L-CAMP: Extremely local high-performance wavelet representations in high spatial dimension

Youngmi Hur and Amos Ron

Abstract— **A new wavelet-based methodology for representing data on regular grids is introduced and studied. The main attraction of this new L-CAMP methodology is in the way it scales with the spatial dimension, making it, thus, highly suitable for the representation of high dimensional data. The specific highlights of the L-CAMP methodology are three. First, it is computed and inverted by fast algorithms with linear complexity and very small constants; moreover, the constants in the complexity bound decay, rather than grow, with the spatial dimension. Second, the representation is accompanied by solid mathematical theory that reveals its performance in terms of the maximal level of smoothness that is accurately encoded by the representation. Third, the localness of the representation, measured as the sum of the volumes of the supports of the underlying mother wavelets, is extreme. An illustration of this last property is done by comparing the L-CAMP system that is marked in this paper as V with the widely used tensor-product biorthogonal 9/7. Both are essentially equivalent in terms of performance. However, the L-CAMP V has in 10D localness score < 29. The localness score of the 9/7 is, in that same dimension, $> 575,000,000,000$.**

Index Terms— **wavelets, multidimensional wavelets, fast wavelet transforms, wavelet frames, Unitary Extension Principle, L-CAMP, performance, fast algorithms, extremely local wavelets.**

I. Introduction

A (dyadic) wavelet system $X(\Psi)$ is a collection of linear functionals defined on $\mathbb{R}^n$ that are obtained by applying integer translations and dyadic dilations to a finite set Ψ of mother wavelets:

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

Here, $(Df)(t) := 2^{n/2} f(2t)$, while $(E^k f)(t) := f(t-k)$. The mother wavelets Ψ are assumed to lie in

$$L_2(\mathbb{R}^n) := \{f : \mathbb{R}^n \to \mathbb{C} : \ \|f\|^2 := \int_{\mathbb{R}^n} |f|^2 < \infty\}.$$

The *wavelet representation* of $f \in L_2(\mathbb{R}^n)$ is then the discrete set of inner products

$$(\langle f, x\rangle)_{x \in X(\Psi)}, \quad \langle f, g\rangle := \int_{\mathbb{R}^n} f(t)\overline{g(t)}\, dt.$$

This work was supported by the US National Science Foundation under Grant ANI-0085984, by the National Institute of General Medical Sciences under Grant NIH-1-R01-GM072000-01, and by the Vilas Foundation at the University of Wisconsin.

Y. Hur is with the Mathematics Department, University of Wisconsin, Madison, WI 53706-1388 USA (e-mail:hur@math.wisc.edu)

A. Ron is with the Computer Sciences Department, University of Wisconsin, Madison, WI 53706-1685 USA (e-mail:amos@cs.wisc.edu)

The wavelet representation is one of the major representations for data defined on regular grids. There are two main reasons for the popularity of this representation. Firstly, its discrete version is computed and inverted by a fast algorithm, the so-called Fast Wavelet Transform (FWT) [16]. Secondly, it is known to provide optimally sparse representations for the "right type" of functions/datasets (see, e.g., [7], [17]). We refer in this paper to this latter issue as "performance", and actually distinguish between two different types of performance, Jackson-type and Bernstein-type. Let us pause momentarily for a brief explanation of these notions.

Jackson-type performance guarantees that the wavelet coefficients decay, as a function of the dilation level j, in a way that corresponds to the smoothness class of f. Bernstein-type performance guarantees that the decay will not be "too fast". Taken together, performance is essential for the correct detection and classification of local singularities: the more subtle the singularity is, the higher performance is needed for its detection. The Jackson-type performance of a wavelet system is intimately related to vanishing moments of its mother wavelets. Recall that the system is said to have s *vanishing moments* if the Fourier transform $\widehat{\psi}$ of each of the mother wavelets has a zero of order s at the origin. A comprehensive discussion of performance is given in the sequel.

We are interested in this paper in wavelet representations in high-dimensions. The construction of effective wavelet representations in high spatial dimension is a challenging problem. At a first glance, the choice falls on the so-called tensor-product constructions, that work by "lifting" a univariate wavelet construction to n-dimensions. These constructions are readily available, their performance is well-understood, they are simple, convenient, and, to a degree, computationally effective. However, as the spatial dimension grows, such constructions become immensely non-local in space. Let us illustrate this by the following simple example. Suppose that our construction is in $\mathbb{R}^5$, and that we require Jackson-type performance $s = 4$, which essentially means that all the wavelets have four vanishing moments. A standard choice would be to use the tensor-product of Daubechies' 8-tap filters (=:Daub 8) [3] (or one of the related biorthogonal ones, like Bior 9/7 or Bior 10/6, [2]). The 5-D tensor-product construction yields 31 mother wavelets. For the case here, the sum of the support volumes of these 31 mother wavelets is approximately 5×10^5. This means that every point in space is visited approximately half million times by the wavelets within a single scale! That does not sound "local" at all, especially in view of the fact

TABLE I
THE PERFORMANCE AND LOCALNESS NUMBERS OF OUR **L-CAMP SYSTEMS I, II, V, VII** ARE COMPARED WITH THE TENSOR-PRODUCTS OF BENCHMARK WAVELETS: DAUB 4, DAUB 8, BIOR 5/3, BIOR 9/7.

	Daub 4	Bior 5/3	**L-CAMP I**	**L-CAMP II**	Daub 8	Bior 9/7	**L-CAMP V**	**L-CAMP VII**
s_J	2	2	**2**	**2**	4	4	**4**	**4**
s_B	0.55	1	**1.41**	**2**	1.62	1.70	**2.02**	**4**
vol(Ψ) ($n = 3$)	189	279	**4.6**	**5.6**	2401	2863	**14.4**	**31.1**
vol(Ψ) ($n = 4$)	1215	2145	**4.8**	**5.8**	36015	46529	**16.7**	**37.6**
vol(Ψ) ($n = 5$)	7533	15783	**4.9**	**5.9**	521017	726607	**18.8**	**43.8**

that the dimension of the cubic polynomial space Π_3^5 in 5 variables is 56. The latter fact implies the existence of a single piecewise-polynomial function whose volume of support is 56 and whose shifts provide approximation order 4. While, admittedly, spline approximation and wavelet decomposition are not exactly comparable, one must be alarmed, and rightly so, by the gap between the two numbers: in comparison with the 56, half a million looks an awfully large number.

Our goal in this paper is to introduce an algorithm that, for a given fixed performance level s (we will deal concretely with $s = 2, 3, 4$) and a spatial dimension n, yields a wavelet system, to which we refer as an L-CAMP system, generated by the mother wavelets $\Psi := \Psi(n, s) \subset L_2(\mathbb{R}^n)$ such that:

(1) The performance of the representation matches the given grade s. One can choose here to accept Jackson-type performance or to insist on Bernstein-type performance.

(2) The representation can be computed by the FWT, hence with linear complexity.

(3) The representation can be inverted by an algorithm which is different from, *and is at least as fast as*, the standard inversion of the FWT.

(4) A complete cycle of one decomposition step and its inversion is not only of linear complexity, but the constant in the $O(N)$ bounds, where N is the size of an initial data to be analysed, does not grow with the spatial dimension. In fact, it decays slightly with the dimension!

(5) The representation is *extremely local*: The L-CAMP mother wavelets Ψ satisfy

$$\mathrm{vol}(\Psi) := \sum_{\psi\in\Psi} \mathrm{vol}(\mathrm{supp}\,\psi) \ll \binom{n+s-1}{n}. \qquad (1)$$

As a glimpse into column 7 of Table IV reveals, the volume $\mathrm{vol}(\Psi)$ grows, at worst, linearly with the spatial dimension n; in some of the constructions, it does not grow at all! □

Note that $\binom{n+s-1}{n}$ in (1) is the dimension of the space Π_{s-1}^n of (s–1)-degree polynomials in n variables, a number that a few paragraphs above was considered to be very small in comparison with the volume of mainstream wavelet systems. So, the L-CAMP representation is even more local than the most local spline approximation scheme. Concretely, in column 7 of Table IV, we see that for the case $n = 5$ and Jackson-type performance grade $s_J = 4$, two of our L-CAMP systems, system V and system VI, satisfy $\mathrm{vol}(\Psi) \approx 19$ (V), and $\mathrm{vol}(\Psi) \approx 24$ (VI). This should be considered a dramatic improvement over the $\mathrm{vol}(\Psi) \approx 5 \times 10^5$ of the tensor-product construct that was detailed earlier. It is even better than the spline's 56. Had we wanted the comparison to look more dramatic, we could have chosen a higher dimension n: for $n = 10$, for example, the L-CAMP volumes of the above constructs are 29 and 39 respectively. The volume of the tensor-product construct is then about $289,000,000,000$...

