CAPlets: wavelet representations without wavelets

Youngmi Hur
Department of Mathematics
University of Wisconsin-Madison
480 Lincoln Dr.
Madison, WI 53706
hur@math.wisc.edu

Amos Ron
Computer Sciences Department
University of Wisconsin-Madison
1210 West Dayton
Madison, WI 53706
amos@cs.wisc.edu

ABSTRACT

MultiResolution (MR) is among the most effective and the most popular approaches for data representation. In that approach, the given data are organized into a sequence of resolution layers, and then the "difference" between each two consecutive layers is recorded in terms of *detail coefficients*. Wavelet decomposition is the best known representation methodology in the MR category. The major reason for the popularity of wavelet decompositions is their implementation and inversion by a fast algorithm, the so-called fast wavelet transform (FWT). Another central reason for the success of wavelets is that the wavelet coefficients capture very accurately the *smoothness class* of the function hidden behind the data. This is essential for the understanding of the performance of key wavelet-based algorithms in compression, in denoising, and in other applications. On the downside, constructing wavelets with good space-frequency localization properties becomes involved as the spatial dimension grows.

An alternative to the sometime-hard-to-construct wavelet representations is the always-easy-to-construct (and slightly older) non-orthogonal pyramidal algorithms. Similar to wavelets, the (linear, regular, isotropic) pyramidal representations are based on some method for linear coarsening (by a decomposition filter) of their data, and a complementary method for linear prediction (by a prediction filter) of the original data from the coarsened one. The first step creates the resolution layers and the second allows for trivial extractions of suitable detail coefficients. The decomposition and reconstruction algorithms in the pyramidal approach are as fast as those of wavelets. In contrast with orthonormal wavelets, the representation is *redundant*, viz. the total number of detail coefficients exceeds the original size of the data: denoting by s the ratio between the size of the data at two consecutive resolution layers, the "redundancy ratio" in the pyramidal representation is $\frac{s}{s-1}$.

In this paper, we introduce and study a general class of pyramidal representations that we refer to as Compression-Alignment-Prediction (CAP) representations. The CAP representation is based on the selection of three filters: the low-pass decomposition filter, the low-pass prediction filter, and the full-pass alignment filter. Like previous pyramidal algorithms, CAP are implemented by a simple, fast, wavelet-like decomposition and a trivial reconstruction. The primary goal of this paper is to establish the precise way in which the CAP representations encode the smoothness class of the underlying function. Remarkably, the CAP coefficients provide the same characterizations of Triebel-Lizorkin spaces and Besov spaces as the wavelet coefficients do, provided that the three CAP filters satisfy certain requirements. This means, at least in principle, that the performance of CAP-based algorithms should be similar to their wavelet counterparts, despite of the fact that, when compared with wavelets, it is much easier to develop CAP representations with "customized" or "optimal" properties. Moreover, upon assuming the prediction filter to be *interpolatory*, we extract from the CAP representation a sister CAMP representation ("M" for "modified"). Those CAMP representations strike a phenomenal balance between performance (viz., smoothness characterization) and space localization.

Our analysis of the CAP representations is based on the existing theory of framelet (redundant wavelet) representations.

AMS (MOS) Subject Classifications: Primary 42C15, Secondary 42C30

Key Words: frames, framelets, wavelets, oblique extension principle, pyramidal representations, Besov spaces, Triebel-Lizorkin spaces, CAP representations, CAPlets, CAMP representations, CAMPlets.

This work was supported by the US National Science Foundation under Grants DMS-9872890, DBI-9983114, and ANI-0085984, and by the National Institute of General Medical Sciences under Grant NIH-1-R01-GM072000-01.

CAPlets: wavelet representations without wavelets

YOUNGMI HUR AND AMOS RON

1. Introduction

1.1. CAP representations

We analyse in this paper hierarchical representations of mesh functions that are defined on n-dimensional regular grids. As is quite customary, we assume the regular grid to be the multi-integer lattice $\mathbb{Z}^n$.

The input for our representation is a sequence $y := y_0 : \mathbb{Z}^n \to \mathbb{C}$, which is considered to be the "data at full resolution". The representation is then computed with the aid of three filters. The first, $h_c : \mathbb{Z}^n \to \mathbb{C}$, is a low-pass filter that is used for coarsification (aka decomposition or compression). The second, $h_r : \mathbb{Z}^n \to \mathbb{C}$, is a low-pass filter that is used for prediction (aka subdivision). The third, $h_a : \mathbb{Z}^n \to \mathbb{C}$, is a (full-pass) filter that is used to enhance the performance of the representation algorithm and is referred to as the alignment filter. Throughout the paper, we assume that the filters h_c, h_r, h_a are all *finitely supported.* The process also employs the downsampling and upsampling operators, that are defined, respectively, in the dyadic dilation case, as follows:

$$y_{\downarrow}(k) := y(2k), \quad k \in \mathbb{Z}^n,$$

and

$$y_{\uparrow}(k) := \begin{cases} y(\frac{k}{2}), & k \in 2\mathbb{Z}^n, \\ 0, & k \in \mathbb{Z}^n \backslash (2\mathbb{Z}^n). \end{cases}$$

Now, with $*$ the discrete convolution (between finitely supported sequences defined on $\mathbb{Z}^n$), and $j \in \mathbb{Z}$, the algorithm first coarsifies a given dataset $y_j : \mathbb{Z}^n \to \mathbb{C}$:

$$y_j \mapsto y_{j-1} := Cy_j := (h_c * y_j)_{\downarrow},$$

then re-aligns y_{j-1} by convolving it with h_a:

$$A : y_{j-1} \mapsto Ay_{j-1} := h_a * y_{j-1},$$

and then uses Ay_{j-1} in order to predict Ay_j:

$$Ay_{j-1} \mapsto PAy_{j-1} := 2^n(h_r * (Ay_{j-1})_{\uparrow}).$$

It then stores the "detail coefficients"

$$d_j := (A - PAC)y_j,$$

and reiterates, now with y_{j-1} as the input dataset. The representation

$$y_0 \to y_{-1} \to y_{-2} \to \cdots$$

belongs to the class of *hierarchical representations* (of y_0), while the detail sequences $(d_j)_{j=0}^{-\infty}$ fall into the category of *pyramidal representations.* The smaller j, the coarser are the "details" recorded by d_j. We refer to the particular pyramidal representation that was described above as regular, linear *Compression-Alignment-Prediction representation*, which we abbreviate as *CAP representation.* The *CAP coefficients* are the detail ones $(d_j)_j$. Note that the representation is *redundant*: if y_0 contains N non-zero coefficients, then the total number of non-zero coefficients in $(d_j)_j$ is about $\frac{2^n}{2^n-1}N$ (assuming that the size of the filters h_c, h_a, h_r is negligible compared to N).

This CAP algorithm is basic, simple and is used (for the choice $A := \mathrm{I}$) in a variety of applications. For example, the now-classical *pyramidal algorithm* of P. J. Burt and E. H. Adelson [BA] is a special case of the CAP representation corresponding to $n = 2$, $h_a = \delta$, $h_r = h_c(-\cdot)$ with h_c chosen from some specific class of low-pass filters. Here δ denotes the *Dirac sequence*, i.e. $\delta(0) = 1$ and $\delta(k) = 0$, for all $k \in \mathbb{Z}^n\backslash 0$. Burt and Adelson refer to the corresponding $(y_j)_{j=0}^{-\infty}$ as the *Gaussian pyramid* and to $(d_j)_{j=0}^{-\infty}$ as the *Laplacian pyramid.* Originally devised for progressive image transmission and image compression applications, the pyramidal algorithm has become increasingly popular in image processing applications such as video-conferencing [TP] and synthesizing texture images [HB].

The CAP representation is also a close relative of the filter bank representation. The theoretical foundation of the latter is rooted in *wavelet theory.** We note that, however, there are notable differences between the CAP representation and the filter bank representation, and that these prevent a direct application of wavelet theory to the CAP setup. Let us elaborate on this point.

The wavelet methodology can be viewed as the addition of another layer to the CAP one: it introduces additional "detail" (high-pass) filters $(h_i)_{i=1}^{L}$, and complementary reconstruction filters $(h_i^{\rm dual})_{i=1}^{L}$ in a way that the detail map $A - PAC$ is decomposed into

$$A - PAC = \sum_{i=1}^{L} R_i W_i. \tag{1.1}$$

Here, $W_i : y \mapsto (h_i * y)_{\downarrow}$, and $R_i : y \mapsto 2^n(h_i^{\rm dual} * (y_{\uparrow}))$. Instead of storing the details at level j as $(A - PAC)y_j$, the details are stored as the "wavelet coefficients" $w_{i,j-1} := W_i y_j$. Here, L must be $\geq 2^n - 1$, since otherwise $A - PAC$ cannot be decomposed as above. Moreover, if $L = 2^n - 1$, significant restrictions must be imposed on the original three filters h_c, h_a, h_r for a decomposition as above to exist. The case $L = 2^n - 1$ includes the *orthogonal pyramidal algorithm* [Ma], [Me1] and the *biorthogonal pyramidal algorithm* [CDF]. A comparison between Burt and Adelson's pyramidal algorithm and the (bi)orthogonal pyramidal algorithms in the context of image processing is given in [DN], [JMR].

Trivially, the detail coefficients d_j of the CAP method are connected to the wavelet coefficients $(w_{i,j-1})_{i=1}^{L}$ via the one-step reconstruction formula

$$d_j = \sum_{i=1}^{L} R_i w_{i,j-1}.$$

One might then wonder whether the recording of the details in terms of wavelet coefficients has any advantage over the much simpler recording of the details of the CAP methods. After all, finding a decomposition of $A - PAC$ as in (1.1) becomes involved in high dimensions, especially for high-performance (hence long) filters h_c, h_r. One of the main reasons for preferring the wavelet representation over the CAP representation is that wavelet theory provides a very powerful and useful interpretation of the wavelet coefficients: it describes the above discrete process in terms of functions defined on $\mathbb{R}^n$, making the wavelet coefficients the coefficients in a suitable series representation of an underlying function. Wavelet theory then shows that the wavelet coefficients detect, in a very subtle and accurate way, the smoothness class of the function that is being represented. These function space characterizations play a crucial role in the analysis of the performance of wavelet-based algorithms, such as compression or denoising via the thresholding of the wavelet coefficients.

---

* Throughout the paper, the notion of a *wavelet* or a *wavelet system* refers to any collection of dilated shifts of a finite set of functions; i.e., the definition does not assume the orthonormality of the wavelet system.

The function spaces that we were alluding to above are known as Triebel-Lizorkin (TL) spaces and Besov spaces (cf. section 3.1). More familiar smoothness spaces are obtained as special cases. We mention in passing the $L_p(\mathbb{R}^n)$ $(1 < p < \infty)$ spaces, the Sobolev spaces $W_p^s(\mathbb{R}^n)$ $(s > 0, 1 < p < \infty)$, and the Hölder spaces. Characterizations of TL and Besov spaces using orthonormal wavelets were studied by Y. Meyer [Me1], and E. Hernández and G. Weiss [HW]. Analogous characterizations that use special types of band-limited wavelets (cf. condition (3.1)) were studied by M. Frazier and B. Jawerth [FJ1], [FJ2]. More recently, G. Kyriazis [K] established characterizations of these spaces that are based on *bi-framelet* systems. Independently, L. Borup, R. Gribonval, and M. Nielsen, [BGN1], [BGN2] derived bi-framelet characterization of $L_p(\mathbb{R}^n)$ and tight-framelet characterization of $W_p^s(\mathbb{R}^n)$. We will provide further details on these characterizations in the body of this article. It is important to note already at this stage that for the above-mentioned smoothness characterizations to be valid, one needs to impose some side conditions on the various filters involved. The side conditions are becoming more demanding with the increase of the smoothness level that is being studied.

We develop in this article a solid theoretical foundation to the CAP algorithm. The theory is then used in order to identify the properties of the filters h_c, h_a, h_r that are pertinent to the performance of the CAP representation. Specifically, we show that the CAP coefficients encode the smoothness class of the underlying function in a way that is completely analogous to the way this is done by wavelet coefficients. At the same time, the side conditions that need to be imposed are far simpler (after all, the CAP method is based only on two low-pass filters and one alignment filter). Moreover, we will show that, frequently, the wavelet decomposition of the detail operator $A-PAC$ is performance-degrading: the wavelet coefficients may fail to provide characterizations of smoothness classes that the CAP coefficients do. As a matter of fact, we anticipated this performance-degrading phenomenon, and that anticipation was the main stimulus behind the current endeavor.

At this point of the discussion, our reader might ask the suggestive question: "wavelets why?". Why bother with the enormous variety of, and the vast literature dealing with, the sometime-very-messy wavelet constructions (viz. the decompositions of $A - PAC$) if the major accomplishment here is a possible performance-degradation? The surprising answer is that, with our performance analysis of the CAP method, one is left with very little incentive to deal with wavelets. The main remaining advantage of wavelets here is that they allow non-redundancy: if one sets their number at the minimum $L = 2^n - 1$, the discrete wavelet representation becomes non-redundant (and the underlying wavelet systems form, almost always, bases). Another possible advantage of wavelets is the ability to employ *high* level of redundancy, something that may be important for applications that deal with feature detection or denoising. A signal analysis expert who reads the presentation so far may also claim that the detail coefficients of the CAP representation are likely to have worse time/space localization properties compared to their wavelet counterparts. While this might seem correct at first glance, the truth is that the issue of localness in time/space is much more subtle and should not be credited to any method. In particular, there are several ways to improve the space localization of the CAP methodology. We discuss one approach in this direction, the CAMP representation. The balance between space localization and performance of some of our CAMP representation represents a significant improvement over the corresponding balance offered by the currently available wavelets.

1.2. Toward setting new benchmarks for simultaneous space-scale localization

This section assumes some familiarity with wavelet representations. It might be skipped without loss of continuity.

Constructing a good MultiResolution (MR) representation requires careful balancing between the time/space localization of the system on the one hand, and its frequency/scale localization on the other hand. In the case of wavelet representations, there are several, closely related, ways to quantify the space localness of the system. One way is to compute the average number N_1 of the non-zero coefficients in the filters (low-pass together with high-pass) involved. This, in general, is connected to the computational effort that is required to compute the representation. Another way is to compute the average size (area/volume) N_2 of the support of the mother wavelets. This represents the amount of "overlapping in space" since it captures, for any fixed scale, the number of times a generic point in space is "examined" by the wavelets. Finally, assuming that the supports of the wavelets are not convex, one may, conservatively, compute the average size N_3 of the *convex hull* of the wavelet supports. This is directly connected to space localization but may sometimes result in overshooting, for example, when the support of the mother wavelet lies inside the union of two narrow strips.

There are also several ways to define the notion of frequency localization in the context of wavelets. The most demanding one is to require the wavelet coefficients to characterize function smoothness up to smoothness order $s > 0$, with s, thus, being the localization parameter. It is best to think about s as the *performance grade* of the wavelet system.

The CAP representation can be recast as a redundant wavelet representation, and in this way it may also be associated it with "space localness" parameters N_1, N_2, N_3 that are defined essentially as above. However, in general it is not very local, since it is based in the successive execution of three filters. Its prowess lies in its universality, as well as in the ease at which one can construct high-performance systems, but not in the resulting space localization. Under additional assumptions on the CAP filters (e.g., that the prediction filter h_r is interpolatory), we derive from the given CAP a sister representation, referred to as CAMP ("M" for "modified"), with the same "performance grade" s, but with better space localization. Some of the CAMP constructions provide a balance between space localization and performance that, when compared to the localizations offered by mainstream wavelet constructions, may be referred to as nothing less than *stunning*.

We discuss in this subsection the space localization of one class of CAMP representations. The representation is available at any spatial dimension n, and for any scalar, integer, dilation Λ (i.e., not only for the dyadic one listed in the previous subsection). The performance grade of this CAMP is $s = 1$, which puts it on par with the biorthogonal wavelet system 3/5 (provided that the 3-tap spline filter is used for reconstruction; otherwise the performance of 3/5 is lower), and its higher dimensional tensor products. However, since we are not aware of an analog of the 3/5 system for non-dyadic dilations, we compare our CAMP representation with the classical Haar system. The performance grade of the Haar system is abysmal, i.e., $s = 0$. On the other hand, it is considered an epitome of the time/space localization notion. Its localization parameters (as defined above), for integer dilation Λ and in n-dimensions, are as follows:

$$N_1 = \Lambda^n, \quad N_2 = N_3 = 1.$$

While our CAMP representation provides significant performance improvement over Haar, it does so at no cost: its localization numbers are anywhere between being comparable up to being far superior to those of Haar. We demonstrate this point for $n = 2$ (and Λ arbitrary), and then for $\Lambda = 2$ (and n arbitrary).

For $n = 2$, our CAMP representation is equivalent to a wavelet/framelet representation with $\Lambda^2 + 1$ filters (one is a low-pass filter, and the others are high-pass). The redundancy of the representation can be controlled by choosing large Λ. The performance is $s = 1$, independently of Λ. Now, the average size of the CAMP filters is

$$N_1 = \frac{(\Lambda-1)(7\Lambda+1)+2}{\Lambda^2+1} \leq 7,$$

which yields for $\Lambda = 2,3,4,5$ the values $N_1 = 3.4, 4.6, 5.2, 5.6$, respectively (compared to $4, 9, 16, 25$ of Haar). The average area of the support of the CAMPlets satisfies

$$N_2 \le \frac{3(5\Lambda - 3)}{\Lambda^3} = O(\Lambda^{-2}).$$

For $\Lambda = 2,3,4,5$, we then have $N_2 \le 2.6, 1.3, .8, .5$, respectively. We do not provide similar estimates for N_3 but note that its limit (as $\Lambda \to \infty$) is $\le .5$.

For $\Lambda = 2$ (and without any restriction on n), the Haar system provides localization numbers

$$N_1 = 2^n, \quad N_2 = N_3 = 1.$$

In contrast, our CAMP construction delivers

$$N_1 = \frac{5 \cdot 2^n - 3}{2^n + 1} < 5,$$

which for $n = 2,3,4$ reads as $N_1 = 3.4, 4.1, 4.5$ (compared with $N_1 = 4, 8, 16$ for Haar). Also,

$$N_2 \le \frac{n+1}{2^{n-2}}, \qquad \text{and} \quad N_3 = \frac{(n+1)(n+4-2^{1-n})}{2^{n+1}},$$

which means that $N_3 = 2.06, 1.69, 1.23, .84$, for $n = 2,3,4,5$. The fact that $N_2 = O(n2^{-n})$ implies that "each point in space is visited $O(n)$ times by the wavelets at a given scale". It is plausible that this is the best possible asymptotics for performance grade $s = 1$ representations, since n is the dimension of linear polynomials in n variables.

For $\Lambda = 2$, we could also recall the localization parameters of the tensor product 3/5 system. For spatial dimension n, those numbers are $N_1 = 4^n$, $N_2 = N_3 = (7^n - 4^n)/(2^n - 1)$. Thus, while our CAMP and the 3/5 deliver the same performance (at least in theory), they differ dramatically in space localization: for $n = 3$, CAMP is about 15-25 times more local than the 3/5 one!

Our CAMP thus sets new standards for time/space localization. In retrospect, none of the standard wavelet constructions have any hope to be very local in space, unless both Λ and n are small: for large Λ and/or large n, such constructs employ at a given scale $\Lambda^n - 1$ mother wavelets that overlap one on top of the other. Ideally, we would like to alter the construction so that the mother wavelets will have essentially disjoint supports. CAMP realizes that dream: its low-pass filter is sizable, and so is the first of its high-pass filters. All the other high-pass filters are tiny: 3-tap ones for $\Lambda = 2$ (and arbitrary n), and 3- or 4-tap ones for $n = 2$ (and arbitrary Λ), which results at mother wavelets whose support is comparable in size to the refinable function at the *fine* scale. Existing wavelet constructions fail to compete with our CAMP since they create mother wavelets whose support size is comparable to the support of the refinable function at *coarse* scale.

We note that the above class of CAMP constructions is based (indirectly) on piecewise-linear splines: the refinable function of the compression filter as well as of the prediction filter is a box spline with $n + 1$ simple directions in $\mathbb{R}^n$. For more details, see §5.2.

Higher performance CAMP constructions have similar characteristics: they employ one mother wavelet with large support, and with all the other mother wavelets having much smaller supports. The size of the latter supports depends on the ratio between the number of non-zero coefficients in h_r and the determinant of the dilation process viz., Λ^n. Thus, our CAMP reduces the problem of constructing high-performance representations with good localization numbers to the following problem:

Goal. *For a given dilation process Λ, find an interpolatory function g, refinable with respect to Λ such that*

(i) The Hölder smoothness of g is as high as possible.
(ii) The ratio between the number of non-zero coefficients in the filter h_r of g and $\det \Lambda$ is as small as possible. □

It is beyond the scope of the present article to provide specific constructions that serve the above goal. We limit ourselves to the general development of the theory that pinpoints the performance of CAP and CAMP representations.

1.3. Notations

Throughout the paper, we use the standard multi-index notation. In particular, for $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$ and $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}_0^n$ ($\mathbb{N}_0 := \mathbb{N} \cup \{0\}$), we let $|t| := \sqrt{t_1^2 + \cdots + t_n^2}$, $t^\beta := t_1^{\beta_1} \cdots t_n^{\beta_n}$, $|\beta| := \beta_1 + \cdots + \beta_n$ and $(\cdot)^{(\beta)} := \partial^{|\beta|}(\cdot)/\partial t_1^{\beta_1} \cdots \partial t_n^{\beta_n}$. The inner product of two vectors t, x in $\mathbb{R}^n$ is denoted by $t \cdot x$. We use the following normalization of the Fourier transform (for, e.g., $f \in L_1(\mathbb{R}^n)$):

$$\widehat{f}(\omega) := \int_{\mathbb{R}^n} f(t) e^{i\omega \cdot t} dt.$$

We denote by $\mathcal{S}(\mathbb{R}^n)$ the Schwartz space of test functions, and by $\mathcal{S}'(\mathbb{R}^n)$ its dual, the space of tempered distributions. Given a function space whose elements are defined on $\mathbb{R}^n$, we sometimes omit the domain $\mathbb{R}^n$ in our notation. Also we denote by $\mathcal{S}'/\mathcal{P}$ the space of equivalence classes of (tempered) distributions modulo polynomials. For any $f, g \in L_2$, we define

$$\langle f, g \rangle := \int_{\mathbb{R}^n} f(t) \overline{g(t)}\, dt.$$

For any $f \in \mathcal{S}'$ and $g \in \mathcal{S}$, we define $\langle f, g \rangle := f(\overline{g})$ with the usual extensions, by means of duality, to the various subspaces of $\mathcal{S}'$.

We let $D : f \mapsto 2^{n/2} f(2\cdot)$ be the unitary dyadic dilation operator, and let $E^t : f \mapsto f(\cdot - t)$ for $t \in \mathbb{R}^n$, be the translation operator. Given any f defined on $\mathbb{R}^n$, we use throughout this paper the notation

$$f_{j,k} := D^j E^k f, \quad j \in \mathbb{Z},\ k \in \mathbb{Z}^n.$$

We denote by χ the characteristic function of the unit cube $[0,1]^n$ and for $j \in \mathbb{Z}, k \in \mathbb{Z}^n$, define $\chi_{\infty,j,k} := \chi(2^j \cdot -k)$.

Throughout the paper, c stands for a generic constant that may change with every occurrence. We use the notation $a \lesssim b$ to mean that there is a constant $c > 0$ such that $a \leq cb$. We also use the notation $a \approx b$ to denote two quantities that satisfy $c_1 a \leq b \leq c_2 a$, for some positive constants. The specific dependence of the constants c_1, c_2 on the problem's parameters is explained in the text, whenever such an explanation is required.

1.4. Wavelet and framelet representations

During the last fifteen years, wavelets became a powerful tool in a variety of areas of Science and Engineering. One of the main reasons for this development is the existence of fast algorithms (FWT) for wavelet-based decomposition and reconstruction of data. Another major reason is the associated mathematical and statistical theories. A main pillar in these theories is the ability to characterize the smoothness class of the underlying function in terms of its "wavelet coefficients".

A (dyadic) wavelet system $X(\Psi)$ in $\mathbb{R}^n$ is the collection of dilated shifts of a finite set of *mother wavelets* $\Psi \subset L_2(\mathbb{R}^n)$:

$$X(\Psi) := \{\psi_{j,k} :\ \psi \in \Psi,\ j \in \mathbb{Z},\ k \in \mathbb{Z}^n\}.$$

We say that $X(\Psi)$ has m *vanishing moments* if the Fourier transform $\widehat{\psi}$ of each mother wavelet $\psi \in \Psi$ has a zero of order m at the origin: $\widehat{\psi} = O(|\cdot|^m)$. The *wavelet coefficients* $T^* f$ of a given function $f : \mathbb{R}^n \to \mathbb{C}$ are defined as

$$T^* f := T^*_{X(\Psi)} f := \{\langle f, x \rangle : x \in X(\Psi)\}.$$

One calls $X(\Psi)$ a *frame* (for $L_2(\mathbb{R}^n)$) if the $\ell_2(X(\Psi))$-norm of T^*f is equivalent to the $L_2(\mathbb{R}^n)$-norm of f, viz.,

$$\sum_{x\in X(\Psi)} |\langle f,x\rangle|^2 \lesssim \|f\|^2_{L_2(\mathbb{R}^n)} \lesssim \sum_{x\in X(\Psi)} |\langle f,x\rangle|^2, \quad \text{for all } f\in L_2(\mathbb{R}^n). \tag{1.2}$$

$X(\Psi)$ is a *Bessel system* if T^* is bounded, i.e., the left-hand side of (1.2) is valid. A complete orthonormal basis $X(\Psi)$ for $L_2(\mathbb{R}^n)$ is called an *orthonormal wavelet.* Clearly, every frame is a Bessel system, and every orthonormal wavelet (or, more generally, a Riesz basis) is a frame. In general, however, a (wavelet) frame needs not to be a basis. This is an important feature which will be fully exploited in the current paper.

Wavelet characterizations of spaces other than $L_2(\mathbb{R}^n)$ are not as simple as their $L_2(\mathbb{R}^n)$-counterpart. For example, the characterization of the Sobolev space $W^s_p(\mathbb{R}^n)$, $1<p<\infty$, $s>0$, roughly goes as follows: for $c_f(j,k) := \left(2^{jn}\sum_{\psi\in\Psi}|\langle f,\psi_{j,k}\rangle|^2\right)^{1/2}$, one defines the *square function*

$$Q_s(c_f)^2 := \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^n} |c_f(j,k)|^2\left(1+2^{2js}\right)\chi_{\infty,j,k}, \tag{1.3}$$

and attempts to prove that the W^s_p-norm of f is equivalent to the L_p-norm of $Q_s(c_f)$. To this end, it is very helpful to find an explicit inverse to $T^*_{X(\Psi)}$. Such an inverse employs the notion of a *dual frame* and goes as follows. First, one defines a map $\Psi\ni\psi\mapsto\psi^{\text{dual}}\in L_2(\mathbb{R}^n)$, and extends it naturally to $X(\Psi)$ (i.e., $(\psi_{j,k})^{\text{dual}} := (\psi^{\text{dual}})_{j,k}$). Assume that $X(\Psi^{\text{dual}})$ is also a frame. The frame $X(\Psi^{\text{dual}})$ is then said to be *dual* to $X(\Psi)$ if one has the perfect reconstruction property:

$$f = T_{X(\Psi^{\text{dual}})}T^*_{X(\Psi)}f := \sum_{x\in X(\Psi)}\langle f,x\rangle x^{\text{dual}}, \quad f\in L_2(\mathbb{R}^n).$$

It is known that the sought-for equivalence between the L_p-norm of the square function $Q_s(c_f)$ and the W^s_p-norm of f is valid whenever the frame $X(\Psi)$ has $m>s$ vanishing moments (and satisfies a minimal smoothness condition), while the dual frame $X(\Psi^{\text{dual}})$ is smooth (e.g., $\Psi^{\text{dual}}\subset C^m(\mathbb{R}^n)$ for $m>s$) and has at least one vanishing moment. The precise statements are given in the body of this article (cf. Theorem 3.5).

Thus, one strives to build wavelet frame systems that have a high number of vanishing moments, and have smooth dual frames. This brings us to the question of how, actually, wavelet systems are constructed. With very few exceptions, wavelet systems are constructed via the tool of MultiResolution Analysis (MRA): one begins with a *refinable* function $\phi\in L_2(\mathbb{R}^n)$, viz., a function whose Fourier transform satisfies a relation

$$\widehat{\phi}(2\cdot) = \tau\widehat{\phi},$$

for some 2π-periodic τ known as the *refinement mask.* One subsequently defines $V_0 := V_0(\phi)\subset L_2(\mathbb{R}^n)$ to be the closed linear span of $(E^k\phi)_{k\in\mathbb{Z}^n}$, and $V_j := V_j(\phi) := D^jV_0$, $j\in\mathbb{Z}$. The mother wavelets $\Psi=\{\psi_1,\ldots,\psi_L\}$ are then (carefully) selected from V_1. This implies [BDR1], [BDR2] that every mother wavelet $\psi_i\in\Psi$ must satisfy a relation of the form

$$\widehat{\psi_i}(2\cdot) = \tau_i\widehat{\phi},$$

for some 2π-periodic τ_i. The dual mother wavelets $\Psi^{\rm dual}$ are constructed similarly, using another refinable function $\phi^{\rm dual}$ with mask $\tau^{\rm dual}$, and corresponding wavelet masks $(\tau_i^{\rm dual})_{i=1}^L$. The two systems are then said to satisfy the Oblique Extension Principle (OEP) [DHRS] if the following conditions are met

$$\Theta(2\cdot)\overline{\tau}\tau^{\rm dual}(\cdot+\nu)+\sum_{i=1}^{L}\overline{\tau}_i\tau_i^{\rm dual}(\cdot+\nu)=\begin{cases}\Theta, & \nu=0,\\ 0, & \nu\in\{0,\pi\}^n\backslash 0.\end{cases} \tag{1.4}$$

Here, Θ is an auxiliary 2π-periodic function whose limit at 0 should equal 1: $\Theta(0)=1$. To complete the OEP conditions, one also assumes that each of the masks involved is bounded, each of the wavelet systems $X(\Psi)$ and $X(\Psi^{\rm dual})$ is Bessel, and that the Fourier transforms of the two refinable functions are continuous at the origin and assume the value 1 there. There is another, more special, extension principle known as the Unitary Extension Principle (UEP) [RS] that corresponds to the case $\Theta=1$ in the OEP. Also, if $\tau^{\rm dual}=\tau$ and $\tau_i^{\rm dual}=\tau_i$ for $i=1,\cdots,L$ in (1.4), the constructed wavelet system is called a *tight framelet*.

