.$$

Likewise, for $q=\infty$,

$$\begin{aligned}\sup_{Q\in\mathbb{P}}\sum_{Q'\in A_Q}\langle\xi_{Q'}\rangle^s\|\psi_{Q'}(D)f\|_{L_p} &\leq d_A\sup_{Q\in\mathbb{P}}\sup_{Q'\in A_Q}\langle\xi_{Q'}\rangle^s\|\psi_{Q'}(D)f\|_{L_p}\\ &= d_A\sup_{Q'\in\mathcal{Q}}\langle\xi_{Q'}\rangle^s\|\psi_{Q'}(D)f\|_{L_p}.\end{aligned}$$

For $1<q<\infty$, Hölder's inequality with $1=1/q+1/q'$ implies

$$\begin{aligned}&\sum_{Q\in\mathbb{P}}\Big(\sum_{Q'\in A_Q}\langle\xi_{Q'}\rangle^s\|\psi_{Q'}(D)f\|_{L_p}\Big)^q\\ &\qquad\leq\sum_{Q\in\mathbb{P}}\Big(\sum_{Q'\in\mathcal{Q}}(\mathbf{1}_{A_Q}(Q'))^{q'}\Big)^{q/q'}\Big(\sum_{Q\in\mathbb{P}}\big(\mathbf{1}_{A_Q}(Q')\langle\xi_{Q'}\rangle^s\|\psi_{Q'}(D)f\|_{L_p}\big)^q\Big)\\ &\qquad\leq d_A^{q-1}\sum_{Q\in\mathbb{P}}\sum_{Q'\in\mathcal{Q}}\mathbf{1}_{A_Q}(Q')\big(\langle\xi_{Q'}\rangle^s\|\psi_{Q'}(D)f\|_{L_p}\big)^q\\ &\qquad\leq d_A\sum_{Q'\in\mathcal{Q}}\langle\xi_{Q'}\rangle^{qs}\|\psi_{Q'}(D)f\|_{L_p}^q.\end{aligned}$$

The lower bound in (4.2) now follows by combining the above estimates with the inequality (A.1). The upper bound can be proved in a similar fashion. □

We conclude by proving Lemma 4.6.

Proof of Lemma 4.6. From (2.9) we notice that there is a constant K such that for every I, $\|g_I\|_{L_p(\mathbb{R})}\leq K$. This estimate, together with the representation (2.8), imply that

$$\sup_{x\in\mathbb{R}^2}\sum_{k\in\mathbb{N}_0^2}|\omega_k^Q(x)|^p\leq C|Q|^{\frac{p}{2}}\quad\text{and}\quad\sup_{k\in\mathbb{N}_0^2}\|\omega_k^Q\|_{L_p}^p\leq C'|Q|^{\frac{p}{2}-1}.\tag{A.2}$$

Suppose $p\leq 1$. Let c be the center of the rectangle Q. Notice that $\operatorname{supp}\big(\mathcal{F}(P_Qf\omega_k^Q)\big)\subset\tilde{Q}$, where $\tilde{Q}$ is the rectangle with center c and twice the side lengths compared to Q. We have (see e.g. [26, p. 18])

$$\begin{aligned}\sum_{k\in\mathbb{N}_0^2}|\langle f,\omega_k^Q\rangle|^p=\sum_{k\in\mathbb{N}_0^2}|\langle P_Qf,\omega_k^Q\rangle|^p&\leq\sum_{k\in\mathbb{N}_0^2}\Big(\int_{\mathbb{R}^2}|e^{ix\cdot c}P_Qf(x)\omega_k^Q(x)|\,dx\Big)^p\\ &\leq C|Q|^{1-p}\sum_{k\in\mathbb{N}_0^2}\int_{\mathbb{R}^2}|P_Qf(x)|^p|\omega_k^Q(x)|^p\,dx\leq C'|Q|^{1-\frac{p}{2}}\|P_Qf\|_{L_p(\mathbb{R}^2)}^p.\end{aligned}$$

Likewise,

$$\|P_Qf\|_{L_p}^p\leq\sum_{n\in\mathbb{N}_0^2}|\langle f,w_{n,Q}\rangle|^p\|w_{n,Q}\|_{L_p}^p\leq C|Q|^{\frac{p}{2}-1}\sum_{n\in\mathbb{N}_0^2}|\langle f,w_{n,Q}\rangle|^p.$$

For $1<p<\infty$ the lemma follows using the two estimates (A.2) for $p=1$, together with Hölder's inequality (see e.g. [23, §2.5]). The case $p=\infty$ is trivial. □

REFERENCES

[1] P. Auscher, G. Weiss, and M. V. Wickerhauser. Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets. In *Wavelets*, volume 2 of *Wavelet Anal. Appl.*, pages 237–256. Academic Press, Boston, MA, 1992.