In Table I, we compare our L-CAMP systems I, II, V, VII with the tensor-products of some of the benchmark wavelet systems: Daub 4, Daub 8, [3] and Bior 5/3, Bior 9/7, [2]. The spatial dimensions in these comparisons are $n = 3, 4, 5$. The performance grades s_J (Jackson-type) and s_B (Bernstein-type) that are listed in the table are defined in the sequel.

The L-CAMP system was discovered as a variation of the CAMP scheme of [13], which by itself was derived as a variation of the class CAP (Compression-Alignment-Prediction) of pyramidal representations in the same paper. "M" in "CAMP" stands for "modified", while "L" in "L-CAMP" stands for "local". The CAP systems are close relatives of Burt-Adelson's Laplacian pyramid algorithm [1], which is used in many image processing applications [23], [21], [12], [8], [9]. Since understanding the Laplacian pyramid and/or CAP/CAMP theory is not necessary for understanding the L-CAMP representation, we do not pursue all these connections. Instead, we construct the L-CAMP systems from scratch and analyse directly their performance.

Throughout the paper, we use the following notation. 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 $|t| := \sqrt{t(1)^2 + \cdots + t(n)^2}$ and $|\beta| := \beta(1) + \cdots + \beta(n)$. 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., $f \in L_1(\mathbb{R}^n)$):

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

Given $f : \mathbb{R}^n \to \mathbb{R}$, we denote

$$f_{j,k} := D^j E^k f = 2^{j\frac{n}{2}} f(2^j \cdot -k), \quad j \in \mathbb{Z},\, k \in \mathbb{Z}^n.$$

Here, as before, we used

$$(Df)(t) = 2^{n/2} f(2t), \quad (E^k f)(t) = f(t - k). \qquad (2)$$

We let χ be the characteristic function of the unit cube, and let $\mathbf{1} := (1, \ldots, 1) \in \mathbb{Z}^n$. Furthermore, we let $\mathbf{E} := \{0, 1\}^n$, the set of the vertices of the unit cube and let $\mathbf{E}' := \mathbf{E}\backslash 0$.

The outline of the paper is as follows. In Section II, we introduce the L-CAMP systems. The extreme localness of these systems, as well as their corresponding performance are discussed there. In Section III, we present an algorithm for computing and inverting the L-CAMP representation, and examine its complexity. In Section IV-A, we review some of the basics of framelet (:= wavelet frame) theory, and use that for a rigorous discussion of the Jackson-type and Bernstein-type performance of wavelet systems. In Section IV-B, the dual system of the L-CAMP system is introduced. The main results concerning the performance of L-CAMP systems are stated in Section IV-C, with their proofs placed in Appendix I. In Section V, we finally construct concrete L-CAMP systems for performance grades $s = 2, 3, 4$. Those constructs are valid in any dimension n, provided that $n > 2$. The lower-dimensional L-CAMP systems are given in Appendix II.

II. The L-CAMP system introduced

The L-CAMP class in this paper is based on the support function $\phi_c := \chi$ of the unit cube. The function ϕ_c is refinable with mask $\tau_c(\omega) = \prod_{l=1}^{n} \frac{1+e^{-i\omega(l)}}{2}$, viz.,

$$\widehat{\phi}_c(2\omega) = \tau_c(\omega)\widehat{\phi}_c(\omega), \quad \omega \in \mathbb{R}^n.$$

Here and later, we use

$$\omega := (\omega(1), \ldots, \omega(n)) \in \mathbb{R}^n$$

to denote the generic point in the frequency domain. Let h_c be the filter associated with $\overline{\tau}_c$ (i.e. $\widehat{h}_c = \overline{\tau}_c$). That is,

$$h_c(-k) := \begin{cases} 2^{-n}, & k \in \mathbf{E}, \\ 0, & \text{otherwise.} \end{cases}$$

More general L-CAMP constructions, that use refinable functions ϕ_c other than χ, are available. However, the choice $\phi_c := \chi$ leads to the most local systems we are able to come up with, makes things concrete and simple, and makes the algorithms faster (in terms of the constants). Most importantly, and in contrast with common misperceptions (that are, perhaps, due to the abysmal performance of the Haar system), the use of piecewise-constant constructs does not impede the ability to obtain high-performance systems; as a matter of fact, one can obtain systems with as high performance as one wishes to. The more sophisticated constructions that we alluded to above, may be important only in specific applications. A review of the graph in Fig. 1 may be useful here: functions in a smoothness class outside the lined area may not be represented well by the piecewise-constant version of our L-CAMP methodology. For example, piecewise-constant L-CAMP lacks the ability to detect most types of *negative* smoothness, a property that might be necessary for some PDE applications. The piecewise-constant L-CAMP, in addition, fails to encode correctly the Hardy space $H_1(\mathbb{R}^n)$ (however, it does encode correctly smoother versions of $H_1(\mathbb{R}^n)$).

The total number of mother wavelets in our L-CAMP construction is 2^n. Recall that, in n-D, the minimal number of mother wavelets is $2^n - 1$, hence that we are slightly redundant. We index the L-CAMP mother wavelets by the set $\mathbf{E}$, i.e., ψ_ν, $\nu \in \mathbf{E}$.

The entire L-CAMP construction is based on two lowpass filters, and on nothing else. One of these filters is used to define the mother wavelet ψ_0, while the other one is used to define all the remaining mother wavelets. We refer to the first one as the *enhancement filter*, and to the latter one as the *main filter*.

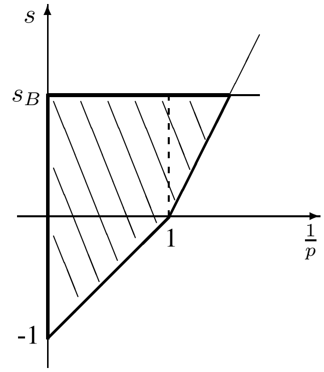

Fig. 1. The polygon with thick boundary captures the performance range of L-CAMP systems that are based on $\phi_c := \chi$. For a pair (p, s) inside the polygon, or on its left vertical boundary, the L-CAMP representation encodes accurately the property of "s derivatives in L_p". While the upper boundary moves upward as the Bernstein-performance grade s_B of the system gets higher, the lower boundary is fixed, and is due to the choice $\phi_c = \chi$.

We start by selecting the *mask* τ_e of the enhancement filter. It can be any trigonometric polynomial (in n-variables). Initially we require that the mask satisfies the relation

$$1 - \overline{\tau_e(2\omega)}\tau_c(\omega) = O(|\omega|^s), \quad \text{near the origin.} \tag{3}$$

We say that τ_e *is of order* s. While a high s is desired here, we require, at a minimum, that $s \geq 2$. The enhancement filter h_e is, then, the filter associated with the mask τ_e, i.e., $\tau_e = \widehat{h}_e$. We further would like the number of taps of h_e to be as small as possible. The third, and final, condition that the enhancement filter should satisfy is detailed later.

Next, we define the first mother wavelet, ψ_0, by the relation

$$\widehat{\psi}_0 := \alpha_n \left(\widehat{\phi}_c(\cdot/2) - \overline{\tau}_e \widehat{\phi}_c \right). \tag{4}$$

Here, $\alpha_n := 2^{-n/2}$. Note that

$$\psi_0 = \alpha_n \left(2^n \phi_c(2\cdot) - \sum_{k \in \mathbb{Z}^n} \overline{h_e(k)} \phi_c(\cdot + k) \right).$$

In order to define the remaining $2^n - 1$ mother wavelets, we choose a univariate mask τ, denote its univariate filter by h, and refer to it as the *main filter*. The main filter should also satisfy three conditions. The first condition is that it is (finite and) supported on the *odd* integers, i.e.,

$$h(2m) = 0, \quad m \in \mathbb{Z}.$$

The second condition is that the filter will have high order of polynomial accuracy N: we say that h has *accuracy* N if

$$h * P = P, \quad \forall P \in \Pi^1_{N-1}. \tag{5}$$

Recall that the main filter is univariate, hence the accuracy test is conducted on univariate polynomials. A third condition that is required of the main filter is detailed later.