It is known that, if the two wavelet systems $X(\Psi)$ and $X(\Psi^{\rm dual})$ satisfy the OEP, and have, each, one vanishing moment, then they form a pair of a wavelet frame and a dual wavelet frame [DHRS]. Each of the so-obtained frames is then known as *framelet*, and the above pair is referred to as a *bi-framelet*. The above construction is particularly attractive in case one chooses the refinable functions to be compactly supported, smooth piecewise-polynomials: B-splines in the univariate case and box splines in higher dimensions. Indeed, the refinement mask of such spline is particularly simple, its smoothness is high compared to the support of the refinement filters, and there are other attractions. Thus, the most attractive constructions are based on refinable functions ϕ, $\phi^{\rm dual}$ that are compactly supported splines.

Two problems that arise in the construction of bi-framelets are relevant to our discussion. First, while there are effective constructions of bi-framelets in the univariate case, the situation becomes quite messy once one deals with higher dimensions. For example, any MRA-based wavelet system in three dimensions requires at least 7 wavelets. In many circumstances (e.g., whenever the two refinable functions are compactly supported splines) the actual number of wavelets is higher, sometimes much higher. And, while the OEP simplifies greatly wavelet constructions, it falls short of being an algorithm to this end: it provides no concrete way for choosing the wavelet masks.

The other problem is deeper, and requires us to recall some elements from the theory of framelets. We mentioned above the high smoothness of the mother wavelets as well as the large number of vanishing moments of the wavelet system as two important features of the system. There is a third parameter, the approximation order of the system: one says that the bi-framelet provides *approximation order* m if for every sufficiently smooth f, as $j\to\infty$,

$$\|f-\sum_{\psi\in\Psi,k\in\mathbb{Z}^n,l<j}\langle f,\psi_{l,k}\rangle\psi_{l,k}^{\rm dual}\|_{L_2(\mathbb{R}^n)}\le c\,\|f\|_{W_2^m(\mathbb{R}^n)}2^{-jm}.$$

For an orthonormal wavelet system, the approximation order coincides with the number of vanishing moments of the system. However, for general framelets, the number of vanishing moments can be as low as half of the approximation order [DHRS]. Now, characterizations of smooth function spaces in terms of wavelet coefficients are saturated at the level of the existing vanishing moments. Our theory in this paper shows that the performance of the CAP coefficients (associated with the same $\phi,\phi^{\rm dual}$) is saturated at the (usually higher) approximation order of the system. This subtle, but critical, understanding was the main stimulus behind the present effort.

1.5. Approximation from principal shift-invariant spaces

Two theories provide the foundation for the CAP methodology. The first, that of framelets, was discussed in the last subsection. The other, approximation from shift-invariant subspaces, is discussed here.

We say that a refinable function $\phi \in L_2(\mathbb{R}^n)$ satisfies the SF (Strang-Fix) conditions of order m if

$$\sum_{\beta \in 2\pi\mathbb{Z}^n\backslash 0} |\widehat{\phi}(\cdot + \beta)|^2 = O(|\cdot|^{2m}), \quad \text{near the origin.} \tag{1.5}$$

One shows that, with τ the refinement mask of ϕ, the SF conditions of order m are implied by the condition

$$\tau \text{ has a zero of order } m \text{ at each } \nu \in \{0, \pi\}^n\backslash 0. \tag{1.6}$$

With a slight abuse of terminology, we refer to (1.6), not (1.5), as the SF conditions of order m in this paper.

Let $\phi_c, \phi_r \in L_2(\mathbb{R}^n)$ be two refinable functions with refinement masks τ_c, τ_r and let Θ be a trigonometric polynomial. Suppose that $\overline{\tau}_c$ is the Fourier series of the compression filter h_c of a given CAP representation, and that τ_r is the Fourier series of the prediction filter h_r of that same CAP. Also suppose that Θ is the Fourier series of the alignment filter h_a of the same CAP. Throughout this paper, we assume the following conditions:

Assumptions 1.7.

(a) The filters h_c, h_r, h_a are finite. This implies (trivially) that τ_c, τ_r, Θ are trigonometric polynomials, and (less trivially) that ϕ_c, ϕ_r are compactly supported. This assumption simplifies much of our presentation, but is not essential: our results are valid under weaker assumptions.

(b) $\widehat{\phi}_c(0) = \widehat{\phi}_r(0) = 1$. This is an essential assumption.

(c) ϕ_c, ϕ_r have some positive Hölder smoothness. In particular, those functions are continuous.

(d) For each of $\phi := \phi_c$ and $\phi := \phi_r$, the SF order defined by (1.5) is the same as the SF order defined by (1.6). This is a mild technical assumption made for pure convenience. It makes some of the *statements* simpler, but does not simplify the actual arguments. In 1D, for example, this assumption is valid once we know that neither of the functions $\sum_{k\in\mathbb{Z}}(-1)^k\phi(\cdot - k)$, $\phi := \phi_c, \phi_r$ is identically 0, *a fortiori* this assumption is valid if the shifts of each of ϕ_c, ϕ_r form a Riesz basis.

(e) The *alignment mask* Θ satisfies $\Theta(0) = 1$ and $|\Theta(\omega)| > 0$, for all $\omega \in \mathbb{T}^n$.

(f) $\Theta - \Theta(2\cdot)\overline{\tau}_c\tau_r$ has a double zero at the origin. We use this assumption only once (in the proof of Lemma 3.15), and even there we could have avoided it with some extra effort. However, CAP systems that do not satisfy this condition are void of any practical value, hence we find it convenient to add this condition to our basic ones. □

We consider approximation schemes of the form

$$G_j : f \mapsto \sum_{k\in\mathbb{Z}^n} \langle f, (\phi_c)_{j,k}\rangle \; (h_a * \phi_r)_{j,k}.$$

Here, $h_a : \mathbb{Z}^n \to \mathbb{C}$ is our finitely supported alignment filter, and

$$h_a * \phi_r := \sum_{l\in\mathbb{Z}^n} h_a(l)\phi_r(\cdot - l).$$

Approximation Theory basics (cf. [JZ1],[JZ2]) tell us then the following:

Result 1.8. *If, for some $m \ge 1$, ϕ_r satisfies the SF conditions of order m, and if $\Theta - \Theta(2\cdot)\tau_r\overline{\tau}_c = O(|\cdot|^m)$ near the origin, then there exists a constant $c > 0$ which is independent of f and j such that, for every $f \in W_2^m(\mathbb{R}^n)$, and for every $j \ge j_0$ (with j_0 any fixed integer)*

$$\|f - G_j f\|_{L_2(\mathbb{R}^n)} \le c\,\|f\|_{W_2^m(\mathbb{R}^n)} 2^{-jm}. \tag{1.9}$$

□

The triangle inequality entails that $(G_j - G_{j-1})f$ satisfies an error bound identical to that in (1.9) (with a different c, though). The refinability of ϕ_r entails that $\operatorname{ran}(G_j - G_{j-1}) \subset V_j(\phi_r) - V_{j-1}(\phi_r) = V_j(\phi_r)$. Thus, one can write

$$2^{jn/2}(G_j - G_{j-1})f = \sum_{k\in\mathbb{Z}^n} d_{j,f}(k)\ (\phi_r)_{j,k}. \tag{1.10}$$

The coefficients $(d_{j,f}(k))_k$ may not be uniquely determined by the above (since we have not assumed $((\phi_r)_{j,k})_k$ to be independent), but one selection of $d_{j,f}$ should be considered canonical: a simple calculation shows that $d_{j,f}$ can be chosen as

$$d_{j,f} = (A - PAC)y_{j,f}, \quad y_{j,f}(k) := 2^{jn/2}\langle f, (\phi_c)_{j,k}\rangle,$$

where

$$Cy := (h_c * y)_\downarrow, \quad Py := 2^n(h_r * (y_\uparrow)), \text{ and } Ay := h_a * y.$$

Moreover, $y_{j-1,f} = Cy_{j,f}$.

Thus, the CAP methodology that was described before calculates iteratively the above $(d_{j,f})$ coefficients. Our goal in this paper is to provide characterizations of functions spaces (made of functions defined on $\mathbb{R}^n$) in terms of the CAP coefficients $(d_{j,f}(k))_{j,k}$. To this end we make the following observation, which is a corollary to Lemma 2.2 (cf. (2.9)):

Observation. *Let $\upsilon \in \{0,1\}^n =: \Gamma$, and define*

$$\widehat{\psi}_\upsilon(2\cdot) := t_\upsilon \widehat{\phi}_c,$$

where

$$t_\upsilon := e_\upsilon\left(\overline{\Theta} - \overline{\Theta}(2\cdot)\tau_c\left(\sum_{\upsilon^*\in\{0,\pi\}^n} e^{i\upsilon\cdot\upsilon^*}\overline{\tau}_r(\cdot + \upsilon^*)\right)\right), \quad e_\upsilon : \omega \mapsto 2^{-n/2}e^{i\upsilon\cdot\omega}.$$

Then, for every $j \in \mathbb{Z}$, $k \in \mathbb{Z}^n$ and $\upsilon \in \Gamma$

$$d_{j,f}(2k - \upsilon) = 2^{jn/2}\langle f, (\psi_\upsilon)_{j-1,k}\rangle.$$

Moreover, the system $X(\Psi)$ with $\Psi := \{\psi_\upsilon\}_{\upsilon\in\Gamma}$ has m vanishing moments where m is the same number used in Result 1.8. □

The observation entails that the CAP coefficients can be identified with $T^*_{X(\Psi)}f$. For this reason, we label the above wavelet system as a *CAPlet system*, and the functions $(\psi_v)_v$ as *mother CAPlets*, or, in short, CAPlets. A natural dual system of this $X(\Psi)$ (cf. (1.10)) is $X(\phi_r)$, which unfortunately, does not cater to our needs: the "wavelet system" $X(\phi_r)$ is not a Bessel system. Rather than going through the hassle of proving directly that the reconstruction operator $T_{X(\phi_r)}$ is bounded on the range of $T^*_{X(\Psi)}$, we will find a new wavelet system $X(\Psi^{\mathrm{dual}})$ which *is* Bessel, so that the pair $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ is a bi-framelet. Note that we have to stick to the refinable function ϕ_c and the wavelet masks $(t_v)_v$: this is the only way to ensure that we use the requisite $X(\Psi)$, and hence that we obtain the correct CAP coefficients. However, there are no restrictions on the dual system: we are free to choose the refinable function there as well as the wavelets. Remarkably, as our next result entails, we can choose $X(\Psi^{\mathrm{dual}})$ in such a way that Ψ^{dual} are as smooth as ϕ_r. In the statement of this result, we use the notation

$$\mathcal{D}^\eta(\mathbb{R}^n).$$

This smoothness class is very similar to the class of functions with Hölder exponent η. For example, condition (c) in Assumptions 1.7 is equivalent to $\phi_c, \phi_r \in \mathcal{D}^0(\mathbb{R}^n)$. Cf. Definition 3.3 for more details. Part (a) of the following theorem is the statement of Theorem 2.11 in §2.2 and part(b) is a corollary to that theorem. The proof of the following theorem is found in §3.3.

Theorem 1.11. *In the above notations, suppose that $\phi_r \in \mathcal{D}^\eta(\mathbb{R}^n)$ for some $\eta \geq 0$. Then there exists a wavelet system $X(\Psi^{\mathrm{dual}})$ associated with a refinable function ϕ^{dual} such that:*

(a) The pair $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ satisfies the OEP, with $\widehat{h_a}$ being the OEP function Θ (cf. (1.4)). In particular, this pair is a bi-framelet.

(b) $\phi^{\mathrm{dual}} \in \mathcal{D}^\eta(\mathbb{R}^n)$. $\square$

1.6. Characterizations of Sobolev spaces via CAP representations

Using the approach detailed in the previous two subsections, we establish in this paper characterizations of Triebel-Lizorkin spaces as well as of Besov spaces in terms of the CAP representations. The main novelty of these characterizations is *not* in the way they are formulated: those formulations are entirely similar to ones known via wavelet coefficients. The main attraction in the results, thus, is the precise formulation of the conditions required of the filters (h_c, h_a, h_r) for the given characterization to hold.

In order to simplify the discussion at this introductory stage, we state here our results for the special case of Sobolev spaces $W_p^s(\mathbb{R}^n)$, $s > 0$, $1 < p < \infty$. The first part of the next theorem follows from Corollary 3.19 in §3.3. The second part of the theorem follows from Proposition 6.17 in §6.4.

Theorem 1.12. *Suppose that $\phi_r \in \mathcal{D}^s(\mathbb{R}^n)$ and $\widehat{h_a} - \widehat{h_a}(2\cdot)\widehat{h_r}\widehat{h_c} = O(|\cdot|^m)$ near the origin, for $m > s$. Then:*

(a) For every $f \in W_p^s(\mathbb{R}^n)$ we have

$$\|f\|_{W_p^s(\mathbb{R}^n)} \lesssim \|Q_s(d_{j,f}(k) : j,k)\|_{L_p(\mathbb{R}^n)} \lesssim \|f\|_{W_p^s(\mathbb{R}^n)},$$

with Q_s the square function, (1.3).

(b) If L_p-norm of the square function $Q_s(d_{j,f}(k) : j,k)$, associated with the CAP coefficients $(d_{j,f})_j$ of a given $f \in L_2(\mathbb{R}^n)$, is finite, then $f \in W_p^s(\mathbb{R}^n)$.

Discussion. The theorem tells us that the refinable function ϕ_c should satisfy only minimal conditions: it should be minimally smooth and its shifts should partition unity as indicated in Assumptions 1.7. In particular, it can be chosen independently of the smoothness parameter s. Thus, for example, in one variable we may always choose the compression filter h_c as the double-averaging one:

$$h_c(k) := \begin{cases} 1 - \frac{|k|}{2}, & |k| \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding refinable function ϕ_c is then the centered hat function (viz. the centerec order-2 B-spline) which satisfies our needs. A similar choice works also in n dimensions (a piecewise-linear box spline based on $n+1$ directions, cf. §5.2).

The underlying smoothness parameter s is reflected in the requirements made of ϕ_r, viz. the prediction filter h_r. The simplest approach here is to choose the refinable function ϕ_r as a sufficiently high-order spline (B-spline in 1D, box spline in higher dimension). *Therefore, we can (always!) choose the filters h_c and h_r to perform purely repeated averages (viz. to be spline filters).*

This leaves us with the last condition

$$\widehat{h_a} - \widehat{h_a}(2\cdot)\widehat{h_r}\widehat{h_c} = \Theta - \Theta(2\cdot)\tau_r\overline{\tau}_c = O(|\cdot|^m).$$

Since this condition replaces the well-known "vanishing moment conditions" from wavelet theory (and also since this condition captures correctly the vanishing moment order of the CAPlets), we refer to the above parameter m as the *vanishing moment order of the CAP representation*, or, in short, the *order of the CAP system.* For symmetric spline filters h_c, h_r and for a trivial choice of h_a ($h_a := \delta$, viz. $\Theta = 1$), one obtains CAP systems of order 2, which are acceptable only when $s < 2$. For higher values of s, one needs CAP systems of higher order, and may employ to this end non-trivial, active, alignment. For example, one can calculate that in one dimension and for arbitrary symmetric spline filters h_c, h_r, a given even CAP order m can be obtained by using a filter h_a with $m-1$ non-zero coefficients.

The use of an alignment filter h_a with $m-1$ non-zero coefficients, say in 1D, extends the support size of the CAPlet filters by $2(m-2)$. This 'double time blurring' is due to the fact that alignment is performed after, and not before, downsampling. As a result, it extends the support interval of the CAPlets by $m-2$. The alternative to using alignment is to make the prediction and/or the compression rules more sophisticated; this is reminiscent of the way one constructs refinable functions with desired properties such as orthonormal shifts. This approach seems to produce shorter "detail filters" hence to avoid the double-blurring associated with alignment. A closer look reveals a different picture: if one replaces the filter h_c, which has l non-zero coefficients by an $(l+m-2)$-filter, then the support interval of the CAPlets extends by $m-2$, i.e., the same extension as in the alignment case. The CAPlets, however, are less smooth in this case.
A more detailed discussion concerning actual CAP systems is contained in §5. See also §3.4. □

Under certain conditions, one can replace a given CAP representation by an associated one which provides the same function space characterization but has better time/space localization. We refer to such variants as *Compression-Alignment-Modified Prediction* (CAMP) representations. The particular CAMP idea that we present in this paper applies to an *interpolatory* prediction filters, as well as to a more general class of prediction filters that we label *pseudo-interpolatory* (cf. Section 2.3 for the definition). Rather than detailing here this important idea to its fullest extent, we discuss a special class of CAMP representations that already illustrates the main theme.

Discussion: an example of CAMP representation. Let $g \in C(\mathbb{R}^n)$ be a real symmetric compactly supported refinable function which is *interpolatory*, that is $g(0) = 1$ and g vanishes everywhere else on $\mathbb{Z}^n$. With τ the refinement mask of g, we choose both h_c and h_r to be the corresponding filter h of g (i.e., $h_c = h_r = h$ and $\widehat{h} = \tau$) and forgo using alignment (i.e., $h_a = \delta$). The CAP representation theorem (Theorem 1.12) can be employed then to show that, for a given $s > 0$ in that theorem, one needs to require that the Hölder smoothness exponent of g is $> s$. Since $A = I$ here, the CAP coefficients here satisfy $d_j = (I - PC)y_j$. Moreover, since h_r is interpolatory, then for $k \in 2\mathbb{Z}^n$ we actually have

$$d_j(k) = (y_j - h * y_j)(k) = y_j(k) - y_{j-1}(\frac{k}{2}).$$

We define the CAMP coefficient $\widetilde{d}_j(k)$, $k \in \mathbb{Z}^n$, as follows:

$$\widetilde{d}_j(k) := \begin{cases} y_j(k) - y_{j-1}(\frac{k}{2}), & k \in 2\mathbb{Z}^n, \\ y_j(k) - 2^n(h * (y_{j\downarrow\uparrow}))(k), & \text{otherwise.} \end{cases}$$

That is, at the even locations the CAMP coefficients are the same as their CAP counterparts, while at the odd locations the CAMP coefficients are based only on trivial compression (i.e., downsampling) compared to the compression operator C employed in CAP.

The *reconstruction process* based on CAMP is simple: given y_{j-1} and $\widetilde{d}_j$, we have

$$y_j(k) = y_{j-1}(\frac{k}{2}) + \widetilde{d}_j(k), \quad k \in 2\mathbb{Z}^n.$$

This recovers the sequence $y_{j\downarrow\uparrow}$. Convolving that sequence with h, multiplying by 2^n, and adding the result to $\widetilde{d}_j$ recover correctly y_j at the non-even points.

Our results show that the above CAMP representation provides the same function space characterizations as its associated CAP (cf. Theorem 3.20).

Now, suppose that we compare the above CAMP representation with a wavelet representation whose low-pass reconstruction filter is h (such construction is expected to be on par with the above CAMP as far as function space characterizations are concerned). The wavelet construction must go through the hassle of finding a dual low-pass filter (for decomposition), and then finding two sets of high-pass filters, each of cardinality $2^n - 1$. And, in order for the wavelet construction to have better time/space localization, its low-pass/high-pass filters should better be shorter than the interpolatory filter h. That seems to be a tall order, perhaps impossible, for standard interpolatory filters. □

This paper is organized as follows. In §2 we introduce CAP representations and CAPlet systems with respect to a general integer dilation. We prove two lemmata (Lemma 2.2 and 2.6), which are used to connect the CAP representation with framelet representations in Theorem 2.11. We also discuss the CAMP modification of the CAP representation, and prove the equivalence of the two representations. In §3 we review and modify the known function space characterizations that are based on framelet systems. We then combine these with the results of §2 to obtain our main findings, that is, the characterization of function spaces in terms of CAP coefficients and CAMP coefficients. In §4 we illustrate the way our results can be used for specific applications, by proving a Jackson inequality for best n-term nonlinear approximation. In §5 we illustrate the main results with the aid of specific examples of 1D, 2D and 3D CAP and CAMP representations. Finally, in the Appendix (§6) we provide proofs for two results that are stated and used in §3.2, and discuss a few subtle, technical, points that are needed in order to make our theory fully rigorous.

2. CAP representations

2.1. The rudiments

Let Λ be an $n \times n$ integer dilation matrix. By definition, this means that Λ is an integer matrix and its spectrum lies outside the closed unit disc. Thus $\Lambda^* := \overline{\Lambda^t} = \Lambda^t$ is the transpose of Λ. *Downsampling* is defined on $\mathbb{C}^{\mathbb{Z}^n}$ by

$$y_{\downarrow}(k) := y(\Lambda k),$$

while *upsampling* is defined by

$$y_{\uparrow}(k) := \begin{cases} y(\Lambda^{-1}k), & k \in \Lambda\mathbb{Z}^n, \\ 0, & \text{otherwise.} \end{cases}$$

Let h_c, h_r and h_a be three sequences in $\mathbb{C}^{\mathbb{Z}^n}$ which we refer to as the *compression filter*, the *prediction filter* and the *alignment filter*, respectively. Then, the *Compression-Alignment-Prediction* (CAP) *representation* with dilation Λ is defined as follows. For a given sequence $y : \mathbb{Z}^n \to \mathbb{C}$, the CAP operators are

$$\begin{aligned} C &: y \mapsto (h_c * y)_{\downarrow}, \\ A &: y \mapsto h_a * y, \\ P &: y \mapsto |\det \Lambda| \, (h_r * (y_{\uparrow})). \end{aligned}$$

We will only deal with finitely supported filters (h_c, h_a, h_r), and for such filters the above operations are trivially well-defined.

Given a sequence $y_j \in \mathbb{C}^{\mathbb{Z}^n}$ (with j some running index), we denote

$$y_{j-1} := Cy_j,$$

use PAy_{j-1} to predict Ay_j, and then record the prediction error in the *detail coefficient sequence of level* j:

$$d_j := (A - PAC)y_j = Ay_j - PAy_{j-1}.$$

We refer to $(d_j)_{j \le 0}$ as the *CAP coefficients of the dataset* y_0. The iterations are illustrated in the following diagram.

$$\begin{tikzcd}
y_0 \arrow[r, "C"] \arrow[d, "A"'] & y_{-1} \arrow[r, "C"] \arrow[d, "A"'] & y_{-2} \arrow[r, "C"] \arrow[d, "A"'] & \cdots \arrow[r, "C"] & y_{-(m-1)} \arrow[r, "C"] \arrow[d, "A"'] & y_{-m} \arrow[d, "A"'] \\
Ay_0 \arrow[d, "I"'] & Ay_{-1} \arrow[d, "I"'] \arrow[dl, "P"'] & Ay_{-2} \arrow[dl, "P"'] & \cdots & Ay_{-(m-1)} \arrow[d, "I"'] & Ay_{-m} \arrow[dl, "P"'] \\
d_0 & d_{-1} & \cdots & \cdots & d_{-(m-1)} &
\end{tikzcd}$$

Note that Ay_0 can be recovered from $y_{-m}, d_{-(m-1)}, \ldots, d_{-1}, d_0$ since $Ay_j = d_j + PAy_{j-1}$, $j = 1-m, \ldots, 0$. We can then recover y_0 with the aid of a single deconvolution step.

The CAP representation is *redundant*. If y_0 contains N non-zero coefficients and the size of each filter is negligible compared to N, then d_0 contains about N non-zero coefficients, while y_{-1} contains approximately $\frac{1}{|\det\Lambda|}N$ non-zero coefficients. Reiterating, we find out, thus, that the complete representation uses approximately $\frac{|\det\Lambda|}{|\det\Lambda|-1}N$ coefficients. Consequently, we assign to the CAP representation a redundancy rate of $\frac{|\det\Lambda|}{|\det\Lambda|-1}$. In the dyadic dilation case (i.e. when $\Lambda = 2I$) this redundancy rate becomes $\frac{2^n}{2^n-1}$.

The analysis of pyramidal representations begins with the assumption that the initial sequence y_0 actually records the inner products between the integer shifts of a refinable function on the one hand, and some other function f (that is being analysed) on the other hand. Initially, we may assume that $f \in L_2(\mathbb{R}^n)$. The specific interpretation of the CAP representation begins with two $L_2(\mathbb{R}^n)$-functions ϕ_c, ϕ_r, that are Λ-*refinable*, i.e., there exist 2π-periodic bounded functions $\tau_c, \tau_r : \mathbb{R}^n \to \mathbb{C}$ such that

$$\widehat{\phi}_c(\Lambda^*\cdot) = \tau_c\widehat{\phi}_c, \quad \widehat{\phi_r}(\Lambda^*\cdot) = \tau_r\widehat{\phi}_r.$$

The above τ_c, τ_r are referred to as the *refinement masks*. In addition, we select another trigonometric polynomial Θ and call it the *alignment mask*. We assume throughout that ϕ_c, ϕ_r and Θ satisfy (a-c) and (e) in Assumptions 1.7.

Next, we define a CAP representation by choosing h_c to be the Fourier coefficients of $\overline{\tau}_c$, h_r to be the Fourier coefficients of τ_r, and h_a to be the Fourier coefficients of Θ. With D the dilation operator and E the translation operator

$$D : g \mapsto |\det\Lambda|^{1/2} g(\Lambda\cdot), \quad E^t : g \mapsto g(\cdot - t), \quad t \in \mathbb{R}^n,$$

we denote as before

$$g_{j,k} := D^j E^k g, \quad g \in L_2(\mathbb{R}^n),\ k \in \mathbb{Z}^n,\ j \in \mathbb{Z}.$$

Now, the CAP representation of $f \in L_2(\mathbb{R}^n)$ goes as follows. We first define

$$y_{j,f}(k) := |\det\Lambda|^{j/2}\langle f, (\phi_c)_{j,k}\rangle, \tag{2.1}$$

and observe that $Cy_{j,f} = y_{j-1,f}$, all j, with C the compression operator of the currently selected CAP. Then, we define accordingly the detail coefficients by

$$d_{j,f} := (A - PAC)y_{j,f}, \quad j \in \mathbb{Z}.$$

We refer to $(d_{j,f})_{j\in\mathbb{Z}}$ as the *CAP coefficients (of the function f)*.

2.2. CAPlets: connecting CAP representations to framelet theory

Consider approximation schemes $(G_j)_{j\in\mathbb{Z}}$ on $L_2(\mathbb{R}^n)$ of the form

$$G_j : f \mapsto \sum_{k\in\mathbb{Z}^n} \langle f, (\phi_c)_{j,k}\rangle\ (h_a * \phi_r)_{j,k},$$

where

$$h_a * \phi_r := \sum_{l\in\mathbb{Z}^n} h_a(l)\phi_r(\cdot - l).$$

Invoking (2.1) and then the associativity of convolution we get

$$|\det\Lambda|^{j/2}\, G_j(f) = \sum_{k\in\mathbb{Z}^n} y_{j,f}(k)\,(h_a * \phi_r)_{j,k} = \sum_{k\in\mathbb{Z}^n} (h_a * y_{j,f})(k)\,(\phi_r)_{j,k} = \sum_{k\in\mathbb{Z}^n} (Ay_{j,f})(k)\,(\phi_r)_{j,k}.$$

The refinability of ϕ_c implies that $y_{j-1,f} = Cy_{j,f}$, while the refinability of ϕ_r implies that, for any sequence y,

$$|\det\Lambda|^{1/2}\sum_{k\in\mathbb{Z}^n} y(k)\,(\phi_r)_{j-1,k} = \sum_{k\in\mathbb{Z}^n}(Py)(k)\,(\phi_r)_{j,k}.$$

Consequently,

$$|\det\Lambda|^{j/2}\,G_{j-1}(f) = |\det\Lambda|^{1/2}\sum_{k\in\mathbb{Z}^n}(ACy_{j,f})(k)\,(\phi_r)_{j-1,k} = \sum_{k\in\mathbb{Z}^n}(PACy_{j,f})(k)\,(\phi_r)_{j,k},$$

and hence

$$|\det\Lambda|^{j/2}\,(G_j - G_{j-1})f = \sum_{k\in\mathbb{Z}^n} d_{j,f}(k)\;(\phi_r)_{j,k},\quad \forall j\in\mathbb{Z},$$

with $d_{j,f}$ the CAP coefficients, i.e., $d_{j,f} = (A - PAC)y_{j,f}$. In fact, the following lemma shows that these CAP coefficients $(d_{j,f})_j$ are the framelet coefficients for a suitably chosen framelet system.