[2] L. Borup and M. Nielsen. Approximation with brushlet systems. *J. Approx. Theory*, 123(1):25–51, 2003.

[3] L. Borup and M. Nielsen. Banach frames for multivariate α-modulation spaces. *J. Math. Anal. Appl.*, 321(2):880–895, 2006.

[4] L. Borup and M. Nielsen. Boundedness for pseudodifferential operators on multivariate α-modulation spaces. *Ark. Mat.*, 44(2):241–259, 2006.

[5] L. Borup and M. Nielsen. Nonlinear approximation in α-modulation spaces. *Math. Nachr.*, 279(1-2):101–120, 2006.

[6] L. Borup and M. Nielsen. Frame decomposition of decomposition spaces. *J. Fourier Anal. Appl.*, 13(1):39–70, 2007.

[7] A. Córdoba and C. Fefferman. Wave packets and Fourier integral operators. *Comm. Partial Differential Equations*, 3(11):979–1005, 1978.

[8] S. Dahlke, M. Fornasier, H. Rauhut, G. Steidl, and G. Teschke. Generalized coorbit theory, banach frames, and the relation to alpha-modulation spaces. *Proc. London Math. Soc. (to appear)*, 2007.

[9] R. A. DeVore, B. Jawerth, and B. J. Lucier. Image compression through wavelet transform coding. *IEEE Trans. Inform. Theory*, 38(2, part 2):719–746, 1992.

[10] R. A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992.

[11] R. A. DeVore and G. G. Lorentz. *Constructive approximation*, volume 303 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.

[12] H. G. Feichtinger. Banach spaces of distributions defined by decomposition methods. II. *Math. Nachr.*, 132:207–237, 1987.

[13] H. G. Feichtinger and P. Gröbner. Banach spaces of distributions defined by decomposition methods. I. *Math. Nachr.*, 123:97–120, 1985.

[14] M. Fornasier. Banach frames for α-modulation spaces. *Appl. Comput. Harmon. Anal.*, 22(2):157–175, 2007.

[15] G. Garrigós and E. Hernández. Sharp Jackson and Bernstein inequalities for N-term approximation in sequence spaces with applications. *Indiana Univ. Math. J.*, 53(6):1739–1762, 2004.

[16] R. Gribonval and M. Nielsen. Some remarks on non-linear approximation with Schauder bases. *East J. Approx.*, 7(3):267–285, 2001.

[17] P. Gröbner. *Banachräume glatter Funktionen und Zerlegungsmethoden*. PhD thesis, University of Vienna, 1992.

[18] E. Hernández and G. Weiss. *A first course on wavelets*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1996. With a foreword by Yves Meyer.

[19] G. Kerkyacharian and D. Picard. Entropy, universal coding, approximation, and bases properties. *Constr. Approx.*, 20(1):1–37, 2004.

[20] S. V. Konyagin and V. N. Temlyakov. A remark on greedy approximation in Banach spaces. *East J. Approx.*, 5(3):365–379, 1999.

[21] E. Laeng. Une base orthonormale de $L^2(\mathbf{R})$ dont les éléments sont bien localisés dans l'espace de phase et leurs supports adaptés à toute partition symétrique de l'espace des fréquences. *C. R. Acad. Sci. Paris Sér. I Math.*, 311(11):677–680, 1990.

[22] F. G. Meyer and R. R. Coifman. Brushlets: a tool for directional image analysis and image compression. *Appl. Comput. Harmon. Anal.*, 4(2):147–187, 1997.

[23] Y. Meyer. *Wavelets and operators*, volume 37 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1992.

[24] B. Nazaret and M. Holschneider. An interpolation family between Gabor and wavelet transformations: application to differential calculus and construction of anisotropic Banach spaces. In *Nonlinear hyperbolic equations, spectral theory, and wavelet transformations*, volume 145 of *Oper. Theory Adv. Appl.*, pages 363–394. Birkhäuser, Basel, 2003.

[25] L. Päivärinta and E. Somersalo. A generalization of the Calderón-Vaillancourt theorem to L^p and h^p. *Math. Nachr.*, 138:145–156, 1988.
[26] H. Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983.

DEPARTMENT OF MATHEMATICAL SCIENCES, AALBORG UNIVERSITY, FREDRIK BAJERSVEJ 7G, DK-9220 AALBORG EAST, DENMARK
E-mail address: `mnielsen@math.aau.dk`