TABLE II

THREE STANDARD CHOICES FOR THE MAIN FILTER. THE FIRST COLUMN LISTS THE ACCURACY. THE FIRST ROW LISTS THE DOMAIN OF THE FILTERS (WHICH IS A SUBSET OF THE ODD INTEGERS). ONLY NON-ZERO VALUES OF THE FILTERS ARE LISTED.

$N \setminus k$	-5	-3	-1	1	3	5
2			$\frac{1}{2}$	$\frac{1}{2}$		
4		-$\frac{1}{16}$	$\frac{9}{16}$	$\frac{9}{16}$	-$\frac{1}{16}$	
6	$\frac{3}{256}$	-$\frac{25}{256}$	$\frac{150}{256}$	$\frac{150}{256}$	-$\frac{25}{256}$	$\frac{3}{256}$

Standard choices for the main filter are listed in Table II. We note that for each filter h with accuracy N in the table, $(\delta+h)/2$ is the Deslauriers-Dubuc interpolatory filter of order N, [6], where δ is the *dirac sequence*.

We lift the main filter to n-dimensions by aligning it along one of the coordinate axes. There are n different ways to do it, i.e., for $l = 1, \ldots, n$,

$$\tau_l(\omega) := \tau(\omega(l)).$$

The mother wavelet ψ_ν, $\nu \in \mathbf{E}'$, is defined by the relation

$$e^{i\nu\cdot\omega}\widehat{\psi}_\nu(2\omega) = \alpha_n\left(1 - \overline{\tau_{\lceil\nu\rceil}}(\omega)\right)\widehat{\phi}_c(\omega). \tag{6}$$

Here, $\alpha_n = 2^{-n/2}$ as before, and

$$\lceil\ \rceil : \mathbf{E}' \to \{1, \ldots, n\}$$

is a map that determines the orientation of τ that is assigned to the ν-mother wavelet. While the assignment $\lceil\ \rceil$ cannot be done at random, there is a great deal of flexibility in choosing it. One way for defining $\lceil\ \rceil$ goes as follows: for each $\nu \in \mathbf{E}'$,

$$\lceil\nu\rceil := \text{the position of the last 1-digit in the vector } \nu. \tag{7}$$

Thus, for example, for $n = 2$,

$$\lceil(0,1)\rceil = \lceil(1,1)\rceil = 2, \quad \lceil(1,0)\rceil = 1.$$

We extend the domain of the map $\lceil\ \rceil$ to $\mathbf{E}$ by defining $\lceil 0\rceil := 0$. Under this convention, the valuation map $\lceil\ \rceil$ of (7) satisfies

$$\lceil\nu - \mathbf{e}_{\lceil\nu\rceil}\rceil < \lceil\nu\rceil, \quad \nu \in \mathbf{E}'. \tag{8}$$

Here, $\mathbf{e}_l$ is the lth vector in the standard basis for $\mathbb{R}^n$.

Note that, if h is (λ-1)-tap, then $\operatorname{supp}\psi_\nu$ is the union of λ cubes that are aligned along the $\lceil\nu\rceil$-axis, each of which with volume 2^{-n}. Hence,

$$\operatorname{vol}(\operatorname{supp}\psi_\nu) = \frac{\lambda}{2^n}, \quad \nu \in \mathbf{E}'.$$

Defining, as before, the *volume of* $\Psi \subset L_2(\mathbb{R}^n)$ to be

$$\operatorname{vol}(\Psi) := \sum_{\psi\in\Psi} \operatorname{vol}(\operatorname{supp}\psi),$$

we obtain that the volume of the L-CAMP mother wavelet set

$$\Psi := \{\psi_\nu : \nu \in \mathbf{E}\} \tag{9}$$

is

$$\operatorname{vol}(\Psi) = \operatorname{vol}(\operatorname{supp}\psi_0) + \frac{\lambda(2^n - 1)}{2^n} \tag{10}$$

$$< \operatorname{vol}(\operatorname{supp}\psi_0) + \lambda. \tag{11}$$

That is, the total volume of Ψ is bounded by $\operatorname{vol}(\operatorname{supp}\psi_0) + \lambda$.

We mentioned so far two conditions that we require of the enhancement filter (high order, and small support; the latter, obviously, is needed for localness and has nothing to do with performance), and two conditions that we require of the main filter (support at odd integers and high polynomial accuracy). Only one additional condition is required here, but it is not as simple as the ones above.

Performance conditions. Our performance analysis of the L-CAMP system is based on the following parameters:

- The order $s \geq 2$ of the enhancement mask (cf. (3)).
- The accuracy $N \geq 2$ of the main filter (cf. (5)).
- The Hölder smoothness α of the n-dimensional refinable function $\widetilde{\phi}$ associated with the mask

$$\tau_e\left(\prod_{l=1}^{n} \frac{1+\tau_l}{2}\right). \tag{12}$$

For Jackson-type performance, we need to assume that

$$\alpha > 0. \tag{J}$$

The performance is then related to $\min\{s, N\}$. In all our concrete constructions, this minimum is s.
Bernstein-type performance is related to $\min\{s, N, \alpha\}$, which, again, will coincide in our constructions with $\min\{s, \alpha\}$. So, for this type of performance we desire that

$$\alpha \geq s, \tag{B}$$

or at least that α does not lag far behind s. □

Remark. Constructing good main filters with short support is easy. Constructing enhancement filters with high orders and small support is not too hard. The true challenge in the L-CAMP theory is to obtain high values of α. The main challenge is related to the fact that the definition of the enhancement filter depends on the underlying dimension, and its support size usually grows with the dimension. All that said, and as we will see later, we provide lower bounds on the smoothness of the refinable function $\widetilde{\phi}$ that do not degrade with the dimension. We had to develop to this end new techniques for estimating the smoothness of refinable functions in arbitrary dimensions. It is beyond the scope of this paper to provide the details of our smoothness estimation machinery; it will be detailed elsewhere. One thing must be clear here: any smoothness analysis that relies on estimating the asymptotics of some numerical experiment (e.g., iterating numerically with the transfer operator) is prohibitive here, due to the need to have the smoothness estimation valid in all dimensions. □

Extreme Localness. The rule of thumb in all our constructions is that the number λ in (11) is smaller, usually much smaller, than the support size of the single mother wavelet ψ_0. This means that the total volume $\operatorname{vol}(\Psi)$ from (11) is dominated by the term $\operatorname{vol}(\operatorname{supp}\psi_0)$. In the constructions we present in this paper this latter number grows no faster than linearly with the spatial dimension n (cf. Table IV).

We might pause again to compare the above with the main-stream tensor-product constructions. If we start, for example, with an orthonormal univariate wavelet whose support length is L, then the tensor-product approach yields $2^n - 1$ mother wavelets, each of which with support volume L^n. Thus, the total volume of mother wavelets in this case is $L^n(2^n - 1)$, which grows exponentially with the spatial dimension n. This exponential growth has nothing to do with scientific reality, and is a mere artifact of the construction methodology. □

III. Fast algorithms for computing and inverting the representation

The L-CAMP representation can be computed and inverted by fast algorithms with linear complexity and small constants.