Lemma 2.2. *Let $\Gamma := \mathbb{Z}^n/(\Lambda\mathbb{Z}^n)$ and let $\Gamma^* := 2\pi((\Lambda^{*-1}\mathbb{Z}^n)/\mathbb{Z}^n)$. For $\upsilon\in\Gamma$, let t_υ be the following trigonometric polynomial:*

$$t_\upsilon := e_\upsilon\left(\overline{\Theta} - \overline{\Theta}(\Lambda^*\cdot)\,\tau_c\left(\sum_{\upsilon^*\in\Gamma^*} e^{i\upsilon\cdot\upsilon^*}\,\overline{\tau}_r(\cdot+\upsilon^*)\right)\right),\quad e_\upsilon:\omega\mapsto|\det\Lambda|^{-1/2}e^{i\upsilon\cdot\omega}$$

and define ψ_υ by $\widehat{\psi}_\upsilon(\Lambda^*\cdot) := t_\upsilon\widehat{\phi}_c$. Then, for $f\in L_2(\mathbb{R}^n)$ and for $\upsilon\in\Gamma$,

$$d_{j,f}(\Lambda k - \upsilon) = |\det\Lambda|^{j/2}\,\langle f,(\psi_\upsilon)_{j-1,k}\rangle,\quad j\in\mathbb{Z}, k\in\mathbb{Z}^n. \tag{2.3}$$

Remark. We refer to each ψ_υ as a *CAPlet* and to each t_υ as the (corresponding) *CAPlet mask.* □

Remark. In view of the above lemma, we make (2.3) the *definition* of the CAP detail coefficient $(d_{j,f}(k))_{j,k}$ of f, for as long as such definition makes sense. Note that this more general definition coincides with the definition in terms the approximation operators G_j for all locally integrable functions f. □

Proof : One possible proof is to translate to the Fourier domain the representation of G_{-1} that was obtained in the remarks preceding this lemma. We choose instead to argue directly from scratch, using the fact that $G_{-1} = D^{-1}G_0D$. We prove only the case $\Theta = 1$. The proof for general Θ is similar and is left to the reader. Also, we may assume without loss that $j = 0$, since the general case j follows from the $(j = 0)$-case by dilation. We will use in the proof the orthogonality of non-trivial characters to the trivial one, viz.,

$$\sum_{\upsilon\in\Gamma} e^{i\upsilon\cdot\upsilon^*} = \begin{cases}|\det\Lambda|, & \text{if } \upsilon^*=0,\\ 0, & \text{if } \upsilon^*\in\Gamma^*\backslash 0,\end{cases}\qquad \sum_{\upsilon^*\in\Gamma^*} e^{i\upsilon\cdot\upsilon^*} = \begin{cases}|\det\Lambda|, & \text{if } \upsilon=0,\\ 0, & \text{if } \upsilon\in\Gamma\backslash 0.\end{cases} \tag{2.4}$$

We will further need two properties of bracket products. Recall that the bracket product of two L_2-functions u, v is defined by

$$[u,v] := \sum_{\beta\in 2\pi\mathbb{Z}^n}(u\overline{v})(\cdot+\beta).$$

The properties we need are as follows:

$$D^*([D^{-*}u, \widehat{\phi_c}]) = \sum_{\upsilon^* \in \Gamma^*} (\overline{\tau}_c[u, \widehat{\phi_c}])(\cdot + \upsilon^*), \quad \widehat{y_{0,f}} = [\widehat{f}, \widehat{\phi_c}], \tag{2.5}$$

where $D^{-*} := (D^*)^{-1}$, and D^* is defined as

$$D^* : g \mapsto |\det \Lambda|^{1/2} g(\Lambda^* \cdot).$$

Both properties are straightforward to check.

Now, (2.5) implies that $\widehat{G_0 f} = [\widehat{f}, \widehat{\phi_c}]\widehat{\phi_r}$, and, also, since $G_{-1}f = D^{-1}G_0 Df$, it further implies that

$$(G_{-1}f)\widehat{\ } = (D^{-1}G_0 Df)\widehat{\ } = D^*([D^{-*}\widehat{f}, \widehat{\phi_c}]\widehat{\phi_r}) = D^*([D^{-*}\widehat{f}, \widehat{\phi_c}])\, \tau_r \widehat{\phi_r} = \tau_r \widehat{\phi_r} \sum_{\upsilon^* \in \Gamma^*} (\overline{\tau}_c[\widehat{f}, \widehat{\phi_c}])(\cdot + \upsilon^*).$$

Denoting $\rho := [\widehat{f}, \widehat{\phi_c}]$, we conclude that

$$(G_0 f - G_{-1}f)\widehat{\ } = \widehat{\phi_r} \left(\rho - \tau_r \sum_{\upsilon^* \in \Gamma^*} (\overline{\tau}_c \rho)(\cdot + \upsilon^*) \right).$$

On the other hand, (2.3) is equivalent to the statement that the Fourier series of the sequence $\Lambda \mathbb{Z}^n \ni k \mapsto d_{0,f}(k - \upsilon)$ is $|\det \Lambda|^{-1/2} \sum_{\tilde{\upsilon}^* \in \Gamma^*} (\overline{t_\upsilon} \rho)(\cdot + \tilde{\upsilon}^*)$, hence that

$$\widehat{d_{0,f}} = \sum_{\tilde{\upsilon}^* \in \Gamma^*} \rho(\cdot + \tilde{\upsilon}^*) \sum_{\upsilon \in \Gamma} e_\upsilon \overline{t_\upsilon}(\cdot + \tilde{\upsilon}^*).$$

We then expand $\sum_{\upsilon \in \Gamma} e_\upsilon \overline{t_\upsilon}(\cdot + \tilde{\upsilon}^*)$ to obtain

$$\begin{aligned}
&\frac{1}{|\det \Lambda|} \sum_{\upsilon \in \Gamma} e^{-i\upsilon \cdot \tilde{\upsilon}^*} - \frac{1}{|\det \Lambda|}\, \overline{\tau}_c(\cdot + \tilde{\upsilon}^*) \sum_{\upsilon^* \in \Gamma^*} \tau_r(\cdot + \upsilon^* + \tilde{\upsilon}^*) \sum_{\upsilon \in \Gamma} e^{-i\upsilon \cdot (\upsilon^* + \tilde{\upsilon}^*)} \\
&= \frac{1}{|\det \Lambda|} \sum_{\upsilon \in \Gamma} e^{-i\upsilon \cdot \tilde{\upsilon}^*} - \overline{\tau}_c(\cdot + \tilde{\upsilon}^*)\tau_r.
\end{aligned}$$

Consequently,

$$\sum_{\tilde{\upsilon}^* \in \Gamma^*} \rho(\cdot + \tilde{\upsilon}^*) \sum_{\upsilon \in \Gamma} e_\upsilon \overline{t_\upsilon}(\cdot + \tilde{\upsilon}^*) = \rho - \sum_{\tilde{\upsilon}^* \in \Gamma^*} \rho(\cdot + \tilde{\upsilon}^*)\overline{\tau}_c(\cdot + \tilde{\upsilon}^*)\tau_r.$$

□

Thus, the CAP representation coincides with the framelet representation that is induced by the framelet system in the above lemma. Our next goal is to complement the CAPlet system by a suitable dual framelet system. In fact, the next lemma exhibits a large class of dual systems. At this point, we will merely introduce all these dual systems, and prove their core connection with the above CAPlet/framelet one. We are not claiming (yet) that these systems are dual to our CAPlet; that further claim will be established in Theorem 2.11.

Lemma 2.6. *Suppose that $(t_\upsilon : \upsilon \in \Gamma)$ are given as in Lemma 2.2. Let ξ be any trigonometric polynomial such that $\xi(0) = 1$ and define*

$$\begin{aligned} t^{\text{dual}} &:= \tau_r \left(1 + \xi \left(1 - \sum_{\upsilon^* \in \Gamma^*} \left(\frac{\Theta(\Lambda^* \cdot)}{\Theta} \overline{\tau}_c \tau_r\right)(\cdot + \upsilon^*)\right)\right), \\ t_\upsilon^{\text{dual}} &:= e_\upsilon \left(1 - \xi\tau_r \left(\sum_{\upsilon^* \in \Gamma^*} e^{i\upsilon\cdot\upsilon^*} \left(\frac{\Theta(\Lambda^* \cdot)}{\Theta} \overline{\tau}_c\right)(\cdot + \upsilon^*)\right)\right), \quad \upsilon \in \Gamma. \end{aligned}$$

Then (cf. (1.4)):

$$\begin{aligned} \sum_{\upsilon \in \Gamma} \overline{t}_\upsilon(\omega) t_\upsilon^{\text{dual}}(\omega) &= \Theta(\omega) - \Theta(\Lambda^*\omega)\overline{\tau}_c(\omega) t^{\text{dual}}(\omega), \\ \sum_{\upsilon \in \Gamma} \overline{t}_\upsilon(\omega) t_\upsilon^{\text{dual}}(\omega + \nu) &= -\Theta(\Lambda^*\omega)\overline{\tau}_c(\omega) t^{\text{dual}}(\omega + \nu), \quad \nu \in \Gamma^* \setminus 0. \end{aligned}$$

Remark. Ideally, we would like to choose ξ to equal 1 in a small neighborhood of the origin, and 0 elsewhere. For this choice, one observes that $t^{\text{dual}} \approx \tau_r$ *everywhere*, while the wavelet masks satisfy $t_\upsilon^{\text{dual}} \approx e_\upsilon$ outside a small neighborhood of 0. So, by a small perturbation of the system $((\phi_r)_{j,k})_{j,k}$ we obtain a bi-framelet system whose action coincides with that of $((\phi_r)_{j,k})_{j,k}$ on the range of $T^*_{X(\Psi)}$ (for the CAPlet set $\Psi := \{\psi_\upsilon\}_{\upsilon\in\Gamma}$). By controlling the support of ξ we are able to reduce gradually the magnitude of the perturbation.

A rigorous treatment along the above lines is given in §3.5. □

Proof : We use the properties in (2.4) to obtain the equalities :

$$\begin{aligned} &\sum_{\upsilon \in \Gamma} \overline{t}_\upsilon(\omega) t_\upsilon^{\text{dual}}(\omega) \\ &= \frac{1}{|\det \Lambda|} \sum_{\upsilon \in \Gamma} \left(\Theta(\omega) - \Theta(\Lambda^*\omega)\overline{\tau_c(\omega)} \sum_{\upsilon^* \in \Gamma^*} e^{-i\upsilon\cdot\upsilon^*} \tau_r(\omega + \upsilon^*)\right) \\ &\qquad\qquad \cdot \left(1 - (\xi\tau_r)(\omega) \sum_{\tilde{\upsilon}^* \in \Gamma^*} e^{i\upsilon\cdot\tilde{\upsilon}^*} \frac{\Theta(\Lambda^*(\omega + \tilde{\upsilon}^*))}{\Theta(\omega + \tilde{\upsilon}^*)} \overline{\tau_c(\omega + \tilde{\upsilon}^*)}\right) \\ &= \Theta(\omega) - \Theta(\Lambda^*\omega)\overline{\tau_c(\omega)}\tau_r(\omega) \left(1 + \xi(\omega) \left(1 - \sum_{\upsilon^* \in \Gamma^*} \frac{\Theta(\Lambda^*(\omega + \upsilon^*))}{\Theta(\omega + \upsilon^*)} \overline{\tau_c(\omega + \upsilon^*)}\tau_r(\omega + \upsilon^*)\right)\right) \\ &= \Theta(\omega) - \Theta(\Lambda^*\omega)\overline{\tau_c(\omega)} t^{\text{dual}}(\omega), \end{aligned}$$

and for any $\nu \in \Gamma^* \setminus 0$,

$$\begin{aligned} &\sum_{\upsilon \in \Gamma} \overline{t}_\upsilon(\omega) t_\upsilon^{\text{dual}}(\omega + \nu) \\ &= \frac{1}{|\det \Lambda|} \sum_{\upsilon \in \Gamma} e^{i\upsilon\cdot\nu} \left(\Theta(\omega) - \Theta(\Lambda^*\omega)\overline{\tau_c(\omega)} \sum_{\upsilon^* \in \Gamma^*} e^{-i\upsilon\cdot\upsilon^*} \tau_r(\omega + \upsilon^*)\right) \\ &\qquad\qquad \cdot \left(1 - (\xi\tau_r)(\omega + \nu) \sum_{\tilde{\upsilon}^* \in \Gamma^*} e^{i\upsilon\cdot\tilde{\upsilon}^*} \frac{\Theta(\Lambda^*(\omega + \nu + \tilde{\upsilon}^*))}{\Theta(\omega + \nu + \tilde{\upsilon}^*)} \overline{\tau_c(\omega + \nu + \tilde{\upsilon}^*)}\right) \\ &= -\Theta(\Lambda^*\omega)\overline{\tau_c(\omega)}\tau_r(\omega + \nu) \left(1 + \xi(\omega + \nu) \left(1 - \sum_{\upsilon^* \in \Gamma^*} \frac{\Theta(\Lambda^*(\omega + \upsilon^*))}{\Theta(\omega + \upsilon^*)} \overline{\tau_c(\omega + \upsilon^*)}\tau_r(\omega + \upsilon^*)\right)\right) \\ &= -\Theta(\Lambda^*\omega)\overline{\tau_c(\omega)} t^{\text{dual}}(\omega + \nu). \end{aligned}$$

$\square$

Suppose now that the trio (h_c, h_a, h_r) satisfies the relation

$$\sum_{\upsilon^*\in\Gamma^*}\left(\frac{\Theta(\Lambda^*\cdot)}{\Theta}\overline{\tau}_c\tau_r\right)(\cdot+\upsilon^*) = \sum_{\upsilon^*\in\Gamma^*}\left(\frac{\widehat{h_a}(\Lambda^*\cdot)}{\widehat{h_a}}\widehat{h_c}\widehat{h_r}\right)(\cdot+\upsilon^*) = 1. \tag{2.7}$$

Then, in our construction above we obtain that $t^{\text{dual}} = \tau_r$. In this case, it is natural to select $\xi = 1$. The result is recorded in the next corollary. Note that (ii) of the corollary applies to the case when the shifts of ϕ_c are biorthogonal to the shifts of ϕ_r, while (iii) of the corollary shows that the CAP representation is unitary whenever $\phi_c = \phi_r$, and the shifts of ϕ_c are orthonormal.

Corollary 2.8. *If (2.7) holds, then the dual masks $(t^{\text{dual}}, (t_\upsilon^{\text{dual}})_\upsilon)$ of the CAPlet system can be chosen as follows:*

$$t^{\text{dual}} := \tau_r, \quad t_\upsilon^{\text{dual}} := e_\upsilon\left(1 - \tau_r\left(\sum_{\upsilon^*\in\Gamma^*} e^{i\upsilon\cdot\upsilon^*}\left(\frac{\Theta(\Lambda^*\cdot)}{\Theta}\overline{\tau}_c\right)(\cdot+\upsilon^*)\right)\right).$$

In particular:

(i) The dual system is also a CAPlet system, that employs no alignment. The compression refinable function in that dual system is ϕ_r (hence the Fourier series of the compression filter is $\overline{\tau_r}$), while the prediction refinable function in the dual system is $\overline{h_a}*\phi_c$, with $\overline{h_a}(k) := \overline{h_a(-k)}$ (hence the Fourier series of the prediction filter is $\frac{\overline{\Theta(\Lambda^*\cdot)}}{\overline{\Theta}}\tau_c$).

(ii) Suppose that (2.7) holds with $\Theta = 1$ (i.e., we use no alignment). Then the dual CAPlet system is obtained from the original CAPlet system by switching the roles of the compression and prediction filters.

(iii) Suppose that we use no alignment, and that $\tau_c = \tau_r$. In this case (2.7) reduces to the CQF condition:

$$\sum_{\upsilon^*\in\Gamma^*}|\tau_c(\cdot+\upsilon^*)|^2 = 1,$$

and our result shows that the CAPlet system is a tight framelet satisfying the UEP.

Remark. While the previous constructions in Lemma 2.2 and Lemma 2.6 are done for any integer dilation matrix Λ, we assume that the dilation is dyadic in the rest of the paper. With some extra effort, the entire analysis in this paper can be done under the assumption that the dilation is isotropic, i.e. $\Lambda^*\Lambda = cI$, for some integer $c > 1$. In general, wavelet systems that are based on anisotropic dilation cannot be used for the characterization of the isotropic Triebel-Lizorkin and Besov spaces studied in this paper. When $\Lambda = 2I$, the combined masks $(\tau_c, (t_\upsilon : \upsilon \in \{0,1\}^n))$, $(t^{\text{dual}}, (t_\upsilon^{\text{dual}} : \upsilon \in \{0,1\}^n))$ of the CAPlet system and its dual system become

$$\begin{aligned}
t_\upsilon &:= e_\upsilon\left(\overline{\Theta} - \overline{\Theta}(2\cdot)\tau_c\left(\sum_{\upsilon^*\in\{0,\pi\}^n} e^{i\upsilon\cdot\upsilon^*}\overline{\tau}_r(\cdot+\upsilon^*)\right)\right),\\
t^{\text{dual}} &:= \tau_r\left(1 + \xi\left(1 - \sum_{\upsilon^*\in\{0,\pi\}^n}\left(\frac{\Theta(2\cdot)}{\Theta}\overline{\tau}_c\tau_r\right)(\cdot+\upsilon^*)\right)\right),\\
t_\upsilon^{\text{dual}} &:= e_\upsilon\left(1 - \xi\tau_r\left(\sum_{\upsilon^*\in\{0,\pi\}^n} e^{i\upsilon\cdot\upsilon^*}\left(\frac{\Theta(2\cdot)}{\Theta}\overline{\tau}_c\right)(\cdot+\upsilon^*)\right)\right),
\end{aligned} \tag{2.9}$$

where $e_\upsilon(\omega) = 2^{-n/2}e^{i\upsilon\cdot\omega}$ and ξ is any trigonometric polynomial that satisfies $\xi(0) = 1$. $\square$

Next, we show that we can always find a dual framelet to the given CAPlet one so that the pair becomes a bi-framelet. To this end, we need the following result from [Me2].

Result 2.10. *Let $\alpha > 0$. Suppose that f has Hölder smoothness α, and that $|f(t)| \le c\,(1+|t|)^{-n-\alpha}$, for every $t \in \mathbb{R}^n$. Then the system $(f_{j,k} : \ j \in \mathbb{Z},\ k \in \mathbb{Z}^n)$ is Bessel if $\widehat{f}(0) = 0$.*

Theorem 2.11. *Let the CAPlet masks $\{t_\upsilon\}_{\upsilon \in \{0,1\}^n}$ be as in (2.9). Let $\Psi := \{\psi_\upsilon\}_{\upsilon\in\{0,1\}^n}$ be the corresponding CAPlets. Then there exists a framelet system $X(\Psi^{\mathrm{dual}})$ associated with a refinable function ϕ^{dual} so that the pair $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ is a bi-framelet.*

Proof: We first prove that $X(\Psi)$ is a Bessel system. For that, it suffices to show that $X(\psi_\upsilon)$ is Bessel, for each υ, which will follow once we verify that ψ_υ satisfies the assumptions needed in Result 2.10. The decay assumption in Result 2.10 is trivially satisfied since Assumption 1.7(a) implies that ψ_υ is compactly supported. The smoothness assumption follows from Assumption 1.7(c). The condition $\widehat{\psi}_\upsilon(0) = 0$ is equivalent to $t_\upsilon(0) = 0$. This latter condition is argued as follows. First, since $\widehat{\phi}_r(0) = \widehat{\phi}_c(0) = 1$ (Assumption 1.7(b)), we have that $\tau_c(0) = \tau_r(0) = 1$, too. Also, $\Theta(0) = 1$, by Assumption 1.7(e). Thus,

$$t_\upsilon(0) = 2^{-n/2} \sum_{\upsilon^* \in \{0,\pi\}^n \setminus 0} e^{i\upsilon\cdot\upsilon^*}\,\overline{\tau}_r(\upsilon^*).$$

This latter sum is 0, since τ_r satisfies the SF condition of order 1, by virtue of its positive smoothness, [R1]. Thus all the conditions of Result 2.10 are satisfied and hence $X(\Psi)$ is Bessel.

In order to prove that the dual system $X(\Psi^{\mathrm{dual}})$ is Bessel, too, we invoke again Result 2.10, hence need to verify that each $\psi_\upsilon^{\mathrm{dual}}$ satisfies the assumptions in that result. The fact that $\widehat{\psi}_\upsilon^{\mathrm{dual}}(0) = 0$ is due to the fact that $t_\upsilon^{\mathrm{dual}}(0) = 0$, with the latter argued as in the $X(\Psi)$ case, using now the additional facts that $\xi(0) = 1$, and that τ_c satisfies the SF conditions of order 1, too.

The requisite decay rate of $\psi_\upsilon^{\mathrm{dual}}$ follows from the facts that t^{dual} as well as $t_\upsilon^{\mathrm{dual}}$ are rational polynomials, and that the denominator of t^{dual}, viz., Θ, does not vanish on $\mathbb{R}^n$. Indeed, these conditions imply first that the refinable ϕ^{dual} must decay exponentially at ∞, and then that the Fourier coefficients of $t_\upsilon^{\mathrm{dual}}$ decay exponentially at ∞ as well. Thus, $\psi_\upsilon^{\mathrm{dual}}$ decays exponentially fast at ∞.

Last, we will need to know that $\psi_\upsilon^{\mathrm{dual}}$ is minimally smooth. This follows from the corresponding smoothness of ϕ^{dual}. The latter follows from the (minimal) smoothness assumption on ϕ_r (Assumption 1.7(c)), but requires a careful selection of the polynomial ξ. We skip this non-trivial argument (i.e., that ξ can be chosen such that ϕ^{dual} has positive Hölder smoothness), since we will prove a more general result (Lemma 3.15) later on.

Thus, for a suitable choice of ξ, $X(\Psi^{\mathrm{dual}})$ satisfies the requirements of Result 2.10, and hence it is a Bessel system. At this point, we can appeal to Lemma 2.6: since we satisfy all the conditions of the OEP, the pair $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ is a bi-framelet. $\square$

2.3. Compression-Alignment-Modified-Prediction (CAMP) representations

In this section, we introduce an important variation of the CAP representation idea. The variation allows us to improve the time/space localness of the representation, without altering its performance, i.e., the class of function space characterizations that the representation provides.

A CAMP representation can be derived from any CAP representation that satisfies one additional condition. The condition is that the prediction operator P of the CAP representation is *pseudo-interpolatory*. This means that the prediction filter h_r is the convolution product of two filters, one interpolatory, $h_{\rm in}$, and another one which is supported on the *even* integers. It is thus possible (and convenient) to write that latter filter in the form $h_{{\rm e}\uparrow}$, for some other filter $h_{\rm e}$. In summary,

$$h_r = h_{{\rm e}\uparrow} * h_{\rm in}. \tag{2.12}$$

We refer to $h_{\rm in}$ and $h_{\rm e}$ as the interpolatory (respectively, enhancement) filter of the CAMP representation. To recall, a filter $h_{\rm in}$ is *interpolatory* if

$$h_{\rm in}(0) = \frac{1}{2^n}, \quad h_{\rm in}(2k) = 0, \ \forall k \in \mathbb{Z}^n \backslash 0.$$

The choice $h_{\rm e} = \delta$ is permitted in the above setup. In that case, h_r coincides with $h_{\rm in}$ and hence is interpolatory, as well. The operator P is labeled then *interpolatory*, too. Thus, our condition is always satisfied whenever the prediction operator is interpolatory.

We start our discussion of CAMP representations by noting that the prediction operator P in the pseudo-interpolatory case can be factored as follows:

$$P = P_{\rm in} H_{\rm e}, \quad H_{\rm e} : y \mapsto h_{\rm e} * y, \quad P_{\rm in} : y \mapsto 2^n (h_{\rm in} * (y_\uparrow)).$$

Then, we recall that the CAP coefficients $(d_j)_{j\in\mathbb{Z}}$ are defined by $d_j = (A - PAC)y_j$. Let C_1 be the trivial coarsification operator, i.e., the one associated with $h_c := \delta$; thus $C_1 y = y_\downarrow$. Now, we define the CAMP representation $(\widetilde{d}_j)_j$ by

$$\widetilde{d}_j(k) := \begin{cases} d_j(k), & \text{if } k \in 2\mathbb{Z}^n, \\ (A - P_{\rm in} C_1 A) y_j(k), & \text{if } k \in \mathbb{Z}^n \backslash (2\mathbb{Z}^n). \end{cases} \tag{2.13}$$

We call this process *Compression-Alignment-Modified-Prediction* (CAMP) *representation* and refer to $(\widetilde{d}_j)_{j\le 0}$ as the *CAMP coefficients of* y_0. Note that we have not altered the way the sequence $(y_j)_{j\le 0}$ is defined, but only the way the detail coefficients are extracted from that sequence.

The CAMP representation is more local than its CAP counterpart. In part this is due to the use of C_1 and $P_{\rm in}$ instead of C and P, and in part, in case $A \neq I$, because we apply alignment *before* coarsening. This provides us with the motivation to look closer at the connection between the CAMP coefficients and the CAP ones.

The connection between the CAMP and CAP coefficients is actually quite simple. Since $P_{\rm in}$ is interpolatory, we have that $C_1 P_{\rm in} = I$, and hence $C_1 P = C_1 P_{\rm in} H_{\rm e} = H_{\rm e}$. Consequently, $C_1 d_j = C_1 (A - PAC) y_j = (C_1 A - H_{\rm e} AC) y_j$. On the other hand, on $\mathbb{Z}^n \backslash (2\mathbb{Z}^n)$, $d_j - \widetilde{d}_j = (P_{\rm in} C_1 A - PAC) y_j = P_{\rm in} (C_1 A - H_{\rm e} AC) y_j$, hence

$$d_j - \widetilde{d}_j = P_{\rm in} C_1 d_j = P_{\rm in} C_1 \widetilde{d}_j, \quad \text{on } \mathbb{Z}^n \backslash (2\mathbb{Z}^n).$$

Thus, the difference between the CAP and CAMP coefficients is controlled by the size of $C_1 d_j$, which appears both in the CAP and CAMP representations, provided that the operator $P_{\rm in}$ is bounded in a suitable sense. Precisely, suppose that our metric/norm is defined by a functional F that acts on the detail coefficients. Suppose that F has, at least on all CAP and CAMP sequences, the following two properties:

$$F(a + b) \le c\,(F(a) + F(b)), \quad F(P_{\rm in} C_1 a) \le c\,F(a). \tag{2.14}$$

Then

$$F((d_j)_j) \le c\,(F((\widetilde{d}_j)_j) + F((d_j - \widetilde{d}_j)_j)) = c\,(F((\widetilde{d}_j)_j) + F((P_{\rm in}C_1\widetilde{d}_j)_j)) \le c\,F((\widetilde{d}_j)_j),$$

and the converse inequality is obtained similarly. Thus, we need to prove that the functionals F that are employed in our norm/metric definitions satisfy the two properties in (2.14). Then we will conclude that the CAMP representation provides the same function space characterizations as its CAP counterpart (up to the equivalence constants, of course).

Our next two lemmata show that, indeed, the two metrics that we need in this paper (one for the characterization of Besov spaces and the other for the characterization of Triebel-Lizorkin spaces) satisfy the above requisite property. As said, this implies that a CAP representation and its associated CAMP representation provide equivalent characterizations to exactly the same range of functions spaces. Since the CAMP representation is always more local than the associated CAP, it should be considered the preferred one, whenever it is available, i.e., whenever h_r is pseudo-interpolatory.

Lemma 2.15. *Let $s \in \mathbb{R}$, $0 < p < \infty$ and $0 < q \le \infty$. For $a := (a_j \in \mathbb{C}^{\mathbb{Z}^n} : j \in \mathbb{Z})$, define $F_{TL}(a)$ by*

$$F_{TL}(a) := \left\| \left(\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \left(2^{js} |a_j(k)| \chi_{\infty,j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)},$$

(with the usual modification for $q = \infty$). Then F_{TL} satisfies the properties in (2.14).

Proof: The first property is straightforward.
To prove the second property, first define $\mathbf{M} := (\mathbf{M}(j,k;l,m) : j,l \in \mathbb{Z},\ k,m \in \mathbb{Z}^n)$ by

$$\mathbf{M}(j,k;l,m) := \begin{cases} h_{\rm in}(k-m), & l = j, \\ 0, & l \ne j. \end{cases}$$

Then $(h_{\rm in} * a_j)(k) = (\mathbf{M}a)(j,k)$. Applying Proposition 6.9 to the current $\mathbf{M}$ with $\sigma := \frac{n}{s}$ gives $F_{TL}((h_{\rm in} * a_j)_j) \le cF_{TL}((a_j)_j)$. Therefore we have

$$F_{TL}(P_{\rm in}C_1 a) = 2^n F_{TL}((h_{\rm in} * (a_{j\downarrow\uparrow}))_j) \le c\,F_{TL}((a_{j\downarrow\uparrow})_j) \le c\,F_{TL}((a_j)_j).$$

□

Lemma 2.16. *Let $s \in \mathbb{R}$, $0 < p, q \le \infty$. For $a := (a_j \in \mathbb{C}^{\mathbb{Z}^n} : j \in \mathbb{Z})$, define $F_B(a)$ by*

$$F_B(a) := \left(\sum_{j \in \mathbb{Z}} \left(2^{-jn} \sum_{k \in \mathbb{Z}^n} |2^{js} a_j(k)|^p \right)^{q/p} \right)^{1/q},$$

(with the usual modification for $p = \infty$ or $q = \infty$). Then F_B satisfies the properties in (2.14).

Proof: The proof is identical to that of Lemma 2.15: one only needs to use Proposition 6.3 with $t := s - \frac{n}{p}$ instead of Proposition 6.9. □

Remark. In both lemmata, we used the fact that the filter h_{in} is finitely supported. The arguments used, however, remain intact for filters h_{in} that satisfy

$$|h_{\text{in}}(k)| \le c\,(1+|k|)^{-\gamma},$$

for all $k \in \mathbb{Z}^n$ and for some $\gamma > \mu$, where $\mu = \dfrac{n}{\min\{1,p,q\}}$ for F_{TL} and $\mu = \dfrac{n}{\min\{1,p\}}$ for F_B. $\square$

Finally, the CAMP representation is also a special framelet representation. The CAMPlets are implicit in (2.13). We note that the operator $P_{\text{in}}C_1A$ (that was used in (2.13) to define the CAMP representation for $\upsilon \in \{0,1\}^n\backslash 0$) is equivalent to applying the compression with the filter h_a followed by the prediction with the filter h_{in}. We do not believe, however, that this (correct) interpretation is insightful: the alignment filter is a full-pass, not low-pass, and the prediction obtained in this way is a prediction of Ay_j, not y_j.

For the discussion in §5, we record the CAMPlets that are behind a CAMP representation. For an integer scalar dilation $\Lambda = \lambda I$, $\lambda > 1$, let $\Gamma := \{0,1,\cdots,\lambda-1\}^n$ and $\Gamma^* := \{0, \frac{1}{\lambda}2\pi, \cdots, \frac{\lambda-1}{\lambda}2\pi\}^n$. Then the CAMPlet masks $(t_\upsilon^M : \upsilon \in \Gamma)$ are

$$t_\upsilon^M := \begin{cases} \lambda^{-n/2}\left(\overline{\Theta} - \overline{(\tau_{\text{e}}\Theta)}(\lambda\cdot)\tau_c\right), & \upsilon = 0, \\ e_\upsilon\overline{\Theta}\left(1 - \left(\displaystyle\sum_{\upsilon^*\in\Gamma^*} e^{i\upsilon\cdot\upsilon^*}\overline{\tau_{\text{in}}}\,(\cdot+\upsilon^*)\right)\right), & \upsilon \in \Gamma\backslash 0, \end{cases}$$

where $e_\upsilon(\omega) = \lambda^{-n/2}e^{i\upsilon\cdot\omega}$ and the masks τ_{e}, τ_{in} are the Fourier series of the filters h_{e}, h_{in}, respectively.

3. CAP characterizations of Triebel-Lizorkin and Besov spaces

Since the CAP coefficients are framelet coefficients with respect to the CAPlet system (2.9) (cf. Lemma 2.2 and Theorem 2.11), we can employ the current theory concerning characterizations of smoothness via wavelet and framelet coefficients. This will allow us to identify accurately the features in the CAP representation that determine its performance, i.e., its ability to encode smoothness. In §3.1 and §3.2, we review and modify the known function space characterizations in terms of wavelet/framelet coefficients. We then apply those characterizations directly to the CAP bi-framelet constructed in Theorem 2.11. In this way we obtain characterizations that involve the dual framelet system, hence should be considered unsatisfactory. To this end, we then obtain (in §3.3) our final characterizations that rely on properties of the CAP filters (h_c, h_a, h_r) and the associated refinable functions ϕ_c, ϕ_r, and on nothing else.

3.1. The Triebel-Lizorkin space $\dot{F}^s_{pq}$ and the Besov space $\dot{B}^s_{pq}$

We recall first (one of) the (equivalent) definition(s) of Triebel-Lizorkin and Besov spaces [T]. Let $\varphi \in \mathcal{S}$ be such that

$$\begin{aligned} &\text{(i) } \operatorname{supp}\widehat{\varphi} \subset \{\tfrac{1}{2} \le |\omega| \le 2\}, \\ &\text{(ii) } |\widehat{\varphi}(\omega)| \ge c > 0, \quad \tfrac{3}{5} \le |\omega| \le \tfrac{5}{3}, \\ &\text{(iii) } |\widehat{\varphi}(\omega)|^2 + |\widehat{\varphi}(\tfrac{\omega}{2})|^2 = 1, \quad 1 < |\omega| < 2. \end{aligned} \tag{3.1}$$

Let $\varphi_j := 2^{jn}\varphi(2^j\cdot)$, for $j \in \mathbb{Z}$.

For $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \le \infty$, the (homogeneous) Triebel-Lizorkin space $\dot F^s_{pq} := \dot F^s_{pq}(\mathbb{R}^n)$ is defined to be the set of all $f \in \mathcal{S}'/\mathcal{P}$ such that

$$\|f\|_{\dot F^s_{pq}(\mathbb{R}^n)} := \left\| \left(\sum_{j\in\mathbb{Z}} (2^{js}|\varphi_j * f|)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} < \infty,$$

with the usual modification for $q = \infty$.

For $s \in \mathbb{R}$, $0 < p \le \infty$, $0 < q \le \infty$, the (homogeneous) Besov space $\dot B^s_{pq} := \dot B^s_{pq}(\mathbb{R}^n)$ is defined to be the set of all $f \in \mathcal{S}'/\mathcal{P}$ such that

$$\|f\|_{\dot B^s_{pq}(\mathbb{R}^n)} := \left(\sum_{j\in\mathbb{Z}} (2^{js}\|\varphi_j * f\|_{L_p(\mathbb{R}^n)})^q \right)^{1/q} < \infty,$$

with the usual modification for $q = \infty$.

In [FJ1], [FJ2], M. Frazier and B. Jawerth showed that the convolution operator in the above definitions of $\dot F^s_{pq}$, $\dot B^s_{pq}$ can be discretized:

Result 3.2. *Let $\varphi \in \mathcal{S}$ be as in (3.1). If $f \in \mathcal{S}'/\mathcal{P}$, then*

$$f = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} \langle f, \varphi_{j,k}\rangle \varphi_{j,k}$$

in the sense of $\mathcal{S}'/\mathcal{P}$. Moreover,

$$\|f\|_{\dot F^s_{pq}(\mathbb{R}^n)} \approx \left\| \left(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^n} \left(2^{js}|\langle f, \varphi_{j,k}\rangle|\chi_{j,k}\right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)},$$

(with the usual modification for $q = \infty$) and

$$\|f\|_{\dot B^s_{pq}(\mathbb{R}^n)} \approx \left(\sum_{j\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}^n} \left(2^{js}2^{jn(1/2-1/p)}|\langle f, \varphi_{j,k}\rangle|\right)^p \right)^{q/p} \right)^{1/q},$$

(with the usual modification for $p = \infty$ or $q = \infty$). □

For $s > 0$, $1 \le p < \infty$ and $0 < q \le \infty$, we define the inhomogeneous Triebel-Lizorkin space $F^s_{pq} := F^s_{pq}(\mathbb{R}^n)$ to be the set of all $f \in \mathcal{S}'$ such that $\|f\|_{F^s_{pq}(\mathbb{R}^n)} := \|f\|_{L_p(\mathbb{R}^n)} + \|f\|_{\dot F^s_{pq}(\mathbb{R}^n)} < \infty$. Similarly, for $s > 0$, $1 \le p \le \infty$ and $0 < q \le \infty$, we define the inhomogeneous Besov space $B^s_{pq} := B^s_{pq}(\mathbb{R}^n)$ to be the set of all $f \in \mathcal{S}'$ such that $\|f\|_{B^s_{pq}(\mathbb{R}^n)} := \|f\|_{L_p(\mathbb{R}^n)} + \|f\|_{\dot B^s_{pq}(\mathbb{R}^n)} < \infty$.

We note that many of the traditional smoothness spaces can be captured by choosing suitably the parameters in a Triebel-Lizorkin or a Besov space. The L_p-space, $1 < p < \infty$ is F^0_{p2}, and is also $\dot F^0_{p2}$. The Hardy space H_p, $0 < p \le 1$ is $\dot F^0_{p2}$. The Sobolev space W^s_p, $s > 0$, $1 < p < \infty$ is F^s_{p2}. Also, $B^1_{\infty,\infty}$ is the Zygmund space, while, more generally, for $s > 0$, $B^s_{\infty\infty}$ is the Hölder space.

In what follows, we present the (essentially known) characterizations of (homogeneous) Triebel-Lizorkin spaces and (homogeneous) Besov spaces in terms of the detail coefficients from bi-framelets. The derivation of the corresponding characterizations of inhomogeneous spaces from their homogeneous counterparts is easy, and is left to the interested reader.

3.2. Characterizations of $\dot F^s_{pq}$, $\dot B^s_{pq}$ in terms of framelet coefficients

Definition 3.3. *Let $\eta \in \mathbb{R}$ be a non-integer, and $\gamma > 0$.*

For $\eta > 0$, we define $\mathcal{R}^\eta_\gamma := \mathcal{R}^\eta_\gamma(\mathbb{R}^n)$ to be the set of all functions f such that

$$\begin{cases} |f^{(\beta)}(t)| \le c\,(1+|t|)^{-\gamma}, & \beta \in \mathbb{N}_0^n \text{ and } |\beta| \le \lfloor \eta \rfloor, \\ |f^{(\beta)}(z) - f^{(\beta)}(t)| \le c\,|z-t|^{\eta - \lfloor \eta \rfloor} \sup_{|u| \le |z-t|} (1+|u-t|)^{-\gamma}, & \beta \in \mathbb{N}_0^n,\ |\beta| = \lfloor \eta \rfloor \text{ and } |z-t| \le 3. \end{cases}$$

For $\eta < 0$, we define $\mathcal{R}^\eta_\gamma := \mathcal{R}^\eta_\gamma(\mathbb{R}^n)$ to be the set of all functions f such that

$$|f(t)| \le c\,(1+|t|)^{-\gamma}.$$

The set of all the compactly supported functions within $\mathcal{R}^\eta_\gamma$ is denoted by $\mathcal{R}^\eta := \mathcal{R}^\eta(\mathbb{R}^n)$ (and is trivially independent of γ). For $\eta \in \mathbb{R}$ and $\gamma > 0$, let $\mathcal{D}^\eta_\gamma := \mathcal{D}^\eta_\gamma(\mathbb{R}^n) := \bigcup_{\zeta > \eta} \mathcal{R}^\zeta_\gamma(\mathbb{R}^n)$ and $\mathcal{D}^\eta := \mathcal{D}^\eta(\mathbb{R}^n) := \bigcup_{\zeta > \eta} \mathcal{R}^\zeta(\mathbb{R}^n)$. □

Definition 3.4. *For $\eta \in \mathbb{R}$, we let $\mathcal{M}^\eta := \mathcal{M}^\eta(\mathbb{R}^n)$ be the set of all locally integrable functions f such that*

$$\int_{\mathbb{R}^n} |t^\beta f(t)|\, dt < \infty \quad \text{and} \quad \int_{\mathbb{R}^n} t^\beta f(t)\, dt = 0,$$

for $\beta \in \mathbb{N}_0^n$ with $|\beta| \le \eta$. That is, $f \in \mathcal{M}^\eta(\mathbb{R})$ for some $\eta \in \mathbb{N}_0$ iff $\widehat{f}$ has a zero of order $\eta + 1$ at the origin.

The characterizations of $\dot F^s_{pq}$, $\dot B^s_{pq}$ using bi-framelets that are given below are generalizations of similar results by G. Kyriazis [K]. Their proofs are presented in the Appendix (§6) and follow closely the approach of [K].

Theorem 3.5. *Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \le \infty$, $\lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s$ and let $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ be a bi-framelet satisfying $\Psi \subset \mathcal{D}^\lambda_\gamma \cap \mathcal{M}^s$, $\Psi^{\mathrm{dual}} \subset \mathcal{D}^s_\gamma \cap \mathcal{M}^\lambda$ with $\gamma > n + \max\{s+\lambda, s, \lambda\}$. Then, for every $f \in \dot F^s_{pq}$,*

$$f = \sum_{\psi \in \Psi} \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \langle f, \psi_{j,k} \rangle \psi^{\mathrm{dual}}_{j,k},$$

in the sense of $\mathcal{S}'/\mathcal{P}$. Moreover,

$$\sum_{\psi \in \Psi} \left\| \left(\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \left(2^{js} |\langle f, \psi_{j,k} \rangle| \chi_{j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \approx \|f\|_{\dot F^s_{pq}(\mathbb{R}^n)}, \tag{3.6}$$

with the usual modification for $q = \infty$.

Theorem 3.7. *Let $s \in \mathbb{R}$, $0 < p \le \infty$, $0 < q \le \infty$, $\lambda := n(\frac{1}{\min\{1,p\}} - 1) - s$ and let $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ be a bi-framelet satisfying $\Psi \subset \mathcal{D}^\lambda_\gamma \cap \mathcal{M}^s$, $\Psi^{\mathrm{dual}} \subset \mathcal{D}^s_\gamma \cap \mathcal{M}^\lambda$ with $\gamma > n + \max\{s+\lambda, s, \lambda\}$. Then, for every $f \in \dot B^s_{pq}$,*

$$f = \sum_{\psi \in \Psi} \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \langle f, \psi_{j,k} \rangle \psi^{\mathrm{dual}}_{j,k},$$

in the sense of $\mathcal{S}'/\mathcal{P}$. Moreover,

$$\sum_{\psi\in\Psi}\left(\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}^n}\left(2^{js}2^{jn(1/2-1/p)}|\langle f,\psi_{j,k}\rangle|\right)^p\right)^{q/p}\right)^{1/q}\approx\|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)},\tag{3.8}$$

with the usual modification for $p=\infty$ or $q=\infty$.

Remark. In view of the statements in Theorems 3.5 and 3.7, it is natural to ask whether the fact that the left-hand side of (3.6) (resp. (3.8)) is finite implies the membership of f in $\dot{F}^s_{pq}$ (resp. $\dot{B}^s_{pq}$). We discuss this issue in §6.4. □

The proofs, §6, of (3.6) and (3.8) are done in two steps. In the first step the inequality $\lesssim$ is established, and in the second step the opposite inequality is proved. We refer to the first inequality as a *Jackson-type* one, and to the latter as *Bernstein-type* one. A scrutiny of the proofs in §6 shows that the Jackson-type inequality requires only a subset of the conditions that we need for the complete equivalence:

Corollary 3.9. *Let $s\in\mathbb{R}$, $0<p,q\le\infty$ and let $\Psi\subset\mathcal{D}^\lambda_\gamma\cap\mathcal{M}^s$ with $\gamma>n+\max\{s+\lambda,s,\lambda\}$, where*

$$\lambda:=n(\frac{1}{\min\{1,p,q\}}-1)-s,\quad (p<\infty)$$

for Triebel-Lizorkin spaces, and

$$\lambda:=n(\frac{1}{\min\{1,p\}}-1)-s$$

for Besov spaces. Then, for Triebel-Lizorkin spaces, we have

$$\sum_{\psi\in\Psi}\left\|\left(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^n}\left(2^{js}|\langle f,\psi_{j,k}\rangle|\chi_{j,k}\right)^q\right)^{1/q}\right\|_{L_p(\mathbb{R}^n)}\lesssim\|f\|_{\dot{F}^s_{pq}(\mathbb{R}^n)}$$

and for Besov spaces, we have

$$\sum_{\psi\in\Psi}\left(\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}^n}\left(2^{js}2^{jn(1/2-1/p)}|\langle f,\psi_{j,k}\rangle|\right)^p\right)^{q/p}\right)^{1/q}\lesssim\|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)}.$$

Remark. Since the mother wavelet set Ψ does not lie in $\mathcal{S}$, then, *a priori*, we do not even have a clear-cut interpretation for the wavelet coefficients of $f\in\dot{F}^s_{pq}$ (or $\dot{B}^s_{pq}$):

$$\langle f,\psi_{j,k}\rangle,\quad j\in\mathbb{Z},k\in\mathbb{Z}^n.$$

This technical point is discussed and settled in the Appendix (cf. Corollary 6.6, Corollary 6.15 and §6.3) □

Among the various conditions that the above theorems require of the bi-framelet we focus on the vanishing moment conditions that are required of Ψ: $\Psi \subset \mathcal{M}^s$. As we mentioned before, the *approximation order* of the framelet representation could be twice larger than the number of vanishing moments. On the other hand, the only property that the framelet coefficients can "see" is the number of vanishing moments: the approximation order is captured only after (at least partial) reconstruction is done. The CAP representation is, indeed, obtained from the framelet representation by employing one step of reconstruction. The switch from framelet coefficients to CAP coefficients is tantamount to switching from a representation whose performance is saturated by the number of vanishing moments, to a representation that is governed by the approximation order. At the same time, once one performs one step of reconstruction, the resulting coefficients are *independent of the particular framelet construction*, and depend only on the refinable functions that were chosen for decomposition and reconstruction. Our introduction of CAP is consistent with the above, and avoids altogether any mentioning of high-pass filters and wavelet/framelets. The latter are needed only for the analysis of CAP!

Thus, the CAP representations are more straightforward, and capture the smoothness parameter of a wider range of functions. Indeed, in all the CAPlet constructions, the approximation order of the expansion can be shown to coincide with $\min\{m, m'\}$ with m' the approximation order provided by the shifts of the refinable function ϕ_r, and with m the CAP order, i.e., the order of the zero $\Theta - \Theta(2\cdot)\tau_r\overline{\tau}_c$ has at the origin. The vanishing moments of the CAPlet masks are now "harmonious" with the above: all the CAPlets in $\Psi := (\psi_\upsilon : \upsilon \in \{0,1\}^n)$ have $\min\{m, m'\}$ vanishing moments.

We derive the function space characterization that the CAP coefficients provide in two steps. The first, that is done in this subsection, consists of applying directly Theorems 3.5 and 3.7 to the framelet representation (Theorem 2.11 in §2.2) of the CAP process. However, the aforementioned framelet representation involves the refinable function ϕ^{dual} and its smoothness. That refinable function is not connected to the CAP representation, and was introduced only in order to recast the CAP representation as a framelet one. Our final characterizations, that are obtained in the next subsection, are purely in terms of the CAP process, i.e., the filters h_c, h_a, h_r and the associated functions ϕ_c and ϕ_r.

For the Triebel-Lizorkin case, we assume, given $s \in \mathrm{I\!R}$, $0 < p < \infty$, and $0 < q \le \infty$, that

$$(3.10)\quad \phi_c \in \mathcal{D}^\lambda, \quad \phi^{\mathrm{dual}} \in \mathcal{D}^s_\gamma, \quad \lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s, \quad \text{for some } \gamma > n + \max\{s+\lambda, s, \lambda\}.$$

Then the *smoothness* assumptions of Theorem 3.5 are satisfied. In order to deal with the vanishing moment requirements, we inspect the masks t_υ and $t_\upsilon^{\mathrm{dual}}$ in (2.9). For t_υ we write

$$e_{-\upsilon} t_\upsilon = (\overline{\Theta} - \overline{\Theta}(2\cdot)\tau_c\overline{\tau}_r) - \overline{\Theta}(2\cdot)\tau_c \left(\sum_{\upsilon^* \in \{0,\pi\}^n \setminus 0} e^{i\upsilon\cdot\upsilon^*}\overline{\tau}_r(\cdot + \upsilon^*) \right) =: R_1 - R_2.$$

By definition, R_1 has a zero of order $m :=$ the order of the given CAP. Thus, if we assume R_2 above to have a zero of order m_2 at the origin, we conclude that $\psi_\upsilon \in \mathcal{M}^{m_2 - 1}$, provided that $m_2 \le m$. Now, the order of the zero R_2 has at the origin matches or exceeds the order l_2 of the SF conditions that ϕ_r satisfies. Thus, we can take here $m_2 := \min\{m, l_2\}$. Similar analysis can be done with respect $\psi_\upsilon^{\mathrm{dual}}$, only that in this case the roles of τ_c and τ_r in R_2 above are reversed.

Thus, by applying Theorem 3.5 we obtain the following result. (We use here the fact that $\mathcal{M}^\eta = \mathcal{M}^{\lfloor \eta \rfloor}$ (cf. Definition 3.4).)

Proposition 3.11. *Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \leq \infty$ and assume that (3.10) holds (with respect to the framelet system constructed in Theorem 2.11). Suppose that ϕ_c satisfies the SF conditions of order $m_1 \geq \lfloor \lambda \rfloor + 1$, ϕ_r satisfies the SF conditions of order $m_2 \geq \lfloor s \rfloor + 1$, and that the CAP system has an order $\geq \max\{m_1, m_2\}$. Then, for every $f \in \dot{F}^s_{pq}$,*

$$f = \sum_{\upsilon \in \{0,1\}^n} \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \langle f, (\psi_\upsilon)_{j,k} \rangle (\psi_\upsilon^{\mathrm{dual}})_{j,k},$$

in the sense of $\mathcal{S}'/\mathcal{P}$. Moreover,

$$\sum_{\upsilon \in \{0,1\}^n} \left\| \left(\sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^n} (2^{js} |\langle f, (\psi_\upsilon)_{j,k} \rangle| \chi_{j,k})^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \approx \|f\|_{\dot{F}^s_{pq}(\mathbb{R}^n)}.$$

□

The analogous result with respect to Besov spaces is obtained in the same way, with the only difference being a slightly different definition of λ:

$$\text{(3.12)} \quad \phi_c \in \mathcal{D}^\lambda, \quad \phi^{\mathrm{dual}} \in \mathcal{D}^s_\gamma, \quad \lambda := n(\frac{1}{\min\{1,p\}} - 1) - s, \quad \text{for some } \gamma > n + \max\{s + \lambda, s, \lambda\}.$$

Proposition 3.13. *Let $s \in \mathbb{R}$, $0 < p \leq \infty$, $0 < q \leq \infty$ and assume that (3.12) holds (with respect to the framelet system constructed in Theorem 2.11). Suppose that ϕ_c satisfies that SF conditions of order $m_1 \geq \lfloor \lambda \rfloor + 1$, ϕ_r satisfies the SF conditions of order $m_2 \geq \lfloor s \rfloor + 1$, and that the CAP system has an order $\geq \max\{m_1, m_2\}$. Then, for every $f \in \dot{B}^s_{pq}$,*

$$f = \sum_{\upsilon \in \{0,1\}^n} \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \langle f, (\psi_\upsilon)_{j,k} \rangle (\psi_\upsilon^{\mathrm{dual}})_{j,k},$$

in the sense of $\mathcal{S}'/\mathcal{P}$. Moreover,

$$\sum_{\upsilon \in \{0,1\}^n} \left(\sum_{j \in \mathbb{Z}} \left(\sum_{k \in \mathbb{Z}^n} \left(2^{js} 2^{jn(1/2 - 1/p)} |\langle f, (\psi_\upsilon)_{j,k} \rangle| \right)^p \right)^{q/p} \right)^{1/q} \approx \|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)}.$$

□

Corollary 3.9 tells us that in order to obtain a Jackson-type inequality we only need that $\Psi \subset \mathcal{D}^\lambda_\gamma \cap \mathcal{M}^s$. In the CAPlet context, we already noted before that the condition $\Psi \subset \mathcal{M}^s$ is implied by assuming that ϕ_r satisfies the SF conditions of order $> s$ and that the CAP system has an order $> s$. We record this result in the following corollary.

Corollary 3.14. *Let ϕ_c, ϕ_r be the refinable functions used to construct a CAP bi-framelet as in Theorem 2.11 (not necessarily satisfying Assumption 1.7(d)). Let $s \in \mathbb{R}$, $0 < p, q \leq \infty$ and let*

$$\lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s, \quad (p < \infty)$$

for Triebel-Lizorkin spaces and

$$\lambda := n(\frac{1}{\min\{1,p\}} - 1) - s$$

for Besov spaces. Assume that $\phi_c \in \mathcal{D}^\lambda$. Suppose that ϕ_r satisfies the SF conditions of order $> s$ and the CAP system has an order $> s$. Then, for Triebel-Lizorkin spaces, we have

$$\sum_{v\in\{0,1\}^n} \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} (2^{js}|\langle f, (\psi_v)_{j,k}\rangle|\chi_{j,k})^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \lesssim \|f\|_{\dot{F}^s_{pq}(\mathbb{R}^n)}$$

and for Besov spaces, we have

$$\sum_{v\in\{0,1\}^n} \left(\sum_{j\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}^n} \left(2^{js} 2^{jn(1/2-1/p)} |\langle f, (\psi_v)_{j,k}\rangle| \right)^p \right)^{q/p} \right)^{1/q} \lesssim \|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)}.$$

3.3. Characterizations of $\dot{F}^s_{pq}$, $\dot{B}^s_{pq}$ in terms of CAP coefficients and CAMP coefficients

In order to use the characterizations in Propositions 3.11 and 3.13, we need to know the smoothness of the refinable function ϕ^{dual} (cf. (3.10) and (3.12)). This can be calculated using the formula in (2.9) for any given τ_c, τ_r, Θ and ξ. However, since the dual system was introduced mainly for streamlining theoretical/technical issues, it is desirable to seek characterization theorems that do not make any direct appeal to properties of that system.

In order to remove the dual system's role in the characterization, we use the following lemma whose proof is given in §3.5.

Lemma 3.15. *Fix a positive non-integer $\eta > 0$. Assume that $\phi_r \in \mathcal{R}^\eta$. Then for any $0 < \varepsilon < \eta$ such that $\eta - \varepsilon$ is non-integer, there exists ξ such that $\phi^{\mathrm{dual}} \in \mathcal{R}^{\eta-\varepsilon}$.*

Having Lemma 3.15 in hand, we are ready to prove Theorem 1.11 (from §1.5).

Proof of Theorem 1.11: From $\phi_r \in \mathcal{D}^\eta$ and from the definition of $\mathcal{D}^\eta$ (cf. Definition 3.3), we see that $\phi_r \in \mathcal{R}^{\eta'}$ for some non-integer $\eta' > \eta$. We choose $\varepsilon > 0$ such that $\eta' - \varepsilon$ is not an integer and that $\eta' - \varepsilon > \eta$. Then from Lemma 3.15, there exists ξ such that $\phi^{\mathrm{dual}} \in \mathcal{R}^{\eta'-\varepsilon}$, which in turn gives $\phi^{\mathrm{dual}} \in \mathcal{D}^\eta$. Now we use this ξ also for the dual wavelet mask t_v^{dual} constructions in (2.9). Then by Theorem 2.11 in §2.2, we know that the framelet system $X(\Psi^{\mathrm{dual}})$ associated with these wavelet masks and this refinable function ϕ^{dual} is a dual system to the CAPlet system $X(\Psi)$, i.e. the pair $(X(\Psi), X(\Psi^{\mathrm{dual}}))$ is a bi-framelet. □

We are ready to state and prove the characterization of Triebel-Lizorkin spaces $\dot{F}^s_{pq}$ in terms of the CAP coefficients $(d_{j,f})_{j\in\mathbb{Z}}$.

Theorem 3.16. *Let (h_c, h_r, h_a) be a CAP triplet, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions. Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \le \infty$ and assume that*

$$\phi_c \in \mathcal{D}^\lambda, \quad \phi_r \in \mathcal{D}^s, \quad \lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s.$$

Suppose that the CAP system has an order $\geq 1+\max\{\lfloor\lambda\rfloor, \lfloor s\rfloor\}$. Then, for every $f \in \dot{F}^s_{pq}$, we have

$$\left\| \left(\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} \left(2^{js}|d_{j,f}(k)|\chi_{\infty,j,k}\right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \approx \|f\|_{\dot{F}^s_{pq}(\mathbb{R}^n)}.$$

Remark. The finiteness of the left-hand side of the above expression may imply that $f \in \dot{F}^s_{pq}$, even when we do not know *a priori* that $f \in \dot{F}^s_{pq}$. We refer to §6.4 for details. □

Proof: We first recall, [R1], that a compactly supported refinable function g for which $\widehat{g}(0) \neq 0$ will satisfy the SF conditions of order m once it lies in $\mathcal{D}^{m-1}(\mathbb{R}^n)$. The "SF order" in the above statement is in the sense of (1.5), and, thanks to Assumption 1.7(d), it is valid also in the stronger (1.6) sense, for each of $g := \phi_c$, $g := \phi_r$. Thus we conclude that ϕ_c satisfies the SF conditions of order $1+\lfloor\lambda\rfloor$, while ϕ_r satisfies the SF conditions of order $1+\lfloor s\rfloor$.

Next, note that $\phi_r \in \mathcal{D}^{\max\{s,0\}}$ from the above assumption together with Assumption 1.7(c). Invoking Theorem 1.11, we select ξ such that $\phi^{\rm dual} \in \mathcal{D}^{\max\{s,0\}}$, and such that $(X(\Psi), X(\Psi^{\rm dual}))$ is a bi-framelet. Here, $\Psi := (\psi_\upsilon)_{\upsilon\in\{0,1\}^n}$ are the current mother CAPlets.

Now, we apply Proposition 3.11 (with respect to the current bi-framelet). Thanks to Lemma 2.2 we identify the CAPlet coefficient $\langle f, (\psi_\upsilon)_{j,k}\rangle$ with the CAP detail coefficient $2^{-(j+1)n/2}d_{j+1,f}(2k-\upsilon)$ in the characterization formula in that proposition. This gives, for every $f \in \dot{F}^s_{pq}$,

$$\|f\|_{\dot{F}^s_{pq}(\mathbb{R}^n)} = \sum_{\upsilon\in\{0,1\}^n} \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} \left(2^{js}|d_{j+1,f}(2k-\upsilon)|\chi_{\infty,j,k}\right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} =: \alpha.$$

For $\upsilon \in \{0,1\}^n$, $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^n$, we define $Q^\upsilon_{j,k} := 2^{-(j+1)}(2k + k_0 - \upsilon + [0,1]^n)$, where $k_0 := (1,1,\cdots,1) \in \mathbb{R}^n$. Then, with $\mathrm{vol}(Q) :=$ the volume of Q,

$$Q^\upsilon_{j,k} \subset 2^{-j}(k+[0,1]^n), \quad \frac{\mathrm{vol}(2^{-j}(k+[0,1]^n))}{\mathrm{vol}(Q^\upsilon_{j,k})} = 2^n =: N. \tag{3.17}$$

Now we claim that, for each $\upsilon \in \{0,1\}^n$,

$$\left\| \left(\sum_{j,k} \left(2^{js}|d_{j+1,f}(2k-\upsilon)|\chi_{Q^\upsilon_{j,k}}\right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \approx \left\| \left(\sum_{j,k} \left(2^{js}|d_{j+1,f}(2k-\upsilon)|\chi_{\infty,j,k}\right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)},$$

where $\chi_{Q^\upsilon_{j,k}}$ denotes the characteristic function of $Q^\upsilon_{j,k}$. Since $\chi_{Q^\upsilon_{j,k}} \leq \chi_{\infty,j,k}$, the inequality $\lesssim$ is trivial. For the other direction, we use (3.17) to obtain that, with M_t the maximal operator defined in (6.10), for any $t > 0$ and $j \in \mathbb{Z}$, $k \in \mathbb{Z}^n$,

$$\chi_{\infty,j,k} \leq N^{1/t} M_t(\chi_{Q^\upsilon_{j,k}}),$$

and invoke (6.11) using t with $0 < t < \min\{p,q\}$. Since

$$\chi_{Q^\upsilon_{j,k}} = \chi(2^{j+1}\cdot -(2k-\upsilon+k_0)),$$

we get

$$\alpha \approx \sum_{\upsilon\in\{0,1\}^n} \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{k'\in 2\mathbb{Z}^n+k_0-\upsilon} \left(2^{js}|d_{j+1,f}(k'-k_0)|\chi_{\infty,j+1,k'}\right)^q \right)^{\frac{1}{q}} \right\|_{L_p}$$

$$\approx \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{\upsilon\in\{0,1\}^n} \sum_{k'\in 2\mathbb{Z}^n+k_0-\upsilon} \left(2^{js}|d_{j+1,f}(k'-k_0)|\chi_{\infty,j+1,k'}\right)^q \right)^{\frac{1}{q}} \right\|_{L_p}$$

$$\approx \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} \left(2^{js}|d_{j,f}(k-k_0)|\chi_{\infty,j,k}\right)^q \right)^{\frac{1}{q}} \right\|_{L_p} \approx \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} \left(2^{js}|d_{j,f}(k)|\chi_{\infty,j,k}\right)^q \right)^{\frac{1}{q}} \right\|_{L_p}.$$

Here, the last equivalence is obtained by applying Proposition 6.9 with $\sigma := \frac{n}{s}$, once to the matrix $\mathbf{M} := (\mathbf{M}(j,k;l,m) : j,l \in \mathbb{Z}, k,m \in \mathbb{Z}^n)$ defined by

$$\mathbf{M}(j,k;l,m) := \begin{cases} 1, & l = j,\ m = k - k_0, \\ 0, & \text{otherwise,} \end{cases}$$

and once to its inverse. (Note that $d_{j,f}(k-k_0) = (\mathbf{M}d)(j,k)$, where $d := (d_{j,f} : j \in \mathbb{Z})$.)