L-CAMP decomposition and reconstruction algorithms. Let h_e be the n-dimensional enhancement filter and let h be the 1-dimensional main filter. For $l = 1, \ldots, n$, let h_l be the lifting of h to an n-dimensional filter in the l-coordinate direction. Let $\lceil\ \rceil$ be the valuation map from (7). Then:

input $y_0 : \mathbb{Z}^n \to \mathbb{C}$

(1) Decomposition:

```
for j = -1, -2, ..., j_0
    y_j(k) = 2^{-n} Σ_{μ∈E} y_{j+1}(2k + μ),  k ∈ Z^n  (*)
    if k ∈ 2Z^n
        d_{j+1}(k) = y_{j+1}(k) - (h_e * y_j)(k/2)
    end
    if k ∈ ν + 2Z^n where ν ∈ E'
        d_{j+1}(k) = y_{j+1}(k) - (h_⌈ν⌉ * y_{j+1})(k)
    end
end
```

(2) Reconstruction:

```
for j = j_0, ..., -1
    if k ∈ 2Z^n
        y_{j+1}(k) = d_{j+1}(k) + (h_e * y_j)(k/2)
    end
    for ⌈ν⌉ = 1, ..., n  (**)
        if k ∈ ν + 2Z^n
            y_{j+1}(k) = d_{j+1}(k) + (h_⌈ν⌉ * y_{j+1})(k)
        end
    end
end
```

We note that the resulted MRA $(y_j)_{j\le 0}$ from the line marked by (⋆) is the MRA associated with χ, that is, assuming $y_0(k) = \langle f, \chi_{0,k}\rangle$, $k \in \mathbb{Z}^n$, for some function f, it follows that

$$y_j(k) = 2^{jn/2}\langle f, \chi_{j,k}\rangle, \quad j < 0,\ k \in \mathbb{Z}^n.$$

After that line, the rest of the decomposition step computes the *detail coefficients* d_{j+1}. Note that we use 2^n different rules to extract the detail coefficients d_{j+1}. In the signal analysis literature, such decomposition methods are known as *Polyphase decomposition* [24], [25], [9]. The novelty of our decomposition algorithm lies in the simple way we define the 2^n rules using a single n-dimensional filter h_e and a single 1-dimensional filter h, as well as in the ability to pin down the precise performance of the representation. The interpretation of the detail coefficients is standard up to normalization, i.e., retaining the same assumption on the initial data y_0, one proves that for $j < 0$, $\nu \in \mathbf{E}$, $k \in \mathbb{Z}^n$,

$$d_{j+1}(\nu + 2k) = 2^{(j+1)n/2}\langle f, (\psi_\nu)_{j,k}\rangle.$$

The reconstruction step does not resemble its FWT counterpart. The crucial step in the reconstruction is the *for loop* marked by (⋆⋆). We observe that, if $h_{\lceil\nu\rceil}(l) \neq 0$, then l must be of the form

$$l = a\,\mathbf{e}_{\lceil\nu\rceil},$$

where $\mathbf{e}_{\lceil\nu\rceil}$ is the unit vector in the $\lceil\nu\rceil$-coordinate direction, and a is an *odd* number. This means that all the values of y_{j+1} that are needed for the computation of $(h_{\lceil\nu\rceil} * y_{j+1})(k)$, $k \in \nu + 2\mathbb{Z}^n$, lie in $\nu - \mathbf{e}_{\lceil\nu\rceil} + 2\mathbb{Z}^n$. Since $\lceil\nu - \mathbf{e}_{\lceil\nu\rceil}\rceil < \lceil\nu\rceil$ for any $\nu \in \mathbf{E}'$ from (8), we have already recovered those values of y_{j+1} previously, hence we are able to compute $y_{j+1}(k)$ as above.

Complexity. We measure complexity by counting the number of "operations" needed in order to fully derive y_j and d_{j+1} from y_{j+1}, and add the number of operations needed for the inversion. Here, we define "an operation" as the need to fetch an entry from some of our arrays/vectors. Thus, for example, computing one entry in y_j from y_{j+1} as in (⋆) requires 2^n operations.

Obviously, the complexity here is linear, i.e., $\sim CM$, with M the number of non-zero entries in y_0, and C some constant. Our goal is to estimate that constant: since M is expected to grow exponentially fast with the dimension, we need, at least, to control very tightly that constant! So, we actually compute the mean number of operations per one single entry in y_0.

We observe that the number of operations required to process the portion of y_j that lies in a cube of lengthsize 2 is about

$$2^n + 2(1 + \text{ tap-size of } h_e) + 2\lambda(2^n - 1). \tag{13}$$

This means that the cost per entry of performing one complete cycle of decomposition/inversion is bounded by

$$1 + 2(\lambda + 2^{-n}(1 + \text{ tap-size of } h_e)). \tag{14}$$

Since the tap-size of the enhancement filters grows very slowly (with the dimension) in comparison with the exponential 2^n, the complexity constant is dominated by the term $1 + 2\lambda$. It is then important to note that we are able to achieve the required level of performance s, simultaneously in *all* spatial dimensions, without increasing λ. The concrete values of λ in the constructions in this paper are $\lambda = 3, 5, 7$ (Table II). Consequently, the constants in the complexity estimation are varying from 7 up to 15, depending on the requisite performance, and independently of the dimension. Cf. Table IV for details.

We would like to point out that the switch from constructing L-CAMP systems with a specified Jackson-level performance to systems that deliver a similar Bernstein-level performance barely changes the above complexity constants: the enhancement filters in the Bernstein case can be factored into $h_e = \widetilde{h}_e * h_{sm}$, where h_{sm} is a simple smoothing filter (cf. (22) and (23) in Section V). The tap size of the $\widetilde{h}_e$ factor in this decomposition is on par with the total size of the enhancement filter in the Jackson case (the latter does not require smoothing). Since the smoothing step has a negligible effect on the complexity constants, we see that "Bernstein performance is free", at least as far as the computation of the representation is concerned. The modified numbers, that account for the smoothing factorization, are listed in column 6 of Table IV. □

When comparing the above details to the details of computing and inverting a tensor-product-based wavelet representation, we must account to the fact that the latter can be computed quite effectively by algorithms that bypass completely the immense non-localness of the presentation. A simple efficient implementation of such systems requires, for an orthonormal univariate wavelet with L taps, about nL operations per entry for decomposition only. We are not aware of an implementation of the decomposition step of the FWT whose constant is similar to ours, i.e., independent of the dimension.

IV. Performance analysis

The L-CAMP scheme was discovered as a variation of the CAMP scheme of [13], which by itself was derived as a variation of the class CAP of pyramidal representations, studied in the same paper. Rather than tracing back that evolution, we will provide here an intrinsic analysis of the L-CAMP performance.

A. The performance of wavelet frames

Let Ψ be a finite subset of $L_2(\mathbb{R}^n)$. The *wavelet system* generated by the *mother wavelets* Ψ is the family

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

The *analysis operator* is defined as

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

the entries of $T^*_{X(\Psi)} f$ are the *wavelet coefficients* of f (with respect to the system $X(\Psi)$). The system $X(\Psi)$ is a *frame* if the analysis operator is bounded above and below, viz., if there exist two positive constants A, B such that

$$A\, \|f\|^2_{L_2(\mathbb{R}^n)} \le \sum_{x \in X(\Psi)} |\langle f, x\rangle|^2 \le B\, \|f\|^2_{L_2(\mathbb{R}^n)}, \qquad (15)$$

for all $f \in L_2(\mathbb{R}^n)$. $X(\Psi)$ is a *Bessel system* if $T^*_{X(\Psi)}$ is bounded, i.e., the right-hand side of (15) is valid.

We pay attention here only to wavelet frames that are derived from a *multiresolution analysis* (MRA) ([16], [17], [20], [4]). One begins with the selection of a function $\phi \in L_2(\mathbb{R}^n)$. With ϕ in hand, one defines

$$V_0 := V_0(\phi)$$

to be the closed linear span of the shifts of ϕ, i.e., V_0 is the smallest closed subspace of $L_2(\mathbb{R}^n)$ that contains $E(\phi) := \{\phi(\cdot - k) : k \in \mathbb{Z}^n\}$. Then, with D the operator of dyadic dilation (cf. (2)),

$$V_j := V_j(\phi) := D^j V_0(\phi), \quad j \in \mathbb{Z}.$$

The primary condition of the MRA setup is that the $(V_j)_j$ sequence is nested:

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots.$$

Whenever this condition holds, one refers to ϕ as a *refinable function*. In addition, one requires that the union $\cup_j V_j$ is dense in $L_2(\mathbb{R}^n)$. However, if ϕ is compactly supported and $\widehat{\phi}(0) \neq 0$, the density condition always holds.

Next, we illustrate the way the "performance" of a wavelet frame $X(\Psi)$ may be graded, and use the L_2-setup[1] to this end. For $\alpha > 0$, let $W_2^\alpha(\mathbb{R}^n)$ be the usual Sobolev space. That is, $W_2^\alpha(\mathbb{R}^n)$ is the set of functions $f \in L_2(\mathbb{R}^n)$ such that

$$|f|_{W_2^\alpha(\mathbb{R}^n)} := \|(|\cdot|^\alpha \widehat{f}\,)^\vee\|_{L_2(\mathbb{R}^n)} < \infty.$$

We would like first the wavelet system $X(\Psi)$ to be a frame and to satisfy

$$\sum_{\psi \in \Psi} \|T^*_{X(\psi)} f\|_{\ell_2(\alpha)} \le A_\alpha |f|_{W_2^\alpha(\mathbb{R}^n)}, \quad \forall f \in W_2^\alpha(\mathbb{R}^n). \qquad (16)$$

Here,

$$\|T^*_{X(\psi)} f\|_{\ell_2(\alpha)} := \left(\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^n} 2^{2j\alpha} |\langle f, \psi_{j,k}\rangle|^2 \right)^{1/2}. \qquad (17)$$

The supremum

$$s_J := \sup\{\alpha > 0 :\ X(\Psi) \text{ satisfies (16) for the given } \alpha\},$$

is one way to quantify the "performance-grade" of a frame $X(\Psi)$. Since the inequality (16) is the counterpart of the Jackson-type inequalities in Approximation Theory, we refer to the above s_J as the *Jackson-type performance* of $X(\Psi)$. It is known that the essential condition Ψ needs to satisfy for having "performance-grade" s_J is that each $\psi \in \Psi$ has s_J *vanishing moments* :

$$\widehat{\psi} = O(|\cdot|^{s_J}), \quad \text{near the origin.}$$

Another way to measure the performance of $X(\Psi)$ is to insist that, in addition to (16), the inverse inequality holds as well:

$$\sum_{\psi \in \Psi} \|T^*_{X(\psi)} f\|_{\ell_2(\alpha)} \ge B_\alpha |f|_{W_2^\alpha(\mathbb{R}^n)}, \quad \forall f \in L_2(\mathbb{R}^n). \qquad (18)$$

For a frame $X(\Psi)$, we define s_B to be

$$\sup\{\alpha > 0 :\ X(\Psi) \text{ satisfies (16) and (18) for the given } \alpha\}.$$

The inequality (18) is the counterpart of the Bernstein-type inequalities in Approximation Theory, and therefore we refer to the above s_B as the *Bernstein-type performance* of $X(\Psi)$. Obviously, $s_B \le s_J$, and usually strict inequality holds. The value of s_B is not connected directly to any easy-to-check property of the system $X(\Psi)$. As a matter of fact, the value of s_B is related to the smoothness of the *dual frame* $X(\Psi^d)$, which we now describe.

[1]Our entire performance analysis in this paper is done for W_2^α, $\alpha > 0$, solely for simplicity. The conditions required for α-performance in the W_2^α sense would imply the same performance in the Besov space $B^\alpha_{p,p}$, [22], sense. We refer to Fig. 1, and, for more details, to [13], [14].

First, one defines a map $\Psi \ni \psi \mapsto \psi^d \in L_2(\mathbb{R}^n)$, and extends it naturally to $X(\Psi)$ (i.e., $(\psi_{j,k})^d := (\psi^d)_{j,k}$). Assume that $X(\Psi^d)$ is also a frame. The frame $X(\Psi^d)$ is then said to be *dual* to $X(\Psi)$ if one has the *perfect reconstruction property*:

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

Here, $T_{X(\Psi^d)}$ is the *synthesis operator* :

$$T_{X(\Psi^d)} : \mathbb{C}^{X(\Psi^d)} \ni a \mapsto \sum_{x \in X(\Psi^d)} a(x) x.$$

Thus, one strives to build wavelet frames that have a high number of vanishing moments, and have smooth dual frames. This brings us to the question of how wavelet systems are constructed. The most general recipe in this regard is known as the Oblique Extension Principle (OEP, [4]). However, in this paper, we will need its special and simpler case, the Unitary Extension Principle (UEP). Both lead to the simultaneous construction of a frame and its dual frame. We describe now the UEP.

The refinability assumption on the function ϕ is equivalent to the condition that

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

for some 2π-periodic function τ, called the *refinement mask*. Let us assume for simplicity that τ is a trigonometric polynomial. We assume that the mother wavelet set $\Psi := \{\psi_1, \ldots, \psi_L\}$ is a subset of $V_1(\phi)$. This amounts to the existence of 2π-periodic functions (=:*wavelet masks*) $(\tau_i)_{i=1}^L$ such that

$$\widehat{\psi}_i(2\cdot) = \tau_i \widehat{\phi}, \quad i = 1, \cdots, L.$$

Again, we assume for simplicity that $(\tau_i)_{i=1}^L$ are trigonometric polynomials. The dual system is constructed similarly, using a dual MRA that is derived from a dual refinable function ϕ^d. Let us assume that the dual refinable function ϕ^d has a trigonometric polynomial refinement mask τ^d. The assumption that $\Psi^d := \{\psi_1^d, \ldots, \psi_L^d\} \subset V_1(\phi^d)$ amounts to the existence of 2π-periodic functions $(\tau_i^d)_{i=1}^L$ such that

$$\widehat{\psi_i^d}(2\cdot) = \tau_i^d \widehat{\phi}^d, \quad i = 1, \cdots, L.$$

Again, we assume that the masks (τ_i^d) are trigonometric polynomials.

Suppose now that the two systems $X(\Psi)$ and $X(\Psi^d)$ are known to be, each, a Bessel system, and they satisfy the Mixed Unitary Extension Principle (MUEP) :

$$\tau(\cdot + \gamma)\overline{\tau^d} + \sum_{i=1}^{L} \tau_i(\cdot + \gamma)\overline{\tau_i^d} = \begin{cases} 1, & \gamma = 0, \\ 0, & \gamma \in \{0, \pi\}^n \backslash 0, \end{cases} \quad (19)$$

and $\widehat{\phi}(0) = \widehat{\phi}^d(0) = 1$. Then $X(\Psi)$ and $X(\Psi^d)$ form a pair of a wavelet frame and a dual wavelet frame [19]. We refer then to the pair $(X(\Psi), X(\Psi^d))$ as a (UEP) bi-framelet.

B. The dual system

Our next goal is to complement the L-CAMP system by a suitable dual 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 L-CAMP system. We are not (yet) claiming that these systems are dual to our L-CAMP systems; that further claim will be established in Lemma 2.

We first define the following partial ordering on $\mathbf{E}$:

$$\nu' \geq \nu \quad \Longleftrightarrow \quad \nu'(l) = \nu(l), \quad l = 1, \ldots, \lceil \nu \rceil,$$

where $\lceil\ \rceil$ is defined as in (7). In particular, $\nu' \geq 0$ for all $\nu' \in \mathbf{E}$. Given a vector $a \in \mathbb{C}^n$, we define

$$a^\nu := \prod_{l=1}^{n} a(l)^{\nu(l)}.$$

Lemma 1: For $\nu \in \mathbf{E}$, let t_ν be the wavelet mask that corresponds to the mother wavelet ψ_ν defined in (4) and (6). Let ξ be any trigonometric polynomial such that $\xi(0) = 1$ and define a new refinement mask

$$\tau^d := \tau_e(2\cdot)\tau_r \left(1 + \xi \left(1 - \frac{\tau_e(2\cdot)}{2^n} \sum_{\nu \in \mathbf{E}} e_\nu \boldsymbol{\tau}^\nu \right)\right), \quad (20)$$

where $\tau_r := \prod_{l=1}^n (1 + \tau_l)/2$, $\boldsymbol{\tau} := (\tau_1, \tau_2, \ldots, \tau_n)$, and $e_\nu(\omega) := e^{i\nu\cdot\omega}$ for $\omega \in \mathbb{T}^n$. We also define dual wavelet masks

$$t_\nu^d = \alpha_n e_{-\nu} \sum_{\nu' \geq \nu} \boldsymbol{\tau}^{\nu' - \nu} (1 - \xi \tau_e(2\cdot)\tau_r e_{\nu'}), \quad \nu \in \mathbf{E}, \quad (21)$$

with $\alpha_n = 2^{-n/2}$ as before. Then the masks $(\tau_c, (t_\nu)_{\nu \in \mathbf{E}})$ and $(\tau^d, (t_\nu^d)_{\nu \in \mathbf{E}})$ satisfy the MUEP condition (19), i.e.

$$\tau_c(\cdot + \gamma)\overline{\tau^d} + \sum_{\nu \in \mathbf{E}} t_\nu(\cdot + \gamma)\overline{t_\nu^d} = \begin{cases} 1, & \text{if } \gamma = 0, \\ 0, & \text{if } \gamma \in \{0, \pi\}^n \backslash 0. \end{cases}$$