For later use, we note that the fact that $f \in \dot{F}^s_{pq}$ was not used in our estimation of α. □

Next, we record the characterization of Besov spaces $\dot{B}^s_{pq}$ in terms of the CAP coefficients $(d_{j,f})_{j\in\mathbb{Z}}$. The proof of this result is omitted, since it employs a proper subset of the arguments used in the proof of its Triebel-Lizorkin's counterpart.

Theorem 3.18. *Let (h_c, h_r, h_a) be a CAP triplet, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions. Let $s \in \mathbb{R}$, $0 < p, q \le \infty$ and assume that*

$$\phi_c \in \mathcal{D}^\lambda, \quad \phi_r \in \mathcal{D}^s, \quad \lambda := n(\frac{1}{\min\{1,p\}} - 1) - s.$$

Suppose that the CAP system has an order $\ge 1 + \max\{\lfloor\lambda\rfloor, \lfloor s\rfloor\}$. Then, for every $f \in \dot{B}^s_{pq}$, we have

$$\left(\sum_{j\in\mathbb{Z}} \left(2^{-jn} \sum_{k\in\mathbb{Z}^n} |2^{js} d_{j,f}(k)|^p \right)^{q/p} \right)^{1/q} \approx \|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)}.$$

□

In case $s \ge 0$ and $p, q \ge 1$, the value of λ in the above theorems is non-positive, which allows us to simplify the assumed conditions. The corollary covers the characterizations of the Sobolev space W^s_p ($s > 0$, $1 < p < \infty$), since this space is equivalent to the intersection of two Triebel-Lizorkin spaces, $\dot{F}^0_{p2}$ and $\dot{F}^s_{p2}$.

Corollary 3.19. *Let (h_c, h_r, h_a) be a CAP triplet, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions. Given $s \geq 0$, and $1 \leq p, q \leq \infty$, assume that $\phi_r \in \mathcal{D}^s$, and that the CAP system is of order $> s$. Then the characterizations in Theorem 3.16 (for $p < \infty$) and Theorem 3.18 remain valid.*

In Section 2.3, we observed that for a pseudo-interpolatory prediction filter h_r (see (2.12) for the definition), the CAP coefficients and the CAMP coefficients are norm-equivalent in any of the discrete norms that are used in the Triebel-Lizorkin and Besov spaces characterizations (cf. Lemma 2.15 and 2.16). However, in case of CAMP, we cannot assume that the refinable function ϕ_r, associated with the pseudo-interpolatory filter h_r, satisfies Assumption 1.7(d). Instead we assume the refinable function $\phi_{\rm in}$ associated with the interpolatory $h_{\rm in}$ (which is a factor of h_r) satisfies the assumption. With that in mind, we conclude the following:

Theorem 3.20. *Let (h_c, h_r, h_a) be a CAP triplet with a pseudo-interpolatory h_r, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions. Let $s \in \mathbb{R}$, $0 < p, q \leq \infty$. Assume that*

$$\phi_c \in \mathcal{D}^\lambda, \quad \phi_r \in \mathcal{D}^s,$$

where

$$\lambda := n(\frac{1}{\min\{1, p, q\}} - 1) - s, \quad (p < \infty)$$

for Triebel-Lizorkin spaces and

$$\lambda := n(\frac{1}{\min\{1, p\}} - 1) - s$$

for Besov spaces. Let $\phi_{\rm in}$ be the refinable function associated with $h_{\rm in}$. Suppose that $\phi_{\rm in}$ satisfies the SF conditions of order $> s$ and that the order of the CAP system is $\geq 1 + \max\{\lfloor\lambda\rfloor, \lfloor s\rfloor\}$. Then the characterizations in Theorem 3.16 (for $p < \infty$) and in Theorem 3.18 remain valid with the CAP coefficients $(d_{j,f})_{j\in\mathbb{Z}}$ replaced by the CAMP coefficients $(\widetilde{d}_{j,f})_{j\in\mathbb{Z}}$.

3.4. How to construct effectively a CAMP system?

Let us assume that $s \geq 0$ and $p, q \geq 1$ in Theorem 3.20. In that case, ϕ_c needs merely to be minimally smooth, something that we already assume in Assumptions 1.7. The two important conditions to satisfy then are: (i) the interpolatory function $\phi_{\rm in}$ should satisfy SF conditions of order $> s$, and (ii) the CAMP system should have an order $> s$, viz, $\Theta - \Theta(2\cdot)\tau_r\overline{\tau}_c = O(|\cdot|^{1+\lfloor s\rfloor})$ near the origin. The assumption that ϕ_r is smooth, which is the most demanding one in standard wavelet and CAP characterizations, is secondary here: we can enforce it by using enhancement. We discuss now in more detail the roles of enhancement and alignment in CAMP constructions.

Enhancement. The enhancement filter $h_{\rm e}$ plays the least important role in the CAMP construction. Its mere role is to make the refinable function ϕ_r smoother. It does not affect the order m of the SF conditions that this function satisfies. As is well-known in wavelet theory (cf. [R1]), the smoothness of the refinable function is usually smaller, sometimes much smaller, than the above m; this is definitely true for interpolatory refinable functions, like our $\phi_{\rm in}$. The enhancement filter can help in bringing the smoothness of ϕ_r to be in line with its SF conditions (since we assume $h_{\rm e}$ to be supported on even integers, the convolution with $\phi_{\rm e}$ cannot improve the SF order. I.e., the SF order of ϕ_r coincides with that of $\phi_{\rm in}$).

In view of the above, the enhancement function $\phi_{\rm e}$ should always be chosen to be a *spline* (since splines deliver the highest smoothness per filter size). It is worthwhile to mention that, although the smoothness of ϕ_r plays a central role in the function space characterizations in the previous subsection, it may be less significant in specific applications (e.g., all applications whose performance analysis does not require Bernstein-type inequalities). While enhancement may be unnecessary in some applications, the cost of enhancement is quite low: only one of the CAMP filters is altered once enhancement is introduced. In a slightly exaggerated language we could thus say that "Bernstein-type inequalities are almost free in CAMP representations", namely, they can be obtained at the cost of additional space-blurring of only a small portion of the detail coefficients.

Alignment. The alignment filter (in both CAP and CAMP constructions) has the single role of bridging "compatibility gaps" between the operator C and P (and $H_{\rm e}$, in case enhancement is used), so that the CAP system will have the requisite order $> s$. In the context of CAP representation, the use of alignment is almost always a better choice than modifying C or P. This can be illustrated by the following 1D example. Suppose that ϕ_c is supported in $[0, K_0]$ for some K_0. Suppose that by modifying h_c we can obtain a new refinable function φ with support $[0, K_1]$, and that in this way we can bring the order of the CAP system to required level. Then, as we observed in the Discussion after Theorem 1.12 in §1.6, we can achieve the same task by keeping ϕ_c intact, and employing instead an alignment filter with $K_1 - K_0 + 1$ non-zero coefficients. At the same time, when using alignment without altering ϕ_c we keep its smoothness (and perhaps some other desired properties) intact.

The CAMP situation is more complicated. One possible claim is that the introduction of alignment degrades the space localization of *all* the detail coefficients, while, in contrast, bridging the compatibility gap between C and P by modifying C alters only one of the CAMP filters. That claim might be countered by the observation that changing C changes not only the detail filters but also the entire MRA hierarchy. However, another claim is that it should be faster to compute the detail coefficients without alignment (i.e., by modifying h_c), since the compression filter and its size are irrelevant to the vast majority of detail coefficients, while the alignment filter is employed for the computations of *all* the detail coefficients.

Finally, enhancement might offer a surprising alternative to alignment: instead of using enhancement in order to obtain a smoother ϕ_r, we may choose the enhancement filter in order to improve the order of the CAP (hence CAMP) system. This will lead to an expected *reduction* in the smoothness of ϕ_r, which may or may not be tolerable. On the upside, we are able in this way to increase of the order of the system without changing the MR $(y_j)_j$, and without changing any of the detail coefficients $d_j(k)$, $k \in \mathbb{Z}^n \backslash (2\mathbb{Z}^n)$. This is particularly attractive for applications that require only Jackson-type inequalities, i.e., vanishing moments. □

CAMPlets. Let us now look more closely at the definition of the CAMPlets, and assume that the filters $h_{\rm e}$ and h_a are symmetric. Assume integer scalar dilation $\Lambda = \lambda I$, $\lambda > 1$. The CAMPlet ψ_0^M is particularly simple:

$$\psi_0^M = (h_a * \phi_c)_{1,0} - \lambda^{-n/2}(h_{\rm e} * h_a * \phi_c),$$

which, for the case $h_a = h_{\rm e} = \delta$, yields $\psi_0^M = (\phi_c)_{1,0} - \lambda^{-n/2}\phi_c$.

All the other CAPMlets are obtained by the following simple algorithm, in which we assume $h_{\rm in}$ to be symmetric.

(a) Compute the restriction h_υ of $h_{\rm in}$ to $\upsilon + \lambda\mathbb{Z}^n$ (i.e., $h_\upsilon := E^{-\upsilon}(((E^\upsilon h_{\rm in})_\downarrow)_\uparrow)$.)
(b) Convolve ϕ_c with the filter h_υ, then multiply by λ^n.
(c) Subtract the result of (b) from ϕ_c.
(d) Compute $E^{-\upsilon}g$, where g is the function obtained in (c).

(e) Convolve the result of (d) with h_a.
(f) Dilate the result of (e) by λ and then multiply by $\lambda^{n/2}$.

In case $h_a = \delta$, the support of the CAMPlet must then be a subset of $\upsilon + \Omega/\lambda$, with

$$\Omega := \mathrm{supp}(|h_{\mathrm{in}}| * |\phi_c|).$$

If either λ is large, and/or n is large, then the filter h_υ may inherit only a tiny portion of the non-zero coefficients of h_{in}. In this case, ψ_υ^M is supported only in a small subset of $\upsilon + \Omega/\lambda$, albeit its support is not convex. □

3.5. Proof of Lemma 3.15

To prove Lemma 3.15, we use a result from [R2] concerning the smoothness of refinable functions. We list below a special and simplified version of that result, that suffices for our purposes.

Result 3.21. *Let h be a finite low-pass filter and let ϕ be the associated refinable function. Assume that $\widehat{\phi}(0) = 1$. For a non-negative integer j, denote $\mathcal{Q}_j := \ell_\infty(\mathbb{Z}^n/2^j)$. The 0-extension of h from $\mathbb{Z}^n$ to $\mathbb{Z}^n/2^j$ defines an element of $\mathcal{Q}_j$, which we still denote by h. Let $\mathcal{C}_h : \mathcal{Q}_j \to \mathcal{Q}_{j+1}$ be the cascade operator*

$$\mathcal{C}_h u(j) := (h * u)(2j).$$

Fix $r > 0$. Then there exists a finite set $V \subset \mathcal{Q}_0$ of finitely supported sequences such that, for any $0 < \eta \le r$, the following two properties are equivalent:
(i) $\phi \in \mathcal{R}^\alpha(\mathbb{R}^n)$ for every $\alpha < \eta$.
(ii) For every $v \in V$, and every $\alpha < \eta$, $\|\mathcal{C}_h^j v\|_{\mathcal{Q}_j} = O(2^{-j\alpha})$.
Moreover, suppose that we create a new filter $h_\xi = h * \xi$, with ξ finite and $\widehat{\xi}(0) = 1$. Then V above can be chosen so that the implication (ii) $\Longrightarrow$ (i) holds with respect to h_ξ, too, i.e., with h and $\mathcal{C}_h$ replaced by h_ξ and $\mathcal{C}_{h_\xi}$, and with ϕ replaced by the refinable function ϕ_ξ associated with h_ξ. □

Proof of Lemma 3.15: We note first that, for any filter h, the jth power of the cascade operator, as an operator from $\mathcal{Q}_0$ to $\mathcal{Q}_j$, trivially satisfies the bound $\|\mathcal{C}_h^j\| \le \|h\|_{\ell_1(\mathbb{Z}^n)}^j$. This implies that, for any two filters h_1, h_2, and any $v \in \mathcal{Q}_0$,

$$\|\mathcal{C}_{h_1 * h_2}^j v\|_{\mathcal{Q}_j} \le \|h_2\|_{\ell_1(\mathbb{Z}^n)}^j \|\mathcal{C}_{h_1}^j v\|_{\mathcal{Q}_j}. \tag{3.22}$$

As for the proof itself, we first write

$$t^{\mathrm{dual}} = \tau_r \left(1 + \xi \left(1 - \sum_{\upsilon^* \in \{0,\pi\}^n} \left(\frac{\Theta(2\cdot)}{\Theta} \overline{\tau}_c \tau_r\right)(\cdot + \upsilon^*)\right)\right) =: \tau_r(1 + \xi\nu_0).$$

Note that ν_0 is a rational polynomial. We assume that ξ can be factored $\xi = \zeta\nu_1$, such that ζ and ν_1 are polynomials, and, moreover, $\nu := \nu_1\nu_0$ is a polynomial, too. Thus,

$$t^{\mathrm{dual}} = \tau_r(1 + \zeta\nu) =: \tau_r\rho.$$

Next, we invoke Result 3.21, with respect to $\phi := \phi_r$. Fix $\alpha < \eta$. Since we assume that $\phi_r \in \mathcal{R}^\eta(\mathbb{R}^n)$, we obtain, for V as in Result 3.21 and $v \in V$,

$$\|\mathcal{C}_{h_r}^j v\|_{\mathcal{Q}_j} = O(2^{-j\alpha}).$$

Let h_ρ and h^{dual} be the Fourier coefficients of ρ and t^{dual} respectively. Set $\delta := \|h_\rho\|_{\ell_1} - 1$. Then, since $h^{\mathrm{dual}} = h_r * h_\rho$, we conclude from (3.22) that

$$\|\mathcal{C}^j_{h^{\mathrm{dual}}} v\|_{\mathcal{Q}_j} \leq \|h_\rho\|^j_{\ell_1(\mathbb{Z}^n)} O(2^{-j\alpha}) = O(2^{-j(\alpha - \log_2(1+\delta))}).$$

Result 3.21 now applies to yield that, by changing V if needed, we can conclude from the above that $\phi^{\mathrm{dual}} \in \mathcal{R}^{\eta-\varepsilon}$ provided that $\varepsilon > \log_2(1+\delta)$.

We are thus left to show that we can make the above δ as small as we wish. Since $\rho = 1 + \zeta\nu$, it suffices to show that we can choose ζ such that the ℓ_1-norm of the coefficients of $\zeta\nu$ is $< \delta$. Note that our only constraint on ζ is that $\zeta(0) = c$ for some non-zero and fixed c. The only property of ν that we will use is that ν has a double zero at the origin. (This follows from Assumptions 1.7(c,d,f)).

We first observe the following : for $A(\omega) := \sum_{k\in\mathbb{Z}^n} a(k)e^{ik\cdot\omega}$, and with N the smallest integer larger than $\frac{n}{2}$,

$$\sum_{k\in\mathbb{Z}^n\backslash 0} |a(k)| \lesssim \sum_{|\beta|=N} \|A^{(\beta)}\|_{L_2(\mathbb{T}^n)}.$$

Indeed,

$$\begin{aligned}\sum_{k\in\mathbb{Z}^n\backslash 0} |a(k)| &\leq \left(\sum_{k\in\mathbb{Z}^n\backslash 0} \frac{1}{|k|^{2N}}\right)^{1/2} \left(\sum_{k\in\mathbb{Z}^n} |k|^{2N}|a(k)|^2\right)^{1/2} \\ &\lesssim \left(\sum_{k\in\mathbb{Z}^n} |k|^{2N}|a(k)|^2\right)^{1/2} \lesssim \sum_{|\beta|=N} \|A^{(\beta)}\|_{L_2(\mathbb{T}^n)},\end{aligned}$$

where the first inequality is from Hölder inequality and the last inequality is from Parseval's identity.

Note that ν has zero of order 2 at the origin. Let $\kappa : \mathbb{R}^n \to \mathbb{R}$ be a function that vanishes on $2\pi\mathbb{Z}^n\backslash 0$, and satisfies, for some $n_0 \in \mathbb{Z}$ and some $M > 2 + \frac{n}{2}$,

$$|\kappa^{(\gamma)}(\omega)| \lesssim \begin{cases} 1, & |\omega| \leq 2^{n_0}, \\ |\omega|^{-M}, & |\omega| \geq 2^{n_0}, \end{cases} \quad \text{for all } |\gamma| \leq N.$$

For example, we can take κ to be the tensor product of the Fourier transforms of B-splines of order M. Let $\zeta := c_0 \sum_{k\in\mathbb{Z}^n} \kappa(2^l(\cdot + 2\pi k))$, for $l \in \mathbb{N}$ with $2^l\pi \geq 2^{n_0+1}$. Note the $\zeta(0)$ is independent of l. Now, the series that defines ζ converges absolutely for all $\omega \in \mathbb{T}^n$, and, for any $|\gamma| \leq N$,

$$\begin{aligned}|\zeta^{(\gamma)}(\omega)| &\leq c_0 2^{l|\gamma|}\left(|\kappa^{(\gamma)}(2^l\omega)| + \sum_{k\in\mathbb{Z}^n\backslash 0} |\kappa^{(\gamma)}(2^l(\omega + 2\pi k))|\right) \\ &\lesssim c_0 2^{l|\gamma|} \begin{cases} |2^l\omega|^{-M} + \sum_{k\in\mathbb{Z}^n\backslash 0} |2^l(\pi k)|^{-M}, & |\omega| \geq 2^{n_0-l}, \\ 1 + \sum_{k\in\mathbb{Z}^n\backslash 0} |2^l(\pi k)|^{-M}, & |\omega| \leq 2^{n_0-l}, \end{cases} \\ &\lesssim c_0 2^{l|\gamma|} \begin{cases} |2^l\omega|^{-M}, & |\omega| \geq 2^{n_0-l}, \\ 1, & |\omega| \leq 2^{n_0-l}. \end{cases}\end{aligned}$$

Thus, with $\Omega_m := \{\omega \in \mathbb{T}^n : 2^{-m-1} \leq |\omega| \leq 2^{-m}\}$, $m \geq m_0$, we have

$$|\zeta^{(\gamma)}(\omega)| \lesssim 2^{l|\gamma|} \begin{cases} 2^{(l-m)(-M)}, & m \leq l - n_0, \\ 1, & m \geq l - n_0, \end{cases} \quad \omega \in \Omega_m.$$

Thus, for any $|\gamma| + |\gamma'| \le N$, since $|\nu^{(\gamma')}(\omega)| \lesssim |\omega|^{2-|\gamma'|}$, we have

$$
\begin{aligned}
\|\zeta^{(\gamma)}\, \nu^{(\gamma')}\|^2_{L_2(\mathbb{T}^n)} &= \sum_{m_0 \le m < \infty} \int_{\Omega_m} |\zeta^{(\gamma)}(\omega)|^2 |\nu^{(\gamma')}(\omega)|^2 d\omega \\
&\lesssim 2^{2l|\gamma|} \left(2^{-2lM} \sum_{m_0 \le m \le l-n_0} 2^{2mM} 2^{-2m(2-|\gamma'|)} 2^{-mn} + \sum_{l-n_0 \le m \le \infty} 2^{-2m(2-|\gamma'|)} 2^{-mn} \right) \\
&= 2^{2l|\gamma|} \left(2^{-2lM} \sum_{m_0 \le m \le l-n_0} 2^{m(2M-4+2|\gamma'|-n)} + \sum_{l-n_0 \le m \le \infty} 2^{m(-4+2|\gamma'|-n)} \right) \\
&\lesssim 2^{l(2|\gamma|-2M+2M-4+2|\gamma'|-n)} + 2^{l(2|\gamma|-4+2|\gamma'|-n)} \approx 2^{l(2(|\gamma|+|\gamma'|)-4-n)},
\end{aligned}
$$

since $2(|\gamma| + |\gamma'|) - 4 - n = 2N - 4 - n < 0$. Note that the constants used in $\lesssim$ and $\approx$ of the above estimation do not depend on l. Thus we get

$$
\sum_{|\beta|=N} \|(\zeta\nu)^{(\beta)}\|^2_{L_2(\mathbb{T}^n)} \approx \sum_{|\beta|=N} \sum_{\gamma+\gamma'=\beta} \|\zeta^{(\gamma)}\, \nu^{(\gamma')}\|^2_{L_2(\mathbb{T}^n)} \lesssim c\, 2^{l(2N-4-n)},
$$

and we can make the left side as small as we wish, by choosing l large enough. $\square$

4. Nonlinear approximation

There are several standard applications to the characterizations that were established in the previous section. As an illustration, we describe in this section one such application: a Jackson-type inequality in best-term non-linear approximation. It is well-known that Jackson-type inequalities require far less than the complete function space characterizations for their derivation. To this end, we state the following a simple corollary of Corollary 3.14:

Corollary 4.1. *Let (h_c, h_r, h_a) be a CAP triplet, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions (not necessarily satisfying Assumption 1.7(d)). Let $s \in \mathbb{R}$, $0 < p \le \infty$, $0 < q \le \infty$ and assume that*

$$
\phi_c \in \mathcal{D}^\lambda, \quad \lambda := n(\frac{1}{\min\{1,p\}} - 1) - s.
$$

Suppose that ϕ_r satisfies the SF conditions of order $> s$ and the CAP system has an order $> s$. Then

$$
\left(\sum_{j\in\mathbb{Z}} \left(2^{-jn} \sum_{k\in\mathbb{Z}^n} |2^{js} d_{j,f}(k)|^p \right)^{q/p} \right)^{1/q} \lesssim \|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)}
$$

(with the usual modification for $p = \infty$ or $q = \infty$).

To see this, we only need to recall Corollary 3.14 and replace the framelet coefficients there by the CAP coefficients using (2.3).

For $1 \le p < \infty$, the error of the best m-term approximation of $f \in L_p(\mathbb{R}^n)$ from $X(\phi_r)$ is defined as

$$
\sigma_m(f)_p := \inf_{g\in\Sigma_m} \|f - g\|_{L_p(\mathbb{R}^n)},
$$

where Σ_m denotes the set of linear combinations of $\{g : g \in X(\phi_r)\}$ with at most m nonzero coefficients. Then we have the following Jackson-type inequality :

Theorem 4.2. *Let $1 \le p < \infty$, $s > 0$ and set $\tau := (\frac{s}{n} + \frac{1}{p})^{-1}$. Also let (h_c, h_r, h_a) be a CAP triplet, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions (not necessarily satisfying Assumption 1.7(d)). Suppose that ϕ_r satisfies the SF conditions of order $> s$ and the CAP system has an order $> s$. Then for every $f \in \dot{B}^s_{\tau\tau}(\mathbb{R}^n) \cap L_p(\mathbb{R}^n)$,*

$$\sigma_m(f)_p \le c\,\|f\|_{\dot{B}^s_{\tau\tau}(\mathbb{R}^n)} m^{-\frac{s}{n}}.$$

The *proof* of the theorem follows almost *verbatim* similar arguments in the literature. We present its main ingredients, and refer to [DJP] for some of the more technical points.

Proof: We first renormalize the CAP coefficients and ϕ_r as follows:

$$d_{f,p}(j,k) := 2^{-jn/p} d_{j,f}(k), \quad (\phi_r)_{p,j,k} := 2^{jn(\frac{1}{p}-\frac{1}{2})}(\phi_r)_{j,k}.$$

From the Assumptions 1.7, we have ([Me1])

$$f = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} d_{f,p}(j,k)(\phi_r)_{p,j,k},$$

with the convergence above valid in L_p. Next, we observe that, since we assume $p \ge 1$ and $s \ge 0$, we have that $\lambda \le 0$ in Corollary 4.1, and hence the assumption $\phi_c \in \mathcal{D}^\lambda$ in that corollary is redundant. Hence we can invoke the corollary, which, together with the fact that $\frac{s}{n} = \frac{1}{\tau} - \frac{1}{p}$, implies that, with some constant $A := A(p,\tau,n)$,

$$\|d_{f,p}\|_{\ell_\tau(\mathbb{Z}\times\mathbb{Z}^n)} := \left(\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} |d_{f,p}(j,k)|^\tau\right)^{1/\tau} = \left(\sum_{j\in\mathbb{Z}} 2^{-jn} \sum_{k\in\mathbb{Z}^n} |2^{js} d_{j,f}(k)|^\tau\right)^{1/\tau} \le A\|f\|_{\dot{B}^s_{\tau\tau}(\mathbb{R}^n)}.$$

By renormalizing f, if necessary, we obtain that $\|f\|_{\dot{B}^s_{\tau\tau}(\mathbb{R}^n)} = 1/A$. With this assumption, $\|d_{f,p}\|_{\ell_\tau(\mathbb{Z}\times\mathbb{Z}^n)} \le 1$ and hence there are at most m coefficients for which $|d_{f,p}(j,k)|^\tau \ge \frac{1}{m}$. Now for each $j \in \mathbb{Z}$, let $K_j := \{k \in \mathbb{Z}^n : |d_{f,p}(j,k)|^\tau \ge \frac{1}{m}\}$ and let $S := \sum_{j\in\mathbb{Z}} \sum_{k\in K_j} d_{f,p}(j,k)(\phi_r)_{p,j,k}$. Then $S \in \Sigma_m$ and the error $E := f - S$ is given as

$$E = \sum_{j\in\mathbb{Z}} \sum_{k\notin K_j} d_{f,p}(j,k)(\phi_r)_{p,j,k} = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} d^+_{f,p}(j,k)(\phi_r)_{p,j,k},$$

where $d^+_{f,p}(j,k) := \begin{cases} d_{f,p}(j,k), & k \notin K_j, \\ 0, & k \in K_j. \end{cases}$ Note that $|d^+_{f,p}(j,k)| \le \varepsilon$ with $\varepsilon := m^{-1/\tau}$.

Now, with M_t the maximal operator defined in (6.10), we have that, for $0 < t < 1$, since ϕ_r has compact support,

$$|\phi_r| \le c\,M_t(\chi),$$

hence, for all $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^n$

$$|(\phi_r)_{p,j,k}| \le c\,M_t(\chi_{p,j,k}), \quad \chi_{p,j,k} := 2^{jn(\frac{1}{p}-\frac{1}{2})}\chi_{j,k}.$$

If we let $\widetilde{E} := \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} d^+_{f,p}(j,k)\chi_{p,j,k}$, then by (6.11) we have $\|E\|_{L_p(\mathbb{R}^n)} \le c\,\|\widetilde{E}\|_{L_p(\mathbb{R}^n)}$. Now we claim that

$$\|\widetilde{E}\|^p_{L_p(\mathbb{R}^n)} \le c\,\|d_{f,p}\|^\tau_{\ell_\tau(\mathbb{Z}\times\mathbb{Z}^n)} \varepsilon^{p-\tau}.$$

This claim is easy to prove when $p = 1$, since

$$\#\{(j,k) \in \mathbb{Z} \times \mathbb{Z}^n : |d_{f,1}(j,k)| \geq 2^{-l}\varepsilon\} \leq 2^{l\tau}\varepsilon^{-\tau}\|d_{f,1}\|^\tau_{\ell_\tau(\mathbb{Z}\times\mathbb{Z}^n)},$$

for any number l. Therefore, since $\tau < p = 1$,

$$\begin{aligned}\|\widetilde{E}\|_{L_1(\mathbb{R}^n)} &\leq \sum_{\{(j,k):|d_{f,1}(j,k)|<\varepsilon\}} |d_{f,1}(j,k)| = \sum_{l=1}^{\infty} \sum_{\{(j,k):2^{-l}\varepsilon\leq|d_{f,1}(j,k)|<2^{-l+1}\varepsilon\}} |d_{f,1}(j,k)| \\ &\lesssim \|d_{f,1}\|^\tau_{\ell_\tau(\mathbb{Z}\times\mathbb{Z}^n)} \sum_{l=1}^{\infty}(2^{-l}\varepsilon)2^{l\tau}\varepsilon^{-\tau} \lesssim \varepsilon^{1-\tau}\|d_{f,1}\|^\tau_{\ell_\tau(\mathbb{Z}\times\mathbb{Z}^n)}.\end{aligned}$$

For $p > 1$, a similar, alas far more technical, argument is required. We refer to [DJP: p.745] for details. Altogether, we obtain

$$\sigma_m(f)_p \leq c\,\|\widetilde{E}\|_{L_p(\mathbb{R}^n)} \leq c\,\varepsilon^{1-\tau/p} = c\,\varepsilon^{\frac{s}{n}\tau} = c\,m^{-\frac{s}{n}}.$$

□

5. Examples of CAP and CAMP representations

Our objective in this section is to illustrate the CAP and CAMP representations with the aid of concrete examples. So, the main point here is not to develop "ideal" CAP/CAMP systems, but to show how the various conditions that appear in the characterization theorems compete one with another, and what the trade-off is among the various constructions that aim at the same performance.

We assess the *performance* of a given CAP/CAMP system using two different methods. The basic performance grade is given in terms of

$$s_a := \min\{s_{sf}, s_{or}\},$$

with s_{or} the order to the CAP system, and s_{sf} the order of the SF conditions satisfied by ϕ_r. As we pointed out earlier in §3.2, s_a captures correctly the order of the vanishing moments of the CAPlets. This is the right performance grade of the system when considering Jackson-type inequality such as Theorem 4.2 from the previous section.

Our other, and more demanding, performance level is the value s_b that is listed in the function space characterizations from §3.3. We assume for simplicity that we are trying to characterize the Triebel-Lizorkin spaces $\dot{F}^s_{pq}$ and the Besov spaces $\dot{B}^s_{pq}$ with $p, q \geq 1$ and $s \geq 0$. Here, as we observed in Corollary 3.19,

$$s_b := \min\{s_{sm}, s_{or}\},$$

with s_{sm} being the Hölder regularity of the prediction refinable function ϕ_r:

$$\phi_r \in \mathcal{R}^\alpha(\mathbb{R}^n) \quad \text{for every } \alpha < s_{sm} \tag{5.1}$$

(cf. Definition 3.3). Note that we always have $s_b \leq s_a$, since $s_{sm} \leq s_{sf}$, [R1].