Proof: We first note that, for any $a, b \in \mathbb{C}^n$, and $\nu \in \mathbf{E}$,

$$b^\nu - a^\nu$$
$$= \sum_{m=1}^{n} (b(m)^{\nu(m)} - a(m)^{\nu(m)}) \prod_{l=1}^{m-1} b(l)^{\nu(l)} \prod_{l=m+1}^{n} a(l)^{\nu(l)}.$$

This implies that

$$b^\nu - a^\nu = \sum_{\nu' \leq \nu} (b(\lceil \nu' \rceil) - a(\lceil \nu' \rceil))\, a^{\nu - \nu'} b^{\nu'_-},$$

where ν'_- is obtained from ν' by replacing the last 1-digit of ν' (whose position is at $\lceil \nu' \rceil$) by 0. We now fix $\nu \in \mathbf{E}'$ and $\gamma \in \{0, \pi\}^n$, and choose $a := \overline{\boldsymbol{\tau}}(\omega)$ and $b = (e^{-i\gamma(l)})_{l=1}^n$. In that notation

$$t_\nu(\omega + \gamma) = \alpha_n e_{-\nu}(\omega) b^{\nu_-} \left(b(\lceil \nu \rceil) - a(\lceil \nu \rceil)\right),$$

and thus

$$\sum_{\nu' \in \mathbf{E}'} t_{\nu'}(\cdot + \gamma)\overline{t_{\nu'}^d} = \alpha_n^2 \sum_{\nu \in \mathbf{E}'} (1 - \overline{\xi \tau_e(2\cdot)\tau_r e_\nu})(e^{-i\gamma\cdot\nu} - \overline{\boldsymbol{\tau}}^{\,\nu}).$$

From this, we obtain

$$\alpha_n \overline{t_0^d} + \sum_{\nu \in \mathbf{E}'} t_\nu(\cdot + \gamma)\overline{t_\nu^d} = \alpha_n^2 \sum_{\nu \in \mathbf{E}} e^{-i\gamma\cdot\nu}(1 - \overline{\xi \tau_e(2\cdot)\tau_r e_\nu}).$$

Since $t_0 = \alpha_n \left(1 - \overline{\tau_e(2\cdot)\tau_c}\right)$ from (4), then, once we observe that $\tau_c = 2^{-n} \sum_{\nu\in\mathbf{E}} e_{-\nu}$ and

$$t_0^d = \alpha_n \left(2^n \tau_r - \xi \tau_e(2\cdot)\tau_r \sum_{\nu\in E} e_\nu \boldsymbol{\tau}^\nu \right),$$

we get

$$\begin{aligned}
&\tau_c(\cdot+\gamma)\overline{\tau^d} + \sum_{\nu\in\mathbf{E}} t_\nu(\cdot+\gamma)\overline{t_\nu^d} \\
&= \alpha_n^2 \sum_{\nu\in\mathbf{E}} e^{-i\gamma\cdot\nu}(1 - \overline{\xi\tau_e(2\cdot)\tau_r e_\nu}) \\
&\qquad + \tau_c(\cdot+\gamma)(\overline{\tau^d} - \alpha_n \overline{t_0^d \tau_e(2\cdot)}) \\
&= \alpha_n^2 \sum_{\nu\in\mathbf{E}} e^{-i\gamma\cdot\nu}(1 - \overline{\xi\tau_e(2\cdot)\tau_r e_\nu}) + \tau_c(\cdot+\gamma)\overline{\xi\tau_e(2\cdot)\tau_r} \\
&= \alpha_n^2 \sum_{\nu\in\mathbf{E}} e^{-i\gamma\cdot\nu} = \begin{cases} 1, & \text{if } \gamma = 0, \\ 0, & \text{if } \gamma \in \{0,\pi\}^n \backslash 0. \end{cases} \qquad \square
\end{aligned}$$

Discussion. We examine the function ξ that was used in (20) and (21) more closely in order to understand the L-CAMP representation in several different complementary ways. Let ϕ^d be the refinable function associated with τ^d.

The simplest choice for ξ is to let $\xi := 0$. That yields

$$\tau^d = \tau_e(2\cdot)\tau_r, \quad t_\nu^d = \alpha_n e_{-\nu} \prod_{l=\lceil\nu\rceil+1}^{n} (1+\tau_l), \quad \nu \in \mathbf{E}.$$

This interpretation shows that we view the synthesis step as forming linear combination with the dilated shifts of the refinable function ϕ^d whose mask is $\tau_e(2\cdot)\tau_r$. However, our proposed reconstruction algorithm seems to be more efficient than the one offered here. Also, as far as performance analysis goes, the synthesis "wavelets" in this interpretation lack vanishing moments; the lack of this property makes the performance analysis awkward at best.

For performance analysis, we need to choose ξ such that $\xi(0) = 1$. For such ξ, all the masks t_ν^d vanish at the origin, which means that each of the dual wavelets ψ_ν^d has at least one vanishing moment. However, the refinement mask τ^d becomes more involved, and as a result the refinable function ϕ^d becomes less smooth. $\square$

C. The performance of L-CAMP systems

We are finally ready to present our performance analysis of L-CAMP systems. We start with the definition of a smoothness class :

Definition 1: Let $\eta > 0$ be a non-integer, and $\gamma > 0$. We define $\mathcal{R}_\gamma^\eta := \mathcal{R}_\gamma^\eta(\mathbb{R}^n)$ to be the set of all functions f such that, with some constant c,

$$|f^{(\beta)}(t)| \le c\,(1+|t|)^{-\gamma}, \quad \beta \in \mathbb{N}_0^n \text{ and } |\beta| \le \lfloor\eta\rfloor$$

and

$$|f^{(\beta)}(z) - f^{(\beta)}(t)| \le c\,|z-t|^{\eta-\lfloor\eta\rfloor} \sup_{|u|\le|z-t|} (1+|u-t|)^{-\gamma},$$

for $\beta \in \mathbb{N}_0^n$, $|\beta| = \lfloor\eta\rfloor$ and $|z-t| \le 3$.

The set of all the compactly supported functions within $\mathcal{R}_\gamma^\eta$ is denoted by $\mathcal{R}^\eta := \mathcal{R}^\eta(\mathbb{R}^n)$ (and is trivially independent of γ). Finally, for positive integers $N = 1, 2, \cdots$, we let $\mathcal{R}^N := \bigcap_{0<\eta<N} \mathcal{R}^\eta$. $\square$

We note that under the above definition, if f is a compactly supported function with Hölder smoothness α, then $f \in \mathcal{R}^\alpha$.

We approach the performance analysis as follows. We first fix an integer $s \ge 2$. We then require the enhancement mask τ_e to be of order s (i.e. to satisfy (3)), and require the main filter h to have accuracy s as well (i.e., to satisfy (5) for $N := s$). Now suppose that we construct the L-CAMP wavelet system using τ_e and h as explained in Section II. Let Ψ be the L-CAMP mother wavelet set as in (9). Then we see that each of the mother wavelets has s vanishing moments: for ψ_0 this is due to the order of τ_e, and for all the other mother wavelets this is due to the accuracy of the main filter h.

The other important information needed for the performance analysis of $X(\Psi)$ is the smoothness of the dual system. To this end, we will show that it suffices to know the smoothness of the refinable function $\widetilde{\phi}$, the refinable function associated with the mask $\tau_e\tau_r$, (cf. (12) and (20)). We note that the standard performance analysis will hinge on the smoothness of the more complicated ϕ^d associated with τ^d (see, e.g., [17], [15]); thus the reduction of the performance analysis to the smoothness of $\widetilde{\phi}$ is an important step here.

Jackson-type performance follows once we make a minimal smoothness assumption on the function $\widetilde{\phi}$:

Theorem 1: Let $s \ge 2$ be an integer. Assume that we have an L-CAMP system that satisfies (3) for the given s, and (5) for $N := s$. Suppose that $\widetilde{\phi} \in \mathcal{R}^\eta$ for some $\eta > 0$. Then $X(\Psi)$ provides $s_J \ge s$.

The proof of the above theorem invokes the following lemma, which might be of independent interest. It guarantees, for the given L-CAMP wavelet system $X(\Psi)$, that there exists a dual wavelet system $X(\Psi^d)$ associated with ϕ^d, so that the pair $(X(\Psi), X(\Psi^d))$ is a bi-framelet, and so that the smoothness of the dual mother wavelets Ψ^d is as close as one wishes to the smoothness of the above $\widetilde{\phi}$.