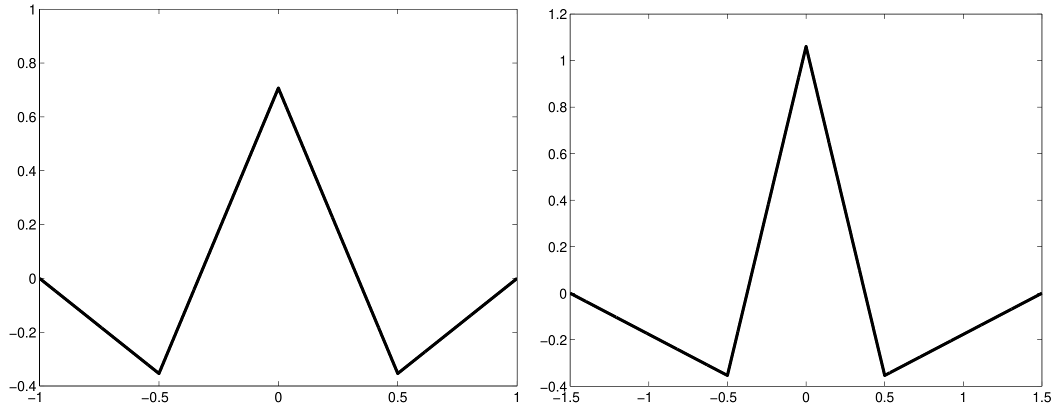

Figure 1. The graphs of ψ^* (left) and ψ^{**} (right).

5.1. Univariate examples

We start by displaying the graphs of two functions that appear in several of our examples. One is denoted as ψ^* and the other as ψ^{**}. Both have two vanishing moments.

5.1.1 CAP and CAMP for $s_a = s_b + 1 = 2$

In view of the discussion above, constructions in this class require a CAP system of order 2, with a prediction function ϕ_r that satisfies the SF conditions of order 2 and has regularity 1 in the sense of (5.1).

Example 5.2: piecewise-linear CAP. We choose ϕ_c and ϕ_r to be the centered hat function. Then $\tau_c(\omega) = \tau_r(\omega) = \cos^2(\frac{\omega}{2})$ and $1 - \tau_r\overline{\tau}_c = O(|\cdot|^2)$ at 0, hence the CAP is of order 2. Since $s_{sm} = 1$, and ϕ_r satisfies the SF conditions of order 2, the system fits the required conditions. The CAPlet masks are

$$\sqrt{2}t_0(\omega) = 1 - \tau_c(\omega)(\overline{\tau}_r(\omega) + \overline{\tau}_r(\omega+\pi)) = \sin^2(\frac{\omega}{2}),$$
$$\sqrt{2}e^{-i\omega}t_1(\omega) = 1 - \tau_c(\omega)(\overline{\tau}_r(\omega) - \overline{\tau}_r(\omega+\pi)) = \sin^2(\frac{\omega}{2})(2+\cos\omega).$$

The filters of (τ_c, t_0, t_1) are $(3,3,5)$-tap, respectively. The CAPlets are $\psi_0 = \psi^*$ and $\psi_1 = E^{-1/2}\psi^{**}$. The support intervals are of lengths 2 and 3, respectively. □

Example 5.3: piecewise-linear CAMP. Since ϕ_r in Example 5.2 is interpolatory, we derive an associated CAMP from it. This leads to a modified version, t_1^M, of t_1 as follows:

$$\sqrt{2}e^{-i\omega}t_1^M(\omega) = 1 - (\overline{\tau}_r(\omega) - \overline{\tau}_r(\omega+\pi)) = 2\sin^2(\frac{\omega}{2}). \tag{5.4}$$

Each of the filters of (τ_c, t_0, t_1^M) is now 3-tap. The new CAPMlet ψ_1^M is $2(E^{-1/2}\psi^*)$. □

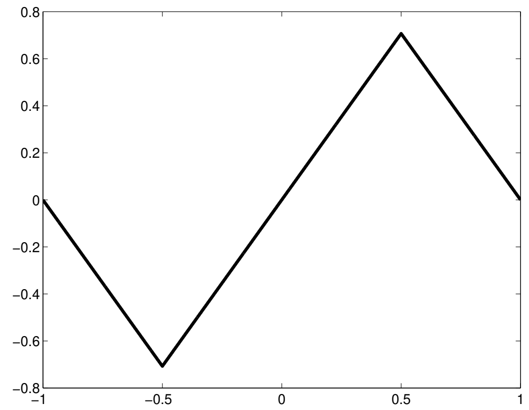

Figure 2. The second wavelet of the piecewise-linear construction of [RS].

Comparison with piecewise-linear tight frames. Next, we compare our CAMP construction with the piecewise linear construction of [RS], and of [DHRS]. Both constructions are of tight frames, while ours is not. They use the same refinable function as here. The framelet filters in [RS] are

$$\tau_1(\omega) := \sin^2(\frac{\omega}{2}), \quad \tau_2(\omega) := -\frac{\sqrt{2}}{2} i \sin(\omega).$$

The first wavelet here coincides with $\sqrt{2}\psi^*$, and the second is depicted in Figure 2. As in our CAMP above, the mother wavelets are supported, each, in an interval of length 2. However, this framelet system has 1 vanishing moment hence only satisfies $s_a = s_b = 1$.

The framelet filters of the system in [DHRS] are

$$\tau_1(\omega) := \sin^2(\frac{\omega}{2}), \quad \tau_2(\omega) := \frac{\sqrt{6}}{3} \sin^2(\frac{\omega}{2})(2 + \cos\omega)$$

and $(\tau_i)_{i=0}^2$ satisfies the OEP with the alignment filter $\Theta(\omega) := (4 - \cos(\omega))/3$. This system has 2 vanishing moments, but its second framelet is supported in a longer interval (viz., $[-1.5, 1.5]$), than our second CAMPlet. That second framelet is $\sqrt{\frac{4}{3}}\psi^{**}$. The first framelet is, again, $\sqrt{2}\psi^*$. □

Example 5.5: still piecewise-linear, but larger dilation. In order to reduce the redundancy of the above CAP/CAMP representations, we may use CAMP with respect to a higher integer dilation Λ. In this case we take $\tau_c = \tau_r =: \tau$, with

$$\tau(\omega) := \left(\frac{\sin(\frac{\Lambda\omega}{2})}{\Lambda \sin(\frac{\omega}{2})} \right)^2.$$

Let h be the *Fejér kernel*, viz., the filter corresponding to τ. The detail coefficient at $k \in \Lambda\mathbb{Z}$ is defined by

$$d_j(k) = y_j(k) - y_{j-1}(\frac{k}{2}) = ((\delta - h) * y_j)(k),$$

which corresponds to the mask $1 - \tau$, with τ as above. The filter $\delta - h$ is a $(2\Lambda - 1)$-tap one, the CAMPlet is (with ϕ the centered hat function)

$$\psi_0^M := \phi_{1,0} - \Lambda^{-1/2}\phi.$$

This CAMPlet is supported in $[-1,1]$. The detail coefficients at other locations are more local. For $k \in m + \Lambda\mathbb{Z}$, $m \in \{1, \ldots, \Lambda-1\}$,

$$\widetilde{d}_j(k) = y_j(k) - \Lambda(h * (y_{j\downarrow\uparrow}))(k) = y_j(k) - \frac{(\Lambda-m)y_j(k-m) + m\,y_j(k-m+\Lambda)}{\Lambda}.$$

This corresponds to $(\Lambda-1)$ filters each of which being a 3-tap one:

$$h_m(k) := \Lambda^{-1} \begin{cases} -m, & k = m-\Lambda, \\ \Lambda, & k = 0, \\ m-\Lambda, & k = m, \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding CAMPlets are supported each in a domain of measure $\leq \frac{6}{\Lambda}$, and whose convex hull is of measure $1 + \frac{2}{\Lambda}$. For example, consider the case $\Lambda = 3$. The CAMPlets $(\psi_0^M, \psi_1^M, \psi_2^M)$ are depicted in Figure 3. They are supported on $[-1,1]$, $[-4/3, 1/3]$ and $[-4/3, 1/3]$, respectively. $\square$

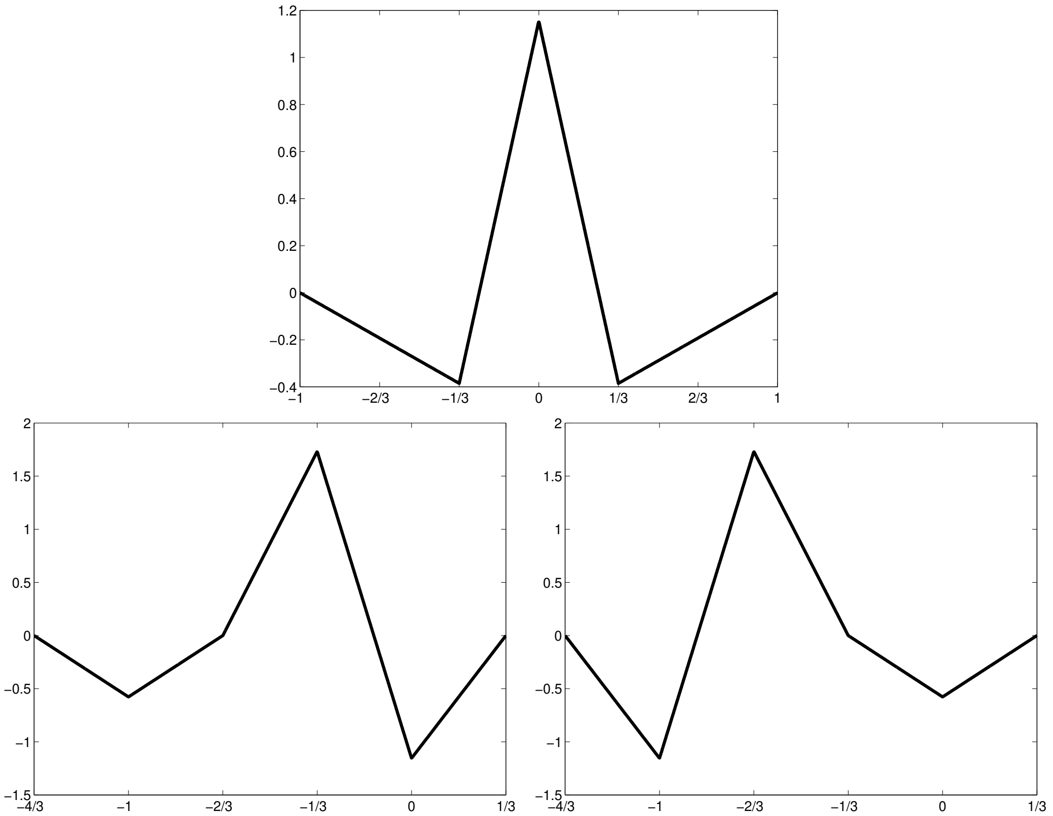

Figure 3. ψ_0^M (top) and ψ_1^M, ψ_2^M (bottom) of Example 5.5 with $\Lambda = 3$.

5.1.2 CAP and CAMP for $s_a = s_b = 2$

Example 5.6: Linear spline for compression, cubic spline for prediction. For this choice $\tau_c(\omega) = \cos^2(\frac{\omega}{2})$, while $\tau_r(\omega) = \cos^4(\frac{\omega}{2})$. Note that $s_{sm} = 3$, which is even more than what we need here. This CAP system is still of order 2, since $1 - \cos^6(\frac{\omega}{2})$ has a double zero at the origin. The CAPlet masks are

$$\sqrt{2}t_0(\omega) = \sin^2(\frac{\omega}{2})(1 + 2\cos^4(\frac{\omega}{2})),$$
$$\sqrt{2}e^{-i\omega}t_1(\omega) = \sin^2(\frac{\omega}{2})(2 + \cos\omega).$$

The CAP filters are 7- and 5-tap respectively, and the CAPlets (ψ_0, ψ_1) are supported on $[-2, 2]$ and $[-2, 1]$, respectively. ψ_0 is depicted in Figure 4 and $\psi_1 = E^{-1/2}\psi^{**}$. □

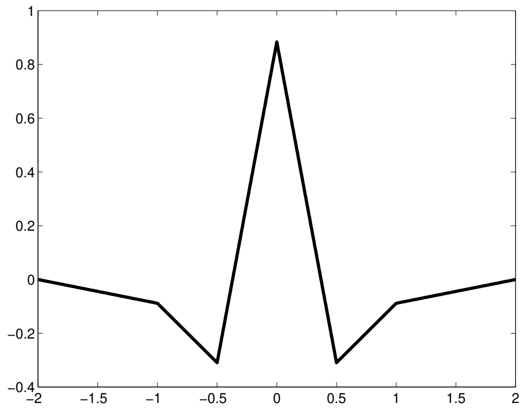

Figure 4. The CAPlet ψ_0 in Example 5.6.

Example 5.7: a CAMP representation. Since the cubic B-spline is not interpolatory, we cannot convert the previous example to CAMP. Instead, while retaining our choice of ϕ_c as the (centered) hat function, we replace ϕ_r by the Deslauriers-Dubuc cardinal interpolant of order 4, i.e.,

$$\tau_r(\omega) := \cos^4(\frac{\omega}{2})(2 - \cos\omega) = \frac{8 + 9\cos(\omega) - \cos(3\omega)}{16}.$$

We do have $s_b = 2$ here, since it is easy to check that this CAP system is of order 2 and it is well-known that $s_{sm} = 2$. Here, h_r is interpolatory, and the CAMP construction yields the following two filters:

$$\sqrt{2}t_0^M(\omega) = 1 - \tau_c(\omega) = \sin^2(\frac{\omega}{2}),$$

$$\sqrt{2}e^{-i\omega}t_1^M(\omega) = 1 - (\overline{\tau}_r(\omega) - \overline{\tau}_r(\omega + \pi)) = 2\,\tau_r(\omega + \pi) = \frac{8 - 9\cos(\omega) + \cos(3\omega)}{8}.$$

The filters of (τ_c, t_0^M, t_1^M) are $(3, 3, 5)$-tap, respectively. The CAMPlets (ψ_0^M, ψ_1^M) are supported in $[-1, 1]$ and $[-2.5, 1.5]$, respectively. Here, $\psi_0^M = \psi^*$ and ψ_1^M is depicted in Figure 5. □

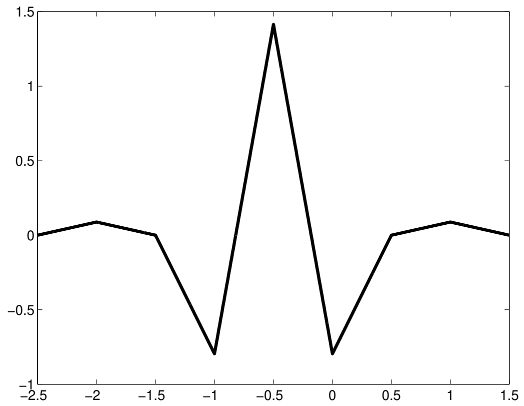

Figure 5. The CAMPlet ψ_1^M in Example 5.7.

Example 5.8: Enhancing Example 5.3. CAMP enhancement might be the method of choice in case a given CAMP construction has an undesired gap between its s_a performance and it s_b performance. Thus, in lieu of upgrading the prediction hat function to either of the cubic spline or the Deslauriers-Dubuc refinable function, we could enhance the CAMP construction of §5.1.1. The centered hat function is our choice for enhancement. The enhancement in this case does not change the order 2 of the underlying CAP. Altogether, we have here

$$\tau_{\mathrm{e}}(\omega) = \tau_c(\omega) = \tau_{\mathrm{in}}(\omega) = \cos^2(\omega/2).$$

Then $\tau_r(\omega) = \cos^2(\omega)\cos^2(\omega/2)$. Thus $s_{sm} = s_b = 2$. The enhancement does not alter t_1^M (see (5.4)), but extends $\sqrt{2}t_0^M$. Instead of the simple $\sin^2(\omega/2)$, we have now

$$\sqrt{2}t_0^M(\omega) = 1 - \cos^2(\omega)\cos^2(\omega/2).$$

This modified ψ_0^M CAMPlet is 7-tap, and is supported in $[-2, 2]$, as one can see from Figure 6. □

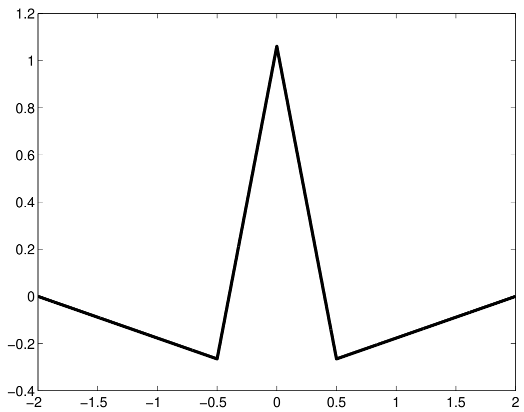

Figure 6. The CAMPlet ψ_0^M in Example 5.8.

5.1.3 CAMP for $s_a = 4$ or $s_b = 4$

The construction in Example 5.7 can be improved to yield better performance. For $s_a = 4$, we will need to improve the CAP order from 2 to 4. For $s_b = 4$ we will also need to enhance the smoothness of the prediction refinable function.

One way for lifting the CAP order is to use alignment. The drawback in using alignment in the context of CAMP is that it extends both CAMP filters. Changing the compression filter, or inserting an enhancement filter change, in contrast, only the filter t_0^M. That said, alignment is not as bad as it may first look: changing τ_c extends the support of the refinable function (which so far was always chosen to be the centered hat function), while inserting enhancement should be carefully tested.* The Deslauriers-Dubuc refinable function of Example 5.7 satisfies the SF conditions of order 4, which fits our needs.

Example 5.9: Let τ_c and τ_r be the same as in Example 5.7, but let

$$\Theta(\omega) := \frac{7 - \cos(\omega)}{6}.$$

Then the order of the CAP system is 4 since $1 - \tau_r(\omega) = O(|\omega|^4)$ and

$$\Theta(\omega) - \Theta(2\omega)\cos^2(\omega/2) = O(|\omega|^4).$$

The CAMPlet masks are

$$\begin{aligned}
\sqrt{2}t_0^M(\omega) &= \Theta(\omega) - \Theta(2\omega)\cos^2(\omega/2) = \frac{1}{24}(14 - 17\cos(\omega) + 2\cos(2\omega) + \cos(3\omega)),\\
\sqrt{2}e^{-i\omega}t_1^M(\omega) &= \Theta(\omega)(2\,\tau_r(\omega+\pi)) = \left(\frac{7-\cos(\omega)}{6}\right)\left(\frac{8 - 9\cos(\omega) + \cos(3\omega)}{8}\right)\\
&= \frac{1}{96}(121 - 142\cos(\omega) + 8\cos(2\omega) + 14\cos(3\omega) - \cos(4\omega)).
\end{aligned}$$

The filters of (τ_c, t_0^M, t_1^M) are (3, 7, 9)-tap, respectively. The CAMPlets (ψ_0^M, ψ_1^M) are supported in $[-2, 2]$ and $[-3, 2]$, respectively, as can be seen from Figure 7. Here $s_a = 4$ but $s_b = 2$. □

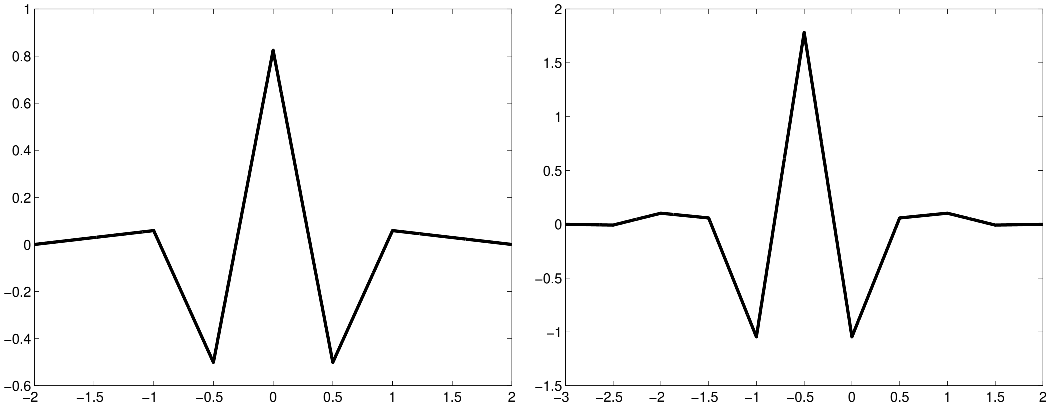

Figure 7. The CAPMlets ψ_0^M (left) and ψ_1^M (right) of Example 5.9.

---

* Whenever enhancement is selected, the shifts of ϕ_r cannot form a Riesz basis. This may appear to have practical implications.

Example 5.10: An alternative to the above is to use the Deslauriers-Dubuc filter for compression, too. That is, we choose $\tau_c = \tau_r =: \tau$ to be the Deslauriers-Dubuc filter and forgo using alignment. Then the order of the CAP system is 4, and again $s_a = 4$ but $s_b = 2$. The CAMPlet masks are

$$\begin{aligned}\sqrt{2}t_0^M(\omega) &= 1 - \tau(\omega) = \tau(\omega + \pi) = \frac{8 - 9\cos(\omega) + \cos(3\omega)}{16},\\ t_1^M(\omega) &= 2e^{i\omega}t_0^M(\omega).\end{aligned}$$

Each of the filters of (τ, t_0^M, t_1^M) is 5-tap. The CAMPlets (ψ_0^M, ψ_1^M) are supported in $[-3, 3]$ and $[-3.5, 2.5]$, respectively. The graph of ψ_0^M is depicted in Figure 8. Note that the two CAMPlets are the same up to normalization and shifting, more precisely, $\psi_1^M = 2(E^{-1/2}\psi_0^M)$. This is always the case in 1D CAMP construction provided that we choose $\tau_c = \tau_r = \tau_{\text{in}}$ and forgo alignment. $\square$

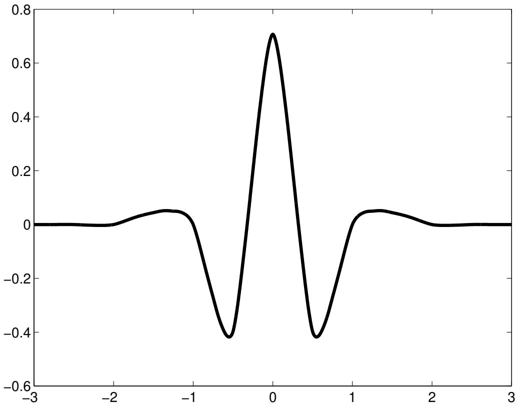

Figure 8. The CAMPlet ψ_0^M in Example 5.10.

Example 5.11: In order to get $s_b = 4$, we use an enhancement filter and then an alignment filter. That is, we choose τ_{in} to be the Deslauriers-Dubuc filter and let

$$\tau_{\text{e}}(\omega) = \tau_c(\omega) = \cos^2(\omega/2), \quad \Theta(\omega) := \frac{11 - 5\cos(\omega)}{6}.$$

Then $\tau_r(\omega) = \tau_{\text{e}}(2\omega)\tau_{\text{in}}(\omega)$. Here $s_a = 4$ as before, since the order of the CAP system is 4 from the fact that and $1 - \tau_{\text{in}}(\omega) = O(|\omega|^4)$ and that

$$\Theta(\omega) - \Theta(2\omega)\cos^2(\omega)\cos^2(\omega/2) = O(|\omega|^4).$$

Since $s_{sm} = 4$, we also have $s_b = 4$. The CAMPlet masks are

$$\begin{aligned}\sqrt{2}t_0^M(\omega) &= \Theta(\omega) - \Theta(2\omega)\cos^2(\omega)\cos^2(\omega/2)\\ &= \frac{1}{96}(142 - 126\cos(\omega) - 24\cos(2\omega) - 7\cos(3\omega) + 10\cos(4\omega) + 5\cos(5\omega)),\\ \sqrt{2}e^{-i\omega}t_1^M(\omega) &= \Theta(\omega)(2\,\tau_{\text{in}}(\omega + \pi)) = \left(\frac{11 - 5\cos(\omega)}{6}\right)\left(\frac{8 - 9\cos(\omega) + \cos(3\omega)}{8}\right)\\ &= \frac{1}{96}(221 - 278\cos(\omega) + 40\cos(2\omega) + 22\cos(3\omega) - 5\cos(4\omega)).\end{aligned}$$

The filters of (τ_c, t_0^M, t_1^M) are $(3, 11, 9)$-tap, respectively. The CAMPlets (ψ_0^M, ψ_1^M) are supported in $[-3, 3]$ and $[-3, 2]$, respectively. Their graphs are found in Figure 9. $\square$

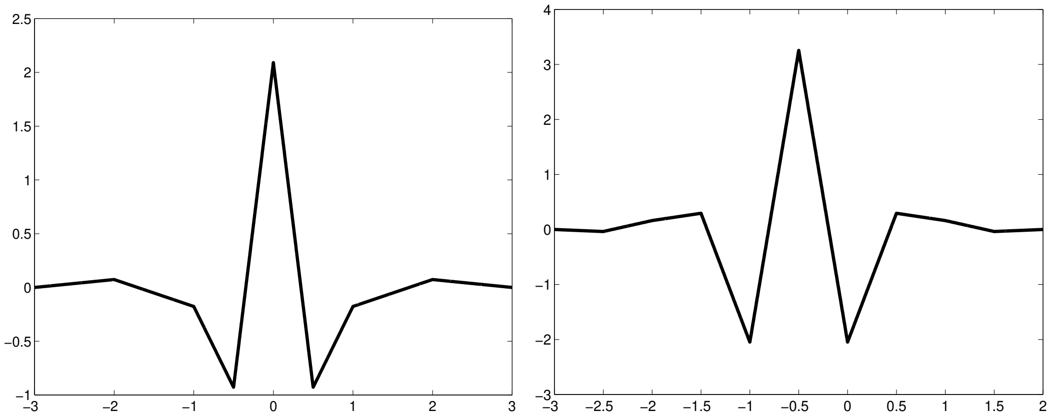

Figure 9. The CAPMlets ψ_0^M (left) and ψ_1^M (right) of Example 5.11.

5.1.4 CAP for $s_a = 4$

We start with the case $s_a = s_b + 1 = 4$.

Example 5.12: Linear spline for compression, cubic spline for prediction, and use alignment. We let $\tau_c(\omega) = \cos^2(\frac{\omega}{2})$, $\tau_r(\omega) = \cos^4(\frac{\omega}{2})$ and let

$$\Theta(\omega) := 1 + \sin^2(\frac{\omega}{2}).$$

Then we have $s_a = s_b + 1 = 4$ since $s_{sm} = 3$ and

$$\Theta(\omega) - \Theta(2\omega)\tau_r(\omega)\overline{\tau}_c(\omega) = O(|\omega|^4).$$

The CAPlet masks are

$$\begin{aligned}\sqrt{2}t_0(\omega) &= \frac{1}{64}(62 - 66\cos(\omega) + \cos(3\omega) + 2\cos(4\omega) + \cos(5\omega)),\\ \sqrt{2}e^{-i\omega}t_1(\omega) &= \frac{1}{16}(19 - 18\cos(\omega) - 4\cos(2\omega) + 2\cos(3\omega) + \cos(4\omega)).\end{aligned}$$

The filters of (τ_c, t_0, t_1) are $(3, 9, 9)$-tap, respectively. The graphs of ϕ_c, ψ_0, ψ_1 are depicted in Figure 10. The CAPlets (ψ_0, ψ_1) are supported in $[-3, 3]$ and $[-3, 2]$, respectively. Note that, in comparison with Example 5.6, we have improved s_a and s_b but the CAPlets are now supported in larger intervals. □

Next, we present two examples that provide $s_a = s_b = 4$. As one might have guessed, none of the examples here beats the support size of the CAMPlets in Example 5.11.

Example 5.13. We choose

$$\tau_c(\omega) := \frac{1}{8}(1 + e^{-i\omega})^3(1 + 4\sin^2(\frac{\omega}{2})), \quad \tau_r(\omega) := \frac{1}{32}e^{i\omega}(1 + e^{-i\omega})^5,$$

and forgo alignment. Standard smoothness analysis of refinable functions (cf. [D]) yields that $\phi_c \in \mathcal{D}^0(\mathbb{R})$. We have $s_a = s_b = 4$ since $s_{sm} = 4$ and

$$1 - \tau_r(\omega)\overline{\tau}_c(\omega) = 1 - \cos^8(\frac{\omega}{2})(1 + 4\sin^2(\frac{\omega}{2})) = O(|\omega|^4).$$

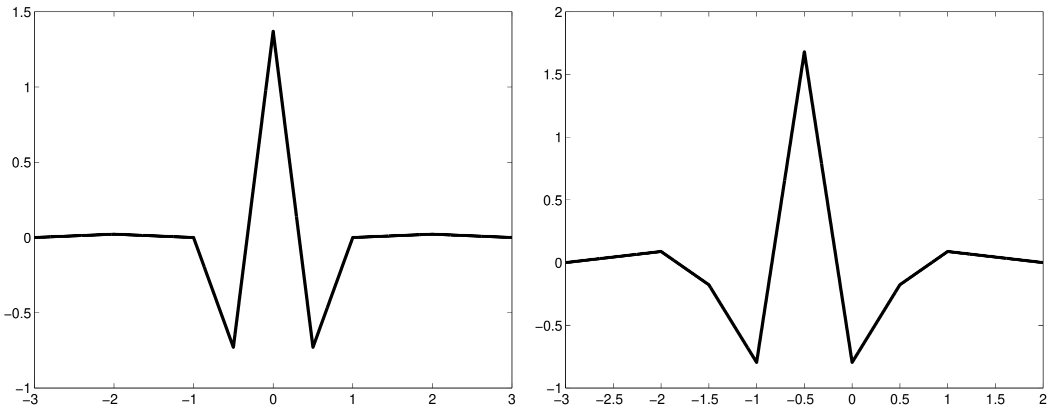

Figure 10. The CAPlets ψ_0 (left) and ψ_1 (right) of Example 5.12.

The CAPlet filters are (as the coefficients of $e^{5i\omega}, e^{4i\omega}, \dots, e^{-4i\omega}$)

$$
\begin{array}{rcl}
t_0 = & \dfrac{\sqrt{2}}{256} & * \; [\; 1 \quad 0 \quad 5 \quad -5 \quad -45 \quad 79 \quad -25 \quad -15 \quad 0 \quad 5 \;], \\
t_1 = & \dfrac{\sqrt{2}}{256} & * \; [\; 5 \quad 0 \quad -15 \quad -25 \quad 79 \quad -45 \quad -5 \quad 5 \quad 0 \quad 1 \;].
\end{array}
$$

The filters of (τ_c, t_0, t_1) are $(4, 8, 8)$-tap, respectively. The graphs of ϕ_c, ψ_0, ψ_1 are depicted in Figure 11. The CAPlets (ψ_0, ψ_1) are supported in intervals of size $(7, 7)$, respectively. The obvious lack of smoothness in the graphs should not mislead here: this is a high-performance system! $\square$

Example 5.14. We choose

$$\tau_c(\omega) := \frac{1}{4}(1 + e^{-i\omega})^2 \rho(\omega), \quad \tau_r(\omega) := \frac{1}{64} e^{2i\omega}(1 + e^{-i\omega})^6 \rho(\omega), \quad |\rho(\omega)|^2 = 1 + 4\sin^2(\frac{\omega}{2}),$$

and forgo alignment. We have $s_a = s_b = 4$ since the standard smoothness analysis of refinable functions (cf. [D]) yields that $s_{sm} = 4$, and

$$1 - \tau_r(\omega)\overline{\tau}_c(\omega) = 1 - \cos^8(\frac{\omega}{2})(1 + 4\sin^2(\frac{\omega}{2})) = O(|\omega|^4).$$

The CAPlet masks are

$$
\begin{aligned}
\sqrt{2} t_0(\omega) &= 1 - \Big(1 + 4\sin^2(\frac{\omega}{2})\Big)\cos^8(\frac{\omega}{2}) + \rho(\omega)\overline{\rho(\omega+\pi)}\cos^2(\frac{\omega}{2})\sin^6(\frac{\omega}{2}), \\
\sqrt{2} e^{-i\omega} t_1(\omega) &= 1 - \Big(1 + 4\sin^2(\frac{\omega}{2})\Big)\cos^8(\frac{\omega}{2}) - \rho(\omega)\overline{\rho(\omega+\pi)}\cos^2(\frac{\omega}{2})\sin^6(\frac{\omega}{2}).
\end{aligned}
$$

For the choice of $\rho(\omega) := \frac{(1+\sqrt{5})}{2} + \frac{(1-\sqrt{5})}{2} e^{-i\omega}$, the filters (τ_c, t_0, t_1) are (as the coefficients of $e^{6i\omega}, e^{5i\omega}, \dots, e^{-5i\omega}$), with $b = \sqrt{5}$,

$$
\begin{array}{rclcccccccccccc}
\tau_c = & \frac{1}{8} & *[0 & 0 & 0 & 0 & 0 & 0 & 1+b & 3+b & 3-b & 1-b & 0 & 0], \\
t_0 = & \frac{\sqrt{2}}{512} & *[0 & 0 & 9-b & 2+4b & -28-4b & -56-4b & 158+10b & -84-4b & -12-4b & 8+4b & 1-b & 2], \\
t_1 = & \frac{\sqrt{2}}{512} & *[2 & 1+b & 8-4b & -12+4b & -84+4b & 158-10b & -56+4b & -28+4b & 2-4b & 9+b & 0 & 0],
\end{array}
$$

and $(4, 10, 10)$-tap, respectively. The graphs of ϕ_c, ψ_0, ψ_1 are depicted in Figure 12. The CAPlets (ψ_0, ψ_1) are supported in intervals of size $(6, 6)$, respectively. $\square$

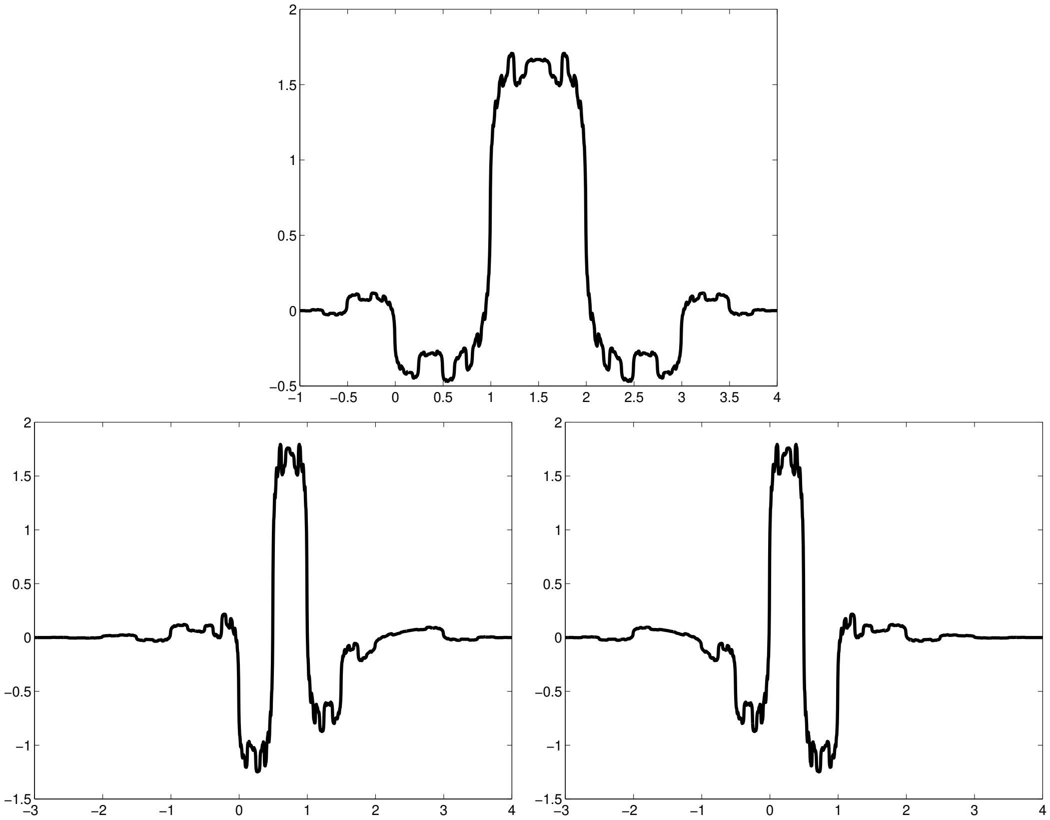

Figure 11. ϕ_c (top) and ψ_0, ψ_1 (bottom) of Example 5.13.

5.2. Multivariate piecewise-linear CAMP constructions: $s_a = s_b + 1 = 2$

We consider now in detail the example that was discussed in §1.2, but restrict our attention to dyadic dilation in 2 and 3 variables. In each of these constructions we have $s_a = 2$ and $s_b = 1$.

5.2.1 Piecewise-linear CAMP in 2 dimensions

We choose the two refinable functions ϕ_c ϕ_r to be the 3-directional box spline ϕ, [BHR], and forgo alignment (which will not help here in any event). Our box spline, whose direction set is

$$\Xi = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

is a continuous piecewise-linear function (hence $s_{sm} = 1$) which is supported in the hexagon

$$H := \Xi[-1/2, 1/2]^3,$$

i.e., in the image of the cube under the above Ξ. ϕ is refinable with mask

$$\tau(\omega_1, \omega_2) = \cos(\frac{\omega_1}{2})\cos(\frac{\omega_2}{2})\cos(\frac{\omega_1 + \omega_2}{2}).$$

It is easy to see that the CAP system associated with the above τ is of order 2, i.e.,

$$1 - \tau^2 = O(|\cdot|^2).$$

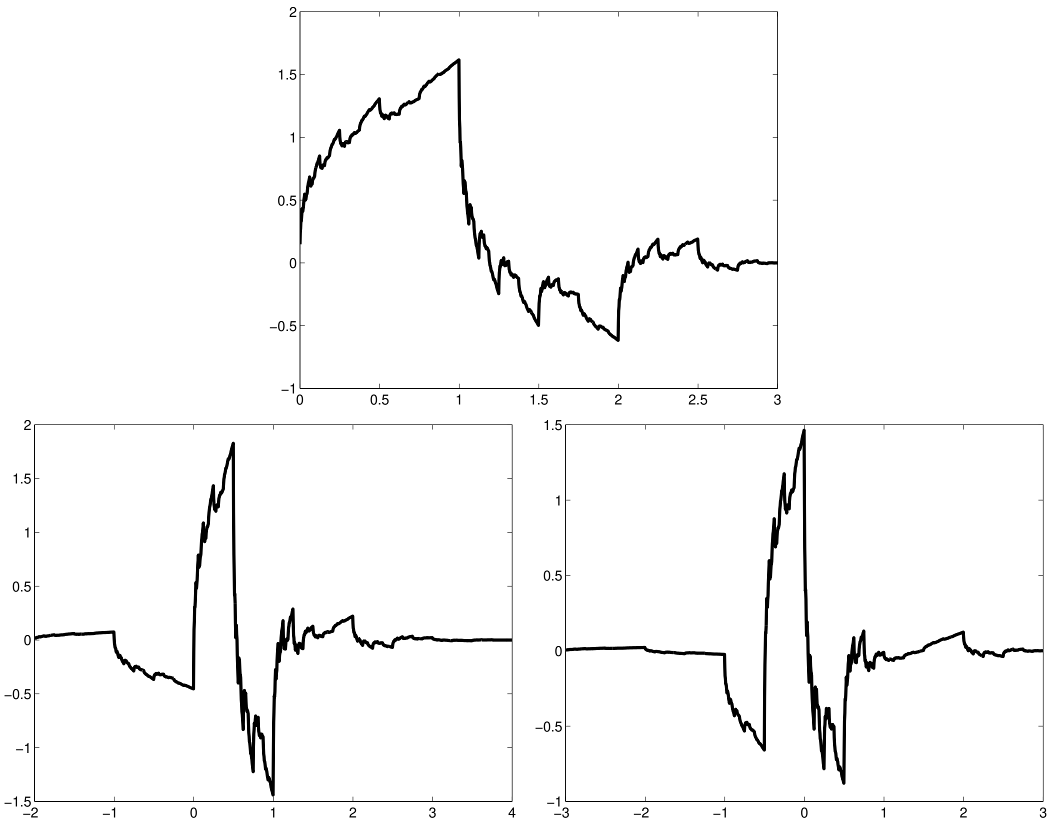

Figure 12. ϕ_c (top) and ψ_0, ψ_1 (bottom) of Example 5.14.

Thus, indeed, this CAP system fits $s_a = 2$ and $s_b = 1$.

The filter h associated with τ has the following non-zero values:

$$h(-1,-1) = h(-1,0) = h(0,-1) = h(0,1) = h(1,0) = h(1,1) = \frac{1}{8} \text{ and } h(0,0) = \frac{1}{4}.$$

Therefore, h is interpolatory, hence an associated CAMP representation is available. Given the dataset y_j, the CAMP detail at an even location $k \in 2\mathbb{Z}^2$ is defined by

$$\widetilde{d}_j(k) = ((\delta - h) * y_j)(k),$$

which corresponds to the 7-tap filter $\delta - h$, and the corresponding CAMPlet

$$\phi_{1,0} - \frac{1}{2}\,\phi,$$

which is supported in the hexagon H, whose area is 3. For $k \in \mathbb{Z}^2 \backslash (2\mathbb{Z}^2)$, the CAMP coefficients are more local:

$$\widetilde{d}_j(k) = [y_j - 4\,h * (y_{j\downarrow\uparrow})](k).$$

Concretely, this means that, with ξ one of the columns in the direction matrix Ξ, and $k \in \xi + 2\mathbb{Z}^2$,

$$\widetilde{d}_j(k) = y_j(k) - \frac{y_j(k+\xi) + y_j(k-\xi)}{2},$$

	3/5	5/3	CAMP in §5.2.1	CAMP in §5.3
performance (s_a)	2.00	2.00	2.00	2.00
performance (s_b)	1.00	negligible	1.00	2.00
average size of high-pass filters	13.00	18.33	4.00	8.50
average area of wavelets' support	11.00	7.00	2.06	5.44

Table 1. Comparison of performance and space localization in 2D.

which corresponds to double-differencing in the ξ direction. Each of the three CAMPlets is supported in a hexagon of area 1.75.

On average, the filters used to compute the CAMP coefficients are ($\frac{3\times 3+7}{4}=$)4-tap, while the area of the CAMPlets' support is 2.06. We note that the associated CAP representation employs 4 filters of average size 11.5.

A mainstream wavelet system with identical performance $s_a = 2$ and $s_b = 1$ is the 3/5 one, with the centered bilinear hat function used for reconstruction. Its three wavelets are supported in rectangles of areas $9, 12, 12$ and the high-pass filters are $15, 15, 25$-tap. A slight reduction in these numbers can be obtained by using the bilinear spline for decomposition: the support areas will be then $9, 6, 6$ while the high-pass filters will be $15, 15, 9$-tap. However, this reverse system, while still having $s_a = 2$, has a very low s_b, somewhere around 0. The comparison is summarized in Table 1.

5.2.2 Piecewise-linear CAMP in 3 dimensions

Now, we take ϕ_c and ϕ_r to be the box spline ϕ with direction set

$$\Xi = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Once again, this box spline is continuous piecewise-linear and is supported in the parallelepiped $\Xi[-1/2, 1/2]^4$ (hence, again, $s_{sm} = 1$). The refinement mask is

$$\tau(\omega_1, \omega_2, \omega_3) = \cos(\frac{\omega_1}{2})\cos(\frac{\omega_2}{2})\cos(\frac{\omega_3}{2})\cos(\frac{\omega_1+\omega_2+\omega_3}{2}),$$

and again, the CAP representation is of order 2 here. As before, this CAP system delivers $s_a = 2$ and $s_b = 1$.

The filter h associated with τ has 15-point support:

$$h(\pm k) = \frac{1}{16},\ k \in \{0,1\}^3\backslash 0, \quad h(0) = \frac{1}{8}.$$

Since h is interpolatory, we consider a CAMP representation. Let y_j, $j \in \mathbb{Z}$, be our dataset. For $k \in 2\mathbb{Z}^3$, the CAMP coefficients are

$$\widetilde{d}_j(k) = ((\delta - h) * y_j)(k),$$

which gives rise to a 15-tap CAMP filter, which corresponds to the CAMPlet $\phi_{1,0} - 2^{-3/2}\,\phi$, with support volume 4. For $k \in \mathbb{Z}^3\backslash(2\mathbb{Z}^3)$, the CAMP coefficients are more local:

$$\widetilde{d}_j(k) = [y_j - 8\,h * (y_{j\downarrow\uparrow})](k).$$

	3/5	5/3	CAMP in §5.2.2
performance (s_a)	2.00	2.00	2.00
performance (s_b)	1.00	negligible	1.00
average size of high-pass filters	55.29	69.29	4.50
average volume of wavelets' support	39.90	16.71	1.69

Table 2. Comparison of performance and space localization in 3D.

The above translates into 7 different rules, depending on the coset of $\mathbb{Z}^3/2\mathbb{Z}^3 = \{0,1\}^3 =: \Gamma$ that k belongs to. For $\eta \in \Gamma\backslash 0$, and $k \in \eta + 2\mathbb{Z}^3$,

$$\widetilde{d}_j(k) = y_j(k) - \frac{y_j(k+\eta) + y_j(k-\eta)}{2}.$$

This shows that each of the other 7 CAMP filters is 3-tap. The volume of its support is ≤ 1.5. In fact, if η is one of the four columns of Ξ, this volume is exactly 1.25.

In summary, the CAMP filters are $\frac{3\times 7+15}{8} = 4.5$-tap on average, and the average volume of their support is $\leq \frac{4\times 1.25+3\times 1.5+4}{8} \leq 1.7$. Table 2 compares this CAMP to the 3D tensor product 3/5, as well as the 3D tensor product 5/3.

5.3. Higher-performance CAMP in 2 dimensions: $s_a = s_b = 2$.

In order to reach $s_b = 2$, we choose ϕ_c as the three-directional box spline from §5.2.1, but need a smoother reconstruction function ϕ_r.

One option is to choose ϕ_r as the tensor product of the Deslauriers-Dubuc cardinal interpolant from §5.1.2. For this case, h_r has 25 non-zero coefficients, and ϕ_r is supported on $[-3,3]^2$, and $s_{sm} = 2$. Since h_r is interpolatory we can resort here to a CAMP construction. We note that $1-\tau_r$ has a 4th order zero at the origin, and since $1-\tau_c$ has a double zero at the origin, the CAP system is of order 2, which implies that the CAMP representation delivers $s_a = s_b = 2$.

There are four CAMPlets, hence four CAMPlet masks, here. The first one corresponds to the filter $\delta - h_c$, hence is a 7-tap filter and results in a CAMPlet ψ_0 supported in the same hexagon as ϕ_c. Two other CAMPlets are supported in the strips $[-0.5, 0.5]\times[-2.5, 1.5]$ and $[-2.5, 1.5]\times[-0.5, 0.5]$, respectively. The exact area of each support is 3.25 and the area of the convex hull is 3.75. Each of the filters is 5-tap, and is supported on one of the coordinate axes.

The 4th CAMPlet is larger, and this is due to the fact that one of the four cosets of h_r has 16 non-zero coefficients. This CAMPlet corresponds to a 17-tap filter, and is supported in the square $[-2.5, 1.5]^2$. The exact area of its support is 12.25, and the area of the convex hull is 15.75.

Note that the sum of the above support areas is 21.75, and the sum of their convex hulls is 26.25. See Table 1.

6. Appendix: Proofs of Theorem 3.5 and 3.7

Throughout this section, we denote

$$2^a_+ := \max\{2^a, 1\}.$$

We first state the following simple fact:

Lemma 6.1. *Let $l \in \mathbb{Z}$. If $k \in \mathbb{Z}^n$ and $\gamma > n$ then*

$$\sum_{m \in \mathbb{Z}^n} \left(1 + \frac{|2^l k - m|}{2^l_+}\right)^{-\gamma} \lesssim 2^{ln}_+.$$

More substantially, we will need the following lemma that is proved in [FJ2]. In the lemma and elsewhere in this section, we use the notation $f \in \mathcal{R}^0_\gamma(\mathbb{R}^n)$ to mean $f \in \mathcal{R}^{-\varepsilon}_\gamma(\mathbb{R}^n)$ for some $0 < \varepsilon < 1$.

Lemma 6.2. *Let η be either 0 or a positive non-integer number. Let $\gamma > n + \eta$. Then, for $j \leq l$,*

$$|\langle \theta_{j,k}, \zeta_{l,m} \rangle| \leq c\, 2^{-(l-j)(\eta + n/2)} \left(1 + \frac{|2^{l-j} k - m|}{2^{l-j}}\right)^{-\gamma}$$

provided that $\theta \in \mathcal{R}^\eta_\gamma$ and $\zeta \in \mathcal{R}^0_\gamma \cap \mathcal{M}^\lambda$, with λ satisfying $\lfloor \lambda \rfloor + 1 > \eta$ if $\eta > 0$, and with any $\lambda \in \mathbb{R}$ if $\eta = 0$.

We begin with the proof of the Besov case, since this proof is more self-contained.

6.1. Proof of Theorem 3.7

For $t \in \mathbb{R}$ and $0 < p, q \leq \infty$, the space $\dot{b}_{pq}(t)$ is defined to be the space of all sequences $h := (h(j,k) : j \in \mathbb{Z}, k \in \mathbb{Z}^n)$ satisfying

$$\|h\|_{\dot{b}_{pq}(t)} := \left(\sum_{j \in \mathbb{Z}} \left(\sum_{k \in \mathbb{Z}^n} (2^{jt} |h(j,k)|)^p\right)^{q/p}\right)^{1/q} < \infty,$$

(with the usual modification for $p = \infty$ or $q = \infty$).

Proposition 6.3. *Let $t \in \mathbb{R}$, $0 < p, q \leq \infty$ and*

$$\mu := \frac{n}{\min\{1, p\}}.$$

Let $\mathbf{A}$ be a complex-valued matrix whose rows and columns are indexed by $\mathbb{Z} \times \mathbb{Z}^n$:

$$\mathbf{A} := (\mathbf{A}(j,k;l,m) : j, l \in \mathbb{Z}, k, m \in \mathbb{Z}^n).$$

Suppose that there exist constants $\gamma > \mu$ and $\varepsilon > 0$ such that, for all j, k, l, m,

$$|\mathbf{A}(j,k;l,m)| \lesssim \frac{2^{-|l-j|\varepsilon}\, 2^{(l-j)(t + \frac{n}{p})}}{2^{(l-j)\mu}_+} \left(1 + \frac{|2^{l-j} k - m|}{2^{l-j}_+}\right)^{-\gamma}. \tag{6.4}$$

Then $\mathbf{A}$ is a bounded endomorphism of $\dot{b}_{pq}(t)$.

Proof of Proposition 6.3: Let $\dot{b}$ be the space of all sequences $h := (h_j \in \mathbb{C}^{\mathbb{Z}^n} : j \in \mathbb{Z})$ such that $\|h\|_{\dot{b}} := \|(\|h_j\|_p : j \in \mathbb{Z})\|_q < \infty$ where $\|\cdot\|_p := \|\cdot\|_{\ell_p(\mathbb{Z}^n)}$ and $\|\cdot\|_q := \|\cdot\|_{\ell_q(\mathbb{Z})}$. Then it suffices to prove that

$$\widetilde{\mathbf{A}} := (\widetilde{\mathbf{A}}_{j,l}(k,m) := \widetilde{\mathbf{A}}(j,k;l,m) := 2^{(j-l)t} \mathbf{A}(j,k;l,m) : j, l \in \mathbb{Z}, k, m \in \mathbb{Z}^n)$$

is bounded on $\dot{b}$.

For every $h \in \dot{b}$ and $j \in \mathbb{Z}$,

$$\|(\widetilde{\mathbf{A}}h)_j\|_p = \|\sum_{l\in\mathbb{Z}} \widetilde{\mathbf{A}}_{j,l}h_l\|_p \lesssim \sum_{l\in\mathbb{Z}} \|\widetilde{\mathbf{A}}_{j,l}\|_p\|h_l\|_p,$$

where $\|\widetilde{\mathbf{A}}_{j,l}\|_p$ is the norm of $\widetilde{\mathbf{A}}_{j,l}$ as an endomorphism of $\ell_p(\mathbb{Z}^n)$. We prove the result in two steps. We first argue that if, for any $0 < p \le \infty$, there exists $\varepsilon := \varepsilon(p)$ s.t. $\|\widetilde{\mathbf{A}}_{j,l}\|_p \lesssim 2^{-|j-l|\varepsilon}$, then $\widetilde{\mathbf{A}}$ is bounded on $\dot{b}$. Then we show that the given condition (6.4) implies $\|\widetilde{\mathbf{A}}_{j,l}\|_p \lesssim 2^{-|j-l|\varepsilon}$.

For $0 < q \le 1$, we have $\|\widetilde{\mathbf{A}}\|_{\dot{b}} \lesssim \left(\sup_{l\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} \|\widetilde{\mathbf{A}}_{j,l}\|_p^q\right)^{1/q}$ since

$$\|\widetilde{\mathbf{A}}h\|_{\dot{b}}^q \lesssim \sum_{j\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \|\widetilde{\mathbf{A}}_{j,l}\|_p^q\|h_l\|_p^q = \sum_{l\in\mathbb{Z}} \|h_l\|_p^q \sum_{j\in\mathbb{Z}} \|\widetilde{\mathbf{A}}_{j,l}\|_p^q \le \left(\sup_{l\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} \|\widetilde{\mathbf{A}}_{j,l}\|_p^q\right) \|h\|_{\dot{b}}^q.$$

Similarly for $q = \infty$, we get $\|\widetilde{\mathbf{A}}\|_{\dot{b}} \lesssim \sup_{j\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \|\widetilde{\mathbf{A}}_{j,l}\|_p$. Thus, it is clear that the condition $\|\widetilde{\mathbf{A}}_{j,l}\|_p \lesssim 2^{-|j-l|\varepsilon}$ implies the boundedness of $\widetilde{\mathbf{A}}$ on $\dot{b}$ when $0 < q \le 1$ or $q = \infty$. The case $1 < q < \infty$ follows then by a standard interpolation argument.

It remains to prove that for any $0 < p \le \infty$, there exists $\varepsilon := \varepsilon(p)$ such that $\|\widetilde{\mathbf{A}}_{j,l}\|_p \lesssim 2^{-|j-l|\varepsilon}$. For $0 < p \le 1$, by the argument from the first part of the proof,

$$\|\widetilde{\mathbf{A}}_{j,l}\|_p^p \lesssim \sup_{m\in\mathbb{Z}^n} \sum_{k\in\mathbb{Z}^n} |\widetilde{\mathbf{A}}_{j,l}(k,m)|^p.$$

We note that $\gamma p > n$ in this case. Invoking condition (6.4) and Lemma 6.1, we see that

$$\begin{aligned}
\|\widetilde{\mathbf{A}}_{j,l}\|_p^p &\lesssim \left(\frac{2^{-(l-j)t}2^{-|l-j|\varepsilon}2^{(l-j)(t+\frac{n}{p})}}{2_+^{(l-j)\mu}}\right)^p \sup_{m\in\mathbb{Z}^n} \sum_{k\in\mathbb{Z}^n} \left(1 + \frac{|2^{j-l}m-k|}{2_+^{j-l}}\right)^{-\gamma p} \\
&\lesssim \left(\frac{2^{(l-j)n/p}\,2^{-|l-j|\varepsilon}}{2_+^{(l-j)n/p}}\right)^p 2_+^{(j-l)n} = 2^{-|j-l|\varepsilon p}.
\end{aligned}$$

For $p = \infty$, we have $\mu = n$. Then, since $\|\widetilde{\mathbf{A}}_{j,l}\|_\infty \lesssim \sup_{k\in\mathbb{Z}^n} \sum_{m\in\mathbb{Z}^n} |\widetilde{\mathbf{A}}_{j,l}(k,m)|$, (6.4) and Lemma 6.1 imply that

$$\begin{aligned}
\|\widetilde{\mathbf{A}}_{j,l}\|_\infty &\lesssim \frac{2^{-|j-l|\varepsilon}}{2_+^{(l-j)n}} \sup_{k\in\mathbb{Z}^n} \sum_{m\in\mathbb{Z}^n} \left(1 + \frac{|2^{l-j}k-m|}{2_+^{l-j}}\right)^{-\gamma} \\
&\lesssim \frac{2^{-|j-l|\varepsilon}}{2_+^{(l-j)n}} 2_+^{(l-j)n} = 2^{-|j-l|\varepsilon}.
\end{aligned}$$

Again, the case $1 < p < \infty$ follows by interpolation, and this finishes the proof. □

Corollary 6.5. *Let $s \in \mathbb{R}$, $0 < p, q \le \infty$ and*

$$\lambda := n(\frac{1}{\min\{1,p\}} - 1) - s, \quad t := s + n(\frac{1}{2} - \frac{1}{p}).$$

Suppose that $\theta \in \mathcal{D}_\gamma^\lambda \cap \mathcal{M}^s$ and $\zeta \in \mathcal{D}_\gamma^s \cap \mathcal{M}^\lambda$ with $\gamma > n + \max\{s+\lambda, s, \lambda\}$. Then the matrix

$$\mathbf{A} := (\mathbf{A}(j,k;l,m) := \delta_{j,k;l,m} \langle \theta_{j,k}, \zeta_{l,m} \rangle : j,l \in \mathbb{Z}, k,m \in \mathbb{Z}^n, \delta_{j,k;l,m} \in \{\pm 1\})$$

defines a bounded operator on $\dot{b}_{pq}(t)$.

Proof of Corollary 6.5: If $\lambda < 0$, we let $\eta := 0$. If $\lambda \geq 0$, we can find a non-integer η such that $\lambda < \eta < \lfloor \lambda \rfloor + 1$, $n + \eta < \gamma$, and $\theta \in \mathcal{R}_\gamma^\eta$. If $s < 0$, we let $u := 0$. If $s \geq 0$, we can find a non-integer u such that $s < u < \lfloor s \rfloor + 1$, $n + u < \gamma$, and $\zeta \in \mathcal{R}_\gamma^u$. From Lemma 6.2, we see that $\mathbf{A}$ satisfies (6.4) with $\mu := \lambda + s + n$ and $\varepsilon := \varepsilon(j,l) := \begin{cases} \eta - \lambda, & l \geq j, \\ u - s, & l < j. \end{cases}$ Applying Proposition 6.3 finishes the proof. □

Corollary 6.6. *Let $s \in \mathbb{R}$, $0 < p, q \leq \infty$ and*

$$\lambda := n(\frac{1}{\min\{1,p\}} - 1) - s, \quad t := s + n(\frac{1}{2} - \frac{1}{p}).$$

Suppose that $\zeta \in \mathcal{D}_\gamma^s \cap \mathcal{M}^\lambda$ for some $\gamma > n + \max\{s+\lambda, s, \lambda\}$. Then for every $\theta \in \mathcal{D}_\gamma^\lambda \cap \mathcal{M}^s$ and $h := (h(l,m) : l \in \mathbb{Z}, m \in \mathbb{Z}^n) \in \dot{b}_{pq}(t)$,

$$\sum_{l,m} |h(l,m)| |\langle \theta, \zeta_{l,m} \rangle| < \infty. \tag{6.7}$$

In particular, the series $\sum_{l,m} h(l,m)\zeta_{l,m}$ converges in $\mathcal{S}'/\mathcal{P}$.

Proof: We first note that the sequence $(|\langle \theta, \zeta_{l,m} \rangle|)_{l,m}$ comprises the $(j = k = 0)$-row of the matrix $\mathbf{A}$ of Corollary 6.5. By that corollary, $\mathbf{A}$ is bounded on $\dot{b}_{pq}(t)$, hence $\mathbf{A}(h')(0,0)$ must be finite for every $h' \in \dot{b}_{pq}(t)$. However, for the choice $h' := |h|$, we have that $\mathbf{A}(h')(0,0) = \sum_{l,m} |h(l,m)| |\langle \theta, \zeta_{l,m} \rangle|$, hence (6.7) is true.

Since any test function of $\mathcal{S}'/\mathcal{P}$ is contained in $\mathcal{D}_\gamma^\lambda \cap \mathcal{M}^s$, the series $\sum_{l,m} h(l,m)\zeta_{l,m}$ converges in $\mathcal{S}'/\mathcal{P}$. □

Proof of Theorem 3.7: Let $t := s + n(1/2 - 1/p)$. Let $\varphi \in \mathcal{S}$ be a function satisfying the conditions in (3.1). We note, Result 3.2, that $T^*_{X(\varphi)}$ is a bijection from $\dot{B}^s_{pq}$ to $\dot{b}_{pq}(t)$ which is inverted by the bijection $T_{X(\varphi)}$. This yields that the sequence $T^*_{X(\psi)} f$, for $f \in \dot{B}^s_{pq}$ and $\psi \in \Psi \subset \mathcal{D}_\gamma^\lambda \cap \mathcal{M}^s$, is well defined (pointwise):

$$(T^*_{X(\psi)} f)(j,k) := \langle f, \psi_{j,k} \rangle = \sum_{l,m} \langle \varphi_{l,m}, \psi_{j,k} \rangle \langle f, \varphi_{l,m} \rangle,$$

since Corollary 6.6 (for $\theta := \psi$ and $\zeta := \varphi$) shows that the right-most sum converges absolutely. (For more detailed discussion about the expression $\langle f, \psi_{j,k} \rangle$, we refer to §6.3.)