Lemma 2: Let $s \ge 2$ be an integer. Assume that we have an L-CAMP system that satisfies (3) for the given s, and (5) for $N := s$. Suppose that $\widetilde{\phi} \in \mathcal{R}^\eta$ for some $\eta > 0$. Then for every $0 < \alpha < \eta$, there exists a wavelet frame $X(\Psi^d)$ associated with a refinable function ϕ^d that corresponds to the mask τ^d in (20), so that the pair $(X(\Psi), X(\Psi^d))$ is a (UEP) bi-framelet and $\Psi^d \subset \mathcal{R}^\alpha$.

From Theorem 1, we see that for Jackson-type performance grade s all we need is that the vanishing moments of the L-CAMP wavelets will be of order s, and that $\widetilde{\phi}$ will be minimally smooth. For the stronger, full-fledged, Bernstein-type performance, the smoothness of $\widetilde{\phi}$ plays a more substantial role:

Theorem 2: Let $s \ge 2$ be an integer. Assume that we have an L-CAMP system that satisfies (3) for the given s, and (5) for $N := s$. Suppose that $\widetilde{\phi} \in \mathcal{R}^\eta$ for some $\eta > 0$. Then $X(\Psi)$ provides $s_J \ge s$ and $s_B \ge \min\{s, \eta\}$.

TABLE III

THE EIGHT L-CAMP SYSTEMS THAT ARE CONSTRUCTED IN THIS PAPER. COLUMN 1 ENUMERATES THOSE SYSTEMS. THE ACCURACY OF THE MAIN FILTER IS LISTED IN COLUMN 2. (SEE TABLE II FOR THE ACTUAL FILTER THAT CORRESPONDS TO THIS ACCURACY.) THE DETAILS OF THE ENHANCEMENT FILTER OCCUPY THE OTHER COLUMNS. WE EXPLAIN IN THE TEXT HOW TO READ THESE DETAILS.

L-CAMP SYSTEMS	accuracy of $h =: N$	order of h_{sm}	$h_G :=$ diagonal part of $\widetilde{h}_e$						h_{A_1}			h_{A_2}			h_{A_3}		
			$-3\mathbf{1}$	$-2\mathbf{1}$	$-\mathbf{1}$	0	$\mathbf{1}$	$2\mathbf{1}$	c_1	$\mathbf{p}_1$	l_1	c_2	$\mathbf{p}_2$	l_2	c_3	$\mathbf{p}_3$	l_3
I	2	0				$\frac{3}{4}$	$\frac{1}{4}$										
II	2	1			$\frac{1}{4}$	$\frac{3}{4}$											
III	4	0				$\frac{9}{16}$	$\frac{9}{16}$	$-\frac{2}{16}$	$\frac{1}{16}$	0	1						
IV	4	1		$-\frac{3}{16}$	$\frac{11}{16}$	$\frac{8}{16}$			$\frac{1}{16}$	$-\mathbf{1}$	-1						
V	4	0			$-\frac{41}{576}$	$\frac{462}{576}$	$\frac{183}{576}$	$-\frac{28}{576}$	$\frac{1}{576}$	$\mathbf{1}$	-3	$\frac{27}{576}$	0	1			
VI	4	0			$-\frac{3}{64}$	$\frac{52}{64}$	$\frac{17}{64}$	$-\frac{2}{64}$	$\frac{2}{64}$	0	-1	$\frac{1}{64}$	0	1	$\frac{1}{64}$	$\mathbf{1}$	1
VII	6	2	$-\frac{157}{2304}$		$\frac{1890}{2304}$	$\frac{976}{2304}$	$-\frac{405}{2304}$		$\frac{4}{2304}$	$-3\mathbf{1}$	3	$\frac{108}{2304}$	0	-1			
VIII	6	2		$-\frac{17}{64}$	$\frac{74}{64}$	$\frac{15}{64}$	$-\frac{8}{64}$		$\frac{2}{64}$	0	-1	$-\frac{1}{64}$	0	1	$\frac{3}{64}$	$-\mathbf{1}$	1

The proofs of the above results (Lemma 2 and Theorems 1,2) are placed in Appendix I.

From Theorem 1 and Theorem 2 (for $\eta := s$), and from the fact that $\widetilde{\phi}$ is compactly supported, we see that the L-CAMP systems satisfying the assumptions in **Performance conditions** (in Section II) have performance grade at least s, in the Jackson sense if we assume (J), and in the Bernstein sense if we assume (B).

V. EXAMPLES OF L-CAMP SYSTEMS

We present eight examples of L-CAMP systems, spanning the range of performance from $s = 2$ to $s = 4$, with performance either in the Jackson sense or in the Bernstein sense. Note that each "system" is actually infinitely many ones, as we cover any possible spatial dimension, and, at least as far as the enhancement filter is concerned, the construction details do depend on the underlying dimension.

We divide the discussion into two. We first present a few tables that contain pertinent information about our eight L-CAMP systems, and explain how to read and interpret the information from those tables. We then sketch the methods we used in order to assess the order of the enhancement filter, and, most importantly, in order to estimate the smoothness of the refinable function $\widetilde{\phi}$.

In Table III, we collect the details that are needed in order to construct the main filters and the enhancement filters of our systems. The system itself is identified in the first column. The accuracy of the main filter is listed in the second column. Using the listed accuracy, one can recover the details of the main filter from Table II. The rest of Table III refers to the enhancement filter, and can be viewed as a set of directions for assembling that filter. Let us explain how those "directions" should be read. Those details are correct whenever the dimension is "large enough", which means $n \geq 2$ for the systems III-IV, and $n \geq 3$ for the systems V-VIII.

The constructed enhancement filters h_e in the table are of the form

$$h_e = h_{sm} * \left(h_G + \sum_{m=1}^{M} h_{A_m} \right) =: h_{sm} * \widetilde{h}_e. \qquad (22)$$

Here, h_{sm} is a smoothing filter, i.e., a filter with Fourier series

$$\widehat{h}_{sm}(\omega) = \left(\frac{1}{2} + \frac{1}{2} e^{-i\mathbf{1}\cdot\omega} \right)^r. \qquad (23)$$

The order r of the smoothing is listed in the 3rd column of Table III. Next, the summand h_G is a filter defined on the diagonal $\mathbb{Z}\mathbf{1}$ of $\mathbb{Z}^n$. The values of h_G can be found in columns 4-9 of Table III. One can see that h_G has small support: it is 2-tap at best, and 4-tap filter at worst.

The other summands, i.e., those that are denoted as h_{A_m}, are obtained by translating, dilating and multiplying by a constant a fixed $(n+1)$-tap filter that we denote as h_A:

$$h_A(k) := \begin{cases} n, & k = 0, \\ -1, & k = \mathbf{e}_l,\ 1 \leq l \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Here, as before, $\mathbf{e}_l$ is the unit vector in the l-coordinate direction. Columns 10-18 of Table III explain how to obtain h_{A_m} from h_A: c_m is a multiplicative constant, $\mathbf{p}_m$ is the translation vector, and l_m is the (unnormalized) dilation parameter. For example, the filter h_{A_1} in system V has parameters $c_1 = 1/576$, $\mathbf{p}_1 = \mathbf{1}$, and $l_1 = -3$, hence it is defined as

$$h_{A_1}(k) := \frac{1}{576} \times \begin{cases} n, & k = \mathbf{1}, \\ -1, & k = \mathbf{1} - 3\mathbf{e}_l,\ 1 \leq l \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Note that the translation vector lies always on the diagonal, hence that all the enhancement filters we construct here are invariant under any permutation of the coordinates. Note also that the number of taps in h_e is primarily determined by the number of summands of the form h_A that appear in its definition, which ranges between 0 and 3 in our constructions. One may also wish to pay attention to the dilation parameters. The L-CAMP system V uses an enhancement filter with fewer taps as compared to system VI. However, one of the dilations in system V definition is -3, making the support of that enhancement filter less local than one might desire. The filter in system VI uses more taps, but its support is more local. Which one may be better for applications we do not know, and it might depend on the details of that application. Hence we listed both options. Similarly, the filter in system VIII uses more taps as compared to system VII, but its support is more local.