Thus, we have proved the identity

$$T^*_{X(\psi)} f = (T^*_{X(\psi)} T_{X(\varphi)}) T^*_{X(\varphi)} f, \quad f \in \dot{B}^s_{pq}.$$

Since $T^*_{X(\psi)} T_{X(\varphi)}$ is bounded on $\dot{b}_{pq}(t)$ by Corollary 6.5 (where, as before, $\theta := \psi$ and $\zeta := \varphi$) we conclude that $T^*_{X(\psi)} : \dot{B}^s_{pq} \to \dot{b}_{pq}(t)$ is bounded, for each $\psi \in \Psi$:

$$\|T^*_{X(\psi)} f\|_{\dot{b}_{pq}(t)} \lesssim \|T^*_{X(\varphi)} f\|_{\dot{b}_{pq}(t)} \approx \|f\|_{\dot{B}^s_{pq}(\mathbb{R}^n)}, \quad f \in \dot{B}^s_{pq}.$$

We note that only the conditions on Ψ (but not on Ψ^{dual}) i.e., $\Psi \subset \mathcal{D}_\gamma^\lambda \cap \mathcal{M}^s$, are used to obtain the above Jackson-type inequality.

Now we prove that, for every $f \in \dot{B}^s_{pq}$,

$$\sum_{\psi\in\Psi}\sum_{l,m}\langle f,\psi_{l,m}\rangle\,\psi_{l,m}^{\mathrm{dual}} = f, \tag{6.8}$$

in the sense of $\mathcal{S}'/\mathcal{P}$. To this end, let $\mathcal{S}_\infty := \{\eta \in \mathcal{S} : \int_{\mathbb{R}^n} \eta(t)t^\alpha dt = 0, \forall \alpha \in \mathbb{N}_0^n\}$. We need to show

$$\langle \sum_{\psi\in\Psi}\sum_{l,m}\langle f,\psi_{l,m}\rangle\,\psi_{l,m}^{\mathrm{dual}},\eta\rangle = \langle f,\eta\rangle, \quad \forall \eta \in \mathcal{S}_\infty :$$

$$\begin{aligned}
\langle \sum_{\psi\in\Psi}\sum_{l,m}\langle f,\psi_{l,m}\rangle\,\psi_{l,m}^{\mathrm{dual}},\eta\rangle &= \sum_{\psi\in\Psi}\sum_{l,m}\langle f,\psi_{l,m}\rangle\langle\psi_{l,m}^{\mathrm{dual}},\eta\rangle \quad (\text{using Corollary 6.6 and } T^*_{X(\psi)}f \in \dot{b}_{pq}(t))\\
&= \sum_{\psi\in\Psi}\sum_{l,m}\sum_{j,k}\langle f,\varphi_{j,k}\rangle\langle\varphi_{j,k},\psi_{l,m}\rangle\langle\psi_{l,m}^{\mathrm{dual}},\eta\rangle \quad (\text{from the definition of } \langle f,\psi_{l,m}\rangle)\\
&= \sum_{j,k}\langle f,\varphi_{j,k}\rangle\sum_{\psi\in\Psi}\sum_{l,m}\langle\varphi_{j,k},\psi_{l,m}\rangle\langle\psi_{l,m}^{\mathrm{dual}},\eta\rangle\\
&= \sum_{j,k}\langle f,\varphi_{j,k}\rangle\langle\varphi_{j,k},\eta\rangle\\
&= \langle f,\eta\rangle.
\end{aligned}$$

For the third equality, we used the fact

$$\sum_{\psi\in\Psi}\sum_{l,m}\left(\sum_{j,k}|\langle f,\varphi_{j,k}\rangle||\langle\varphi_{j,k},\psi_{l,m}\rangle|\right)|\langle\psi_{l,m}^{\mathrm{dual}},\eta\rangle| < \infty,$$

which follows from the facts that $\mathbf{M} := (\mathbf{M}(j,k;l,m) := |\langle\psi_{j,k},\varphi_{l,m}\rangle| : j,l \in \mathbb{Z}, k,m \in \mathbb{Z}^n)$ is bounded on $\dot{b}_{pq}(t)$ by Corollary 6.5, and that $(|\langle f,\varphi_{j,k}\rangle|)_{j,k} \in \dot{b}_{pq}(t)$, and by Corollary 6.6 (for $\theta := \eta$ and $\zeta := \psi^{\mathrm{dual}}$). For the fourth equality, we used that $\sum_{\psi\in\Psi} T_{X(\psi^{\mathrm{dual}})}T^*_{X(\psi)}\varphi_{j,k} = \varphi_{j,k}$ in the sense of L_2, and thus in the sense of $\mathcal{S}'$. Finally, the last equality is due to the fact that $\sum_{j,k}\langle f,\varphi_{j,k}\rangle\varphi_{j,k} = f$ in the sense of $\mathcal{S}'/\mathcal{P}$ (cf. Result 3.2).

Finally, from (6.8), we have that for every $f \in \dot{B}^s_{pq}$,

$$\langle f,\varphi_{j,k}\rangle = \sum_{\psi\in\Psi}\sum_{l,m}\langle\psi_{l,m}^{\mathrm{dual}},\varphi_{j,k}\rangle\langle f,\psi_{l,m}\rangle, \quad \forall j,k.$$

That is,

$$T^*_{X(\varphi)}f = \sum_{\psi\in\Psi}(T^*_{X(\varphi)}T_{X(\psi^{\mathrm{dual}})})T^*_{X(\psi)}f.$$

Since, for each ψ^{dual}, $T^*_{X(\varphi)}T_{X(\psi^{\mathrm{dual}})}$ is bounded by Corollary 6.5, we obtain

$$\|T^*_{X(\varphi)}f\|_{\dot{b}_{pq}(t)} \lesssim \sum_{\psi\in\Psi}\|(T^*_{X(\varphi)}T_{X(\psi^{\mathrm{dual}})})T^*_{X(\psi)}f\|_{\dot{b}_{pq}(t)} \lesssim \sum_{\psi\in\Psi}\|T^*_{X(\psi)}f\|_{\dot{b}_{pq}(t)}, \quad f \in \dot{B}^s_{pq}.$$

Invoking Result 3.2 one more time, we obtain the stated result. $\square$

6.2. Proof of Theorem 3.5

For $\sigma \in (-\infty,\infty]\backslash 0$, $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^n$, define $\chi_{\sigma,j,k} := 2^{jn(1/\sigma-1/2)}\chi_{j,k}$. Note that $\|\chi_{\sigma,j,k}\|_{L_\sigma(\mathbb{R}^n)} = \|\chi\|_{L_\sigma(\mathbb{R}^n)}$ if $\sigma \in (0,\infty]$.

For $\sigma \in (-\infty,\infty]\backslash 0$, $0 < p < \infty$ and $0 < q \le \infty$, the space $\dot{f}_{pq}(\sigma)$ is defined to be the set of all sequences $h := (h(j,k) : j \in \mathbb{Z}, k \in \mathbb{Z}^n)$ satisfying

$$\|h\|_{\dot{f}_{pq}(\sigma)} := \left\| \left(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^n} \left(|h(j,k)|\chi_{\sigma,j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} < \infty,$$

(with the usual modification for $q = \infty$).

Proposition 6.9. *Let $\sigma \in (-\infty,\infty]\backslash 0$, $0 < p < \infty$, $0 < q \le \infty$. Let* **A** *be as in Proposition 6.3, and assume that (6.4) holds for some $\gamma > \mu$ and $\varepsilon > 0$, and with t and μ there replaced by*

$$t := n(\frac{1}{\sigma} - \frac{1}{p}), \quad \mu := \frac{n}{\min\{1,p,q\}}.$$

Then, **A** *is a bounded endomorphism of $\dot{f}_{pq}(\sigma)$.*

The proof of the proposition employs maximal functions. Recall that, for a locally integrable function f, the maximal operator $M_t, t > 0$, is defined by

$$(6.10) \qquad (M_t f)(y) := \left(\sup_{Q\ni y} |Q|^{-1} \int_Q |f(z)|^t \, dz \right)^{1/t},$$

where the supremum is taken over all cubes with sides parallel to the axes.

The famous Fefferman-Stein inequality, [FS], states that if $0 < p < \infty$, $0 < q \le \infty$, and $0 < t < \min\{p,q\}$, then for any sequence $(f_j)_{j\in\mathbb{Z}}$ of functions,

$$(6.11) \qquad \left\| \left(\sum_{j\in\mathbb{Z}} (M_t f_j)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \le c \left\| \left(\sum_{j\in\mathbb{Z}} |f_j|^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)},$$

(with the usual modification for $q = \infty$). We also need the following complementary (and quite simple) result:

Lemma 6.12. *Let $0 < t \le 1$ and $\gamma > \frac{n}{t}$. For any $l \in \mathbb{Z}$, any sequence $(h(l,m) : m \in \mathbb{Z}^n)$ of complex numbers, and $y \in 2^{-j}(k + [0,1]^n)$, we have, with $\chi_{\infty,j,k} := \chi(2^j \cdot -k)$,*

$$\sum_{m\in\mathbb{Z}^n} |h(l,m)| \left(1 + \frac{|2^{l-j}k - m|}{2_+^{l-j}} \right)^{-\gamma} \le c\, 2_+^{(l-j)\frac{n}{t}} \, M_t \left(\sum_{m\in\mathbb{Z}^n} |h(l,m)|\chi_{\infty,l,m} \right)(y).$$

Another lemma that we need is listed below. This particular lemma is a simple consequence of the fact that convolution with ℓ_1-sequences is bounded in ℓ_p for $p \ge 1$:

Lemma 6.13. *Let $\rho > 0$ and $0 < q \leq \infty$. If $a \in \ell_q(\mathbb{Z})$, and if the sequence $b \in \mathbb{C}^{\mathbb{Z}}$ satisfies*

$$|b_j| \leq \sum_{l \in \mathbb{Z}} 2^{-|j-l|\rho} a_l, \quad j \in \mathbb{Z},$$

then

$$\|b\|_{\ell_q} \leq c \, \|a\|_{\ell_q}.$$

Proof of Proposition 6.9: The proof is presented for the case $q < \infty$. A similar argument establishes the case $q = \infty$.

For every $h \in \dot{f}_{pq}(\sigma)$,

$$(\mathbf{A}h)(j,k) = \sum_{l \in \mathbb{Z}, m \in \mathbb{Z}^n} \mathbf{A}(j,k;l,m) \, h(l,m)$$

and

$$\begin{aligned}
\|\mathbf{A}h\|_{\dot{f}_{pq}(\sigma)} &= \left\| \left(\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \left(|(\mathbf{A}h)(j,k)| \chi_{\sigma,j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} \\
&\leq \left\| \left(\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} \left(\sum_{l \in \mathbb{Z}, m \in \mathbb{Z}^n} |\mathbf{A}(j,k;l,m)| \, |h(l,m)| \chi_{\sigma,j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)}.
\end{aligned}$$

Since $\gamma > \mu$, we can choose t so that

$$\mu < \frac{n}{t} < \min\{\mu + \varepsilon, \gamma\}.$$

From (6.4), Lemma 6.12, Lemma 6.13 and (6.11), we see that $\|\mathbf{A}h\|_{\dot{f}_{pq}(\sigma)}$ is bounded by

$$\begin{aligned}
&\left\| \left(\sum_{j,k} \left(\sum_{l,m} \frac{2^{-|l-j|\varepsilon} 2^{(l-j)\frac{n}{\sigma}}}{2_+^{(l-j)\mu}} \left(1 + \frac{|2^{l-j}k - m|}{2_+^{l-j}} \right)^{-\gamma} |h(l,m)| \, \chi_{\sigma,j,k} \right)^q \right)^{1/q} \right\|_{L_p} \\
&\lesssim \left\| \left(\sum_{j,k} \left(\sum_{l} \frac{2^{-|l-j|\varepsilon} 2^{(l-j)\frac{n}{\sigma}}}{2_+^{(l-j)\mu}} 2_+^{(l-j)\frac{n}{t}} \, M_t \left(\sum_m |h(l,m)| \chi_{\infty,l,m} \right) \chi_{\sigma,j,k} \right)^q \right)^{1/q} \right\|_{L_p} \\
&= \left\| \left(\sum_{j} \left(\sum_{l} 2^{-|j-l|\varepsilon} 2_+^{(l-j)(\frac{n}{t} - \mu)} M_t \left(\sum_m |h(l,m)| \chi_{\sigma,l,m} \right) \right)^q \right)^{1/q} \right\|_{L_p} \\
&\lesssim \left\| \left(\sum_{j} \left(M_t \left(\sum_m |h(j,m)| \chi_{\sigma,j,m} \right) \right)^q \right)^{1/q} \right\|_{L_p} \lesssim \left\| \left(\sum_{j} \left(\sum_m |h(j,m)| \chi_{\sigma,j,m} \right)^q \right)^{1/q} \right\|_{L_p} = \|h\|_{\dot{f}_{pq}(\sigma)}.
\end{aligned}$$

$\square$

Corollary 6.14. *Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \le \infty$ and*

$$\lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s, \quad \frac{1}{\sigma} := \frac{s}{n} + \frac{1}{2}.$$

Suppose that $\theta \in \mathcal{D}^{\lambda}_{\gamma} \cap \mathcal{M}^{s}$ and $\zeta \in \mathcal{D}^{s}_{\gamma} \cap \mathcal{M}^{\lambda}$ with $\gamma > n + \max\{s+\lambda, s, \lambda\}$. Then the matrix

$$\mathbf{A} := (\mathbf{A}(j,k;l,m) := \langle \theta_{j,k}, \zeta_{l,m} \rangle : j,l \in \mathbb{Z}, k,m \in \mathbb{Z}^n)$$

defines a bounded operator on $\dot{f}_{pq}(\sigma)$.

Proof of Corollary 6.14: If $\lambda < 0$, we let $\eta := 0$. If $\lambda \ge 0$, we can find a non-integer η such that $\lambda < \eta < \lfloor \lambda \rfloor + 1$, $n + \eta < \gamma$, and $\theta \in \mathcal{R}^{\eta}_{\gamma}$. If $s < 0$, we let $u := 0$. If $s \ge 0$, we can find a non-integer u such that $s < u < \lfloor s \rfloor + 1$, $n + u < \gamma$, and $\zeta \in \mathcal{R}^{u}_{\gamma}$. From Lemma 6.2, we see that $\mathbf{A}$ satisfies (6.4) with $t := n(\frac{1}{\sigma} - \frac{1}{p}) = s + n(\frac{1}{2} - \frac{1}{p})$, $\mu := \lambda + s + n$ and $\varepsilon := \varepsilon(j,l) := \begin{cases} \eta - \lambda, & l \ge j, \\ u - s, & l < j. \end{cases}$
Applying Proposition 6.9 finishes the proof. □

Corollary 6.15. *Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \le \infty$ and*

$$\lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s, \quad \frac{1}{\sigma} := \frac{s}{n} + \frac{1}{2}.$$

Suppose that $\theta \in \mathcal{D}^{s}_{\gamma} \cap \mathcal{M}^{\lambda}$ for some $\gamma > n + \max\{s+\lambda, s, \lambda\}$. Then for every $\zeta \in \mathcal{D}^{\lambda}_{\gamma} \cap \mathcal{M}^{s}$ and $h := (h(l,m) : l \in \mathbb{Z}, m \in \mathbb{Z}^n) \in \dot{f}_{pq}(\sigma)$,

$$\sum_{l,m} |h(l,m)||\langle \theta_{l,m}, \zeta \rangle| < \infty.$$

In particular, the series $\sum_{l,m} h(l,m)\theta_{l,m}$ converges in $\mathcal{S}'/\mathcal{P}$.

Proof: Identical to Corollary 6.6; we only need to appeal to Proposition 6.9 instead of Proposition 6.3. □

Theorem 3.5 can now be proved by following *verbatim* the proof of Theorem 3.7, with the only change being the modified definition of λ. Of course, the applications of Corollary 6.5 and Corollary 6.6 should be replaced by applications of Corollary 6.14 and Corollary 6.15, respectively.

6.3. The interpretation of $\langle f, \psi \rangle$

As we pointed out in the Remark after Theorem 3.7, we need to provide rigorous meaning to the expression

$$\langle f, \psi \rangle,$$

whenever $f \in \dot{F}^{s}_{pq}$ (resp. $\dot{B}^{s}_{pq}$) and ψ satisfying the conditions of Theorem 3.5 (resp. Theorem 3.7). We borrow the interpretation of this pairing that is argued by M. Frazier and B. Jawerth in [FJ2]. Here are the details.

Let $\varphi \in \mathcal{S}$ be a function satisfying the conditions in (3.1). Then, [P], for every $f \in \mathcal{S}'$, there exists a sequence of polynomials (P_N) such that

$$f = \lim_{N\to\infty} \left(\sum_{j=-N}^{\infty} \widetilde{\varphi}_j * \varphi_j * f + P_N \right), \quad \widetilde{\varphi}_j(t) := \overline{\varphi_j(-t)}$$

converges in $\mathcal{S}'$. Furthermore, if we assume that $f \in \dot{F}^s_{pq}$ (resp. $\dot{B}^s_{pq}$), we can find polynomials $P_N \in \mathcal{P}_\eta$, $\eta = \max\{s - \frac{n}{p}, -1\}$, such that

$$g := \lim_{N\to\infty} \left(\sum_{j=-N}^{\infty} \widetilde{\varphi}_j * \varphi_j * f + P_N \right)$$

converges in $\mathcal{S}'$. That is, g is a representative of the equivalence class $f + \mathcal{P}$. It is easy to see that the representative g is well defined modulo $\mathcal{P}_\eta$. In other words, by identifying the equivalence class $f + \mathcal{P}$ with its "canonical" representative g above, the elements of $\dot{F}^s_{pq}$ (resp. $\dot{B}^s_{pq}$) can be regarded as equivalence classes of distributions modulo $\mathcal{P}_\eta$.

Note that if ψ satisfies the conditions of Theorem 3.5 (resp. Theorem 3.6), $\int_{\mathbb{R}^n} t^\alpha \psi(t)dt = 0$ for $|\alpha| \le \eta$. Now we define

$$\langle f, \psi\rangle := \lim_{N\to\infty} \left\langle \sum_{j=-N}^{N} \widetilde{\varphi}_j * \varphi_j * f + P_N, \psi \right\rangle = \sum_{j\in\mathbb{Z}} \langle \widetilde{\varphi}_j * \varphi_j * f, \psi\rangle = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}^n} \langle f, \varphi_{j,k}\rangle \langle \varphi_{j,k}, \psi\rangle,$$

whenever this last sum converges absolutely. The fact, [FJ1], that for every $f \in \mathcal{S}'$ and $t \in \mathbb{R}^n$, $(\widetilde{\varphi}_j * \varphi_j * f)(t) = \sum_{k\in\mathbb{Z}^n} \langle f, \varphi_{j,k}\rangle \varphi_{j,k}(t)$ is used for the last equality.

6.4. A remark on the finiteness of the norm of the CAP detail coefficients

Theorem 3.16 asserts that, given a function f is the relevant TL space, the discrete norm induced by its CAP detail coefficients is finite, and equivalent to the TL norm itself; Theorem 3.18 makes an analogous assertion. Neither of the statements address the question whether an object f whose CAP coefficients have finite norm must lie in the corresponding TL/Besov space. For general elements in $\mathcal{S}'/\mathcal{P}$ the question is not well-posed, since the CAP detail coefficients may not be well-defined. For that reason, one needs to impose *some* restrictions on f. In this section we address the above question for functions $f \in L_2$.

We start with the relevant result concerning *framelet* coefficients.

Lemma 6.16. *Let $s \in \mathbb{R}$, $0 < p, q \le \infty$. Suppose that $(X(\Psi), X(\Psi^{\text{dual}}))$ is a bi-framelet satisfying*

$$\Psi^{\text{dual}} \subset \mathcal{D}^s_\gamma \cap \mathcal{M}^\lambda, \qquad \text{for some } \gamma > n + \max\{s+\lambda, s, \lambda\},$$

where

$$\lambda := n(\frac{1}{\min\{1,p,q\}} - 1) - s, \quad (p < \infty)$$

for Triebel-Lizorkin spaces and

$$\lambda := n(\frac{1}{\min\{1,p\}} - 1) - s$$

for Besov spaces. If $f \in L_2$, and if

$$\sum_{\psi\in\Psi} \left(\sum_{j\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}^n} \left(2^{js} 2^{jn(1/2-1/p)} |\langle f, \psi_{j,k}\rangle| \right)^p \right)^{q/p} \right)^{1/q} < \infty$$

(with the usual modification for $p = \infty$ or $q = \infty$), then $f \in \dot{B}^s_{pq}$. Similarly, if $f \in L_2$, and if, for $p < \infty$,

$$\left\| \left(\sum_{j\in\mathbb{Z}, k\in\mathbb{Z}^n} \left(2^{js} |\langle f, \psi_{j,k}\rangle| \chi_{j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} < \infty, \quad \forall \psi \in \Psi$$

(with the usual modification for $q = \infty$), then $f \in \dot{F}^s_{pq}$.

Proof: We prove the Besov case. The proof of the Triebel-Lizorkin case is similar. Since $f \in L_2$ and $(X(\Psi),\, X(\Psi^{\mathrm{dual}}))$ is a bi-framelet, $T^*_{X(\psi)} f$ is well-defined, for every $\psi \in \Psi$, and

$$\sum_{\psi \in \Psi} T_{X(\psi^{\mathrm{dual}})} T^*_{X(\psi)} f = f,$$

in the sense of L_2. Thus we get

$$T^*_{X(\varphi)} f = \sum_{\psi \in \Psi} (T^*_{X(\varphi)} T_{X(\psi^{\mathrm{dual}})}) T^*_{X(\psi)} f.$$

Since, for each ψ^{dual}, with $t := s + n(1/2 - 1/p)$, $T^*_{X(\varphi)} T_{X(\psi^{\mathrm{dual}})}$ is bounded by Corollary 6.5, we obtain

$$\|T^*_{X(\varphi)} f\|_{\dot{b}_{pq}(t)} \lesssim \sum_{\psi \in \Psi} \|(T^*_{X(\varphi)} T_{X(\psi^{\mathrm{dual}})}) T^*_{X(\psi)} f\|_{\dot{b}_{pq}(t)} \lesssim \sum_{\psi \in \Psi} \|T^*_{X(\psi)} f\|_{\dot{b}_{pq}(t)}.$$

The stated result follows now from Result 3.2. □

We can now borrow some of the arguments that were already used in the proof of Theorem 3.16 to obtain the following result.

Proposition 6.17. *Let (h_c, h_r, h_a) be a CAP triplet, and let $\tau_c, \tau_r, \phi_c, \phi_r, \Theta$ be the associated masks and their refinable functions (not necessarily satisfying Assumption 1.7(d)). Let $s \in \mathbb{R}$, $0 < p, q \le \infty$, and let*

$$\lambda := n(\frac{1}{\min\{1, p, q\}} - 1) - s$$

for Triebel-Lizorkin spaces, and

$$\lambda := n(\frac{1}{\min\{1, p\}} - 1) - s$$

for Besov spaces. Assume that $\phi_r \in \mathcal{D}^s$. Suppose that ϕ_c satisfies the SF conditions of order $> \lambda$ and the CAP system has an order $> \lambda$. If $f \in L_2$, and if

$$\left(\sum_{j \in \mathbb{Z}} \left(2^{-jn} \sum_{k \in \mathbb{Z}^n} |2^{js} d_{j,f}(k)|^p \right)^{q/p} \right)^{1/q} < \infty$$

(with the usual modification for $p = \infty$ or $q = \infty$), then $f \in \dot{B}^s_{pq}$. Similarly, if $f \in L_2$, and if, for $p < \infty$,

$$\left\| \left(\sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^n} \left(2^{js} |d_{j,f}(k)| \chi_{\infty,j,k} \right)^q \right)^{1/q} \right\|_{L_p(\mathbb{R}^n)} < \infty$$

(with the usual modification for $q = \infty$), then $f \in \dot{F}^s_{pq}$.

Proof: We prove only the Triebel-Lizorkin case, since the proof of the Besov case is similar but simpler.

We first recall, from the proof Theorem 3.16 (see the estimation of α there), that

$$\left\|\left(\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^n}\left(2^{js}|d_{j,f}(k)|\chi_{\infty,j,k}\right)^q\right)^{1/q}\right\|_{L_p(\mathbb{R}^n)}$$
$$\approx \sum_{\upsilon\in\{0,1\}^n}\left\|\left(\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^n}\left(2^{js}|d_{j+1,f}(2k-\upsilon)|\chi_{\infty,j,k}\right)^q\right)^{1/q}\right\|_{L_p(\mathbb{R}^n)}.$$

Since $\phi_r \in \mathcal{D}^{\max\{s,0\}}$, by Theorem 1.11 we can select ξ such that $\phi^{\text{dual}} \in \mathcal{D}^{\max\{s,0\}}$, and such that $(X(\Psi), X(\Psi^{\text{dual}}))$ is a bi-framelet, where $\Psi := (\psi_\upsilon)_{\upsilon\in\{0,1\}^n}$ are the current mother CAPlets. As we observed several times, the SF order condition of ϕ_c together with the CAP order condition gives $\Psi^{\text{dual}} \subset \mathcal{M}^\lambda$. Since $f \in L_2$, by Lemma 2.2 we identify the CAP detail coefficient $d_{j+1,f}(2k-\upsilon)$ with the CAPlet coefficient $2^{(j+1)n/2}\langle f, (\psi_\upsilon)_{j,k}\rangle$. Now the CAP bi-framelet $(X(\Psi), X(\Psi^{\text{dual}}))$ satisfies the finiteness assumption in Lemma 6.16, thus we obtain $f \in \dot{F}^s_{pq}$. □

References

[BA] P. J. Burt and E. H. Adelson (1983), "The Laplacian pyramid as a compact image code", *IEEE Trans. Comm.* **31**, 532–540.

[BDR1] C. de Boor, R. DeVore, and A. Ron (1994), "Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^n)$", *Trans. Amer. Math. Soc.* **341(2)**, 787–806.

[BDR2] C. de Boor, R. DeVore, and A. Ron (1994), "The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^n)$", *J. Funct. Anal.* **119(1)**, 37–78.

[BGN1] L. Borup, R. Gribonval, and M. Nielsen (2004), "Tight wavelet frames in Lebesgue and Sobolev spaces", *J. Funct. Spaces Appl.* **2(3)**, 227–252.

[BGN2] L. Borup, R. Gribonval, and M. Nielsen (2004), "Bi-framelet systems with few vanishing moments characterize Besov spaces", *Appl. Comput. Harmon. Anal.* **17(1)**, 3–28

[BHR] C. de Boor, K. Höllig, and S. D. Riemenschneider (1993), *Box Splines*, Springer-Verlag.

[CDF] A. Cohen, I. Daubechies, and J. C. Feauveau (1992), "Biorthogonal bases of compactly supported wavelets", *Comm. Pure Appl. Math.* **45(5)**, 485–560.

[D] I. Daubechies (1992), *Ten Lectures on Wavelets*, SIAM.

[DHRS] I. Daubechies, B. Han, A. Ron, and Z. Shen (2003), "Framelets: MRA-based constructions of wavelet frames", *Appl. Comput. Harmon. Anal.* **14(1)**, 1–46.

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

[DN] G. Davis and A. Nosratinia (1998), "Wavelet-based image coding: an overview", *Applied and computational control, signals, and circuits.* **1**, 369–434.

[FJ1] M. Frazier and B. Jawerth (1985), "Decomposition of Besov spaces", *Indiana Univ. Math. J.* **34(4)**, 777–799.

[FJ2] M. Frazier and B. Jawerth (1990), "A discrete transform and decompositions of distribution spaces", *J. Funct. Anal.* **93(1)**, 34–170.

[FS] C. Fefferman and E. Stein (1971), "Some maximal inequalities", *Amer. J. Math.* **93**, 107–115.

[HB] D. Heeger and J. Bergen (1995), "Pyramid-based texture analysis/synthesis", *Computer Graphics Proceeding, SIGGRAPH 95*, 229–238.

[HW] E. Hernández and G. Weiss (1996), *A first course on wavelets*, CRC Press.

[JMR] S. Jaffard, Y. Meyer and R. Ryan (2001), *Wavelets: Tools for Science & Technology*, SIAM.

[JZ1] K. Jetter and D. X. Zhou (1995), "Order of linear approximation from shift invariant spaces", *Constr. Approx.* **11(4)**, 423–438.

[JZ2] K. Jetter and D. X. Zhou (1998), "Order of linear approximation on finitely generated shift invariant spaces", preprint.

[K] G. Kyriazis (2003), "Decomposition systems for function spaces", *Studia Math.* **157(2)**, 133–169.

[Ma] S. Mallat (1989), "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Analysis and Machine Intelligence* **11(7)**, 674–693.

[Me1] Y. Meyer (1992), *Wavelets and operators*, Cambridge University Press.

[Me2] Y. Meyer and R. Coifman (1997), *Wavelets: Calderón-Zygmund and multilinear operators*, Cambridge University Press.

[P] J. Peetre (1976), "New thoughts on Besov spaces", *Duke University Mathematics Series* **1**, Mathematics Department, Duke University.

[R1] A. Ron (1997), "Smooth refinable functions provide good approximation orders", *SIAM J. Math. Anal.* **28(3)**, 731–748.

[R2] A. Ron (1998), "Wavelets and their associated operators", in *Approximation theory IX (ed. C. K. Chui and L. L. Schumaker), Vol. II*, Vanderbilt Univ. Press, 283–317.

[RS] A. Ron and Z. Shen (1997), "Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator", *J. Funct. Anal.* **148(2)**, 408–447.

[T] H. Triebel (1983), *Theory of function spaces*, Birkhäuser.

[TP] S. Toelg and T. Poggio (1994), "Towards an example-based image compression architecture for video-conferencing", *A.I. Memo No. 1494, M.I.T.*