TABLE IV

SOME OF THE BASIC PROPERTIES OF THE EIGHT SYSTEMS FROM TABLE III. THE TABLE HERE HIGHLIGHTS THE SMALL CONSTANTS IN THE COMPLEXITY BOUNDS, AND THE EXTREME LOCALNESS. THE PERFORMANCE IS LISTED AS WELL, TOGETHER WITH THE BEST ESTIMATE WE GOT FOR THE SMOOTHNESS OF $\widetilde{\phi}$.

L-CAMP SYSTEMS	λ = $N+1$	tap-size of h_{sm}	tap-size of $\widetilde{h}_e$	tap-size of h_e = vol(supp ψ_0)	tap-size of $\widetilde{h}_e$ + tap-size of h_{sm}	vol(Ψ) $<$	Hölder smoothness of $\widetilde{\phi} \geq$	s_J	s_B
I	3	1	2	2	3	5	1.4150	2	$\geq$ 1.4150
II	3	2	2	3	4	6	2.4150	2	2
III	5	1	n+3	n+3	n+4	n+8	2.3561	3	$\geq$ 2.3561
IV	5	2	n+3	2n+4	n+5	2n+9	3.1926	3	3
V	5	1	2n+4	2n+4	2n+5	2n+9	2.0227	4	$\geq$ 2.0227
VI	5	1	3n+4	3n+4	3n+5	3n+9	2.0342	4	$\geq$ 2.0342
VII	7	3	2n+4	6n+7	2n+7	6n+14	4.3353	4	4
VIII	7	3	3n+4	7n+6	3n+7	7n+13	3.7604	4	$\geq$ 3.7604

In Table IV, we list the properties of each L-CAMP system ($n \geq 2$ for the systems III-IV, $n \geq 3$ for the systems V-VIII). Here, we let $\widetilde{\phi}$ be the refinable function associated with the refinement mask $\tau_e\tau_r$ (cf. (12) and (20)). In column 8 of the table, we list the smoothness of $\widetilde{\phi}$, which plays an important role in the performance analysis (cf. Theorems 1,2). As said, the smoothness estimation techniques are too involved to be covered in detail here; instead we will sketch the main ingredients of the smoothness estimation machinery. The performance grade of each L-CAMP system is listed in the last two columns of the table.

The number $\lambda := 1 +$ (tap-size of h) is shown in column 2 of the table. This number, together with the number of taps of h_e (listed in column 5), is important for the computation of the complexity of the algorithm (cf. (14)) as well as for the total volume of the L-CAMP mother wavelet set (cf. (11)).

For spatial dimension $n = 3, 4, 5$, we compared the L-CAMP systems I,II,V,VII in Table I with some of the mainstream wavelets. There, we used formula (10) for the computation of the volume $\text{vol}(\Psi)$ of the L-CAMP systems.

The details provided in Table III and Table IV exclude $n = 1$ for the systems III-VIII, and $n = 2$ for the systems V-VIII. The missing lower-dimensional counterparts of the listed systems can be obtained from the high-dimensional filters in a trivial way. We present some of these lower-dimensional filters in Appendix II. We note here that the lower bound for the smoothness of $\widetilde{\phi}$ (column 8 in Table IV) for these lower dimensional constructs remains essentially unchanged.

How did we construct the above main filters and enhancement filters? We start by choosing the main filter. The primary property of the main filter is to have accuracy N, which determines the number of vanishing moments that the mother wavelets ψ_ν, $\nu \in \mathbf{E}'$ will have. Note that in constructions VII and VIII the accuracy is $N = 6$, while the last wavelet, ψ_0, has only 4 zero moments, hence the Jackson performance is $s_J = 4$. This is not an oversight: a higher value of N leads also to higher smoothness of the tensor-product function ϕ_r (associated with the mask τ_r) which is a convolution factor in the function $\widetilde{\phi}$ whose smoothness is critical here. So, we choose here $N = 6$ in order to reach the requisite smoothness that is needed for performance $s_B = 4$. In theory, we could have taken in all of our constructions very large values of N, and obtained in this way very smooth $\widetilde{\phi}$. We avoided doing that since such approach creates mother wavelet with small volume of support (since the volume grows only linearly with the accuracy (cf. (11)), regardless of the dimension), but with large diameter for their support. While there is no decisive reason to avoid such constructs, we preferred to keep N, hence λ, as small as we can, and to pay, instead, careful attention to the construction of h_e.

The enhancement filter is constructed to achieve a given order s, (3). The order s dictates the Taylor expansion of τ_e around the origin up to degree $s-1$. Since the filter τ_c is invariant under permutation of the variables, we choose τ_e to have this property, too (this is done by choosing the support of h_e to have this invariance). This reduces significantly the number c of conditions that τ_e needs to satisfy for order s: $c = 2, 4, 7$ conditions, for $s = 2, 3, 4$ respectively. We wanted to put the support of the enhancement filter on the diagonal, since this leads to the smallest possible invariant support. However, it is easy to see that an h_e supported on the diagonal can satisfy only s of the Taylor expansion conditions. To this end, we used, for $s > 2$, the extra summands of the form h_A. The number of summands of this form usually equals to $c - s$. In constructions V and VII we were able to do with fewer summands ($c - s = 3$, while we used only two summands), by carefully selecting the supports of the summands.

The smoothness analysis was done in retrospect, i.e., after the construction was completed. It consists of three main ingredients. The most subtle one was to study the effect of convolving the diagonal part of the enhancement filter with the tensor-product filter h_r (whose mask is τ_r). We developed to this end a decomposition technique of the multivariate filters into the sum of 2-tap ones. Another subtle point was to prove that each of the summands of the form h_A in the definition of h_e has only minor effect on the smoothness of $\widetilde{\phi}$, an effect that is independent of the dimension. The third part was to account for the (positive) effect of the smoothing. While not trivial either, that part was less innovative since it was done by employing some of the tools that were used to the same end in the study of the smoothness of box splines, [5].

Appendix I
Proofs of Lemma 2 and Theorems 1, 2

In this section, we prove Lemma 2 and Theorems 1, 2. The proofs are mostly done by specializing to our context some of the more general results from [10], [11], [18], [13], [14].

For the proof of Lemma 2, we recall the following results:

Result 1 ([14]*):* Suppose that ψ is a finite linear combination of integer translates of $\chi(2\cdot)$, and satisfies $\widehat{\psi}(0)=0$. Then $X(\psi)$ is a Bessel system.

Result 2 ([13]*):* Suppose that ζ is some fixed trigonometric polynomial which has a zero of order 2 at the origin. Let ϕ_0 be some refinable function with a refinement mask τ_0 that satisfies $\phi_0 \in \mathcal{R}^\eta$, for some $\eta>0$. Then for any $0<\eta'<\eta$, there exists a trigonometric polynomial ξ such that $\xi(0)=1$, and such that the refinable function ϕ with mask $\tau_0(1+\xi\zeta)$ belongs to $\mathcal{R}^{\eta'}$.

Result 3 ([18]*):* Suppose that $\psi \in \mathcal{R}^\eta$ for some $\eta>0$, and satisfies $\widehat{\psi}(0)=0$. Then $X(\psi)$ is a Bessel system.

Finally, we make the following simple observation:

Lemma 3: Let τ_A and τ_B be some trigonometric polynomials that satisfy $\tau_A(0)=\tau_B(0)=1$. Let ϕ_1 be the refinable function associated with the mask $\tau_A\tau_B$. Then the refinable function ϕ_0 associated with the mask $\tau_A(2\cdot)\tau_B$ is at least as smooth as ϕ_1.

Proof: We first note that there is unique refinable function corresponding to $\tau_A(2\cdot)\tau_B$. Now from

$$\tau_A(2\cdot)\widehat{\phi}_1(2\cdot)=\tau_A(2\cdot)\tau_A\tau_B\widehat{\phi}_1=\tau_A(2\cdot)\tau_B\left(\tau_A\widehat{\phi}_1\right),$$

and from the uniqueness, we see that $\widehat{\phi}_0=\tau_A\widehat{\phi}_1$, and thus ϕ_0 is at least as smooth as ϕ_1. □

Proof of Lemma 2: We see immediately that the L-CAMP system $X(\Psi)$ is Bessel from Result 1. Now let $0<\alpha<\eta$ be fixed.

We choose a number u such that $\alpha<u<\eta$. Then $\widetilde{\phi}\in\mathcal{R}^u$. From Lemma 3 (for $\phi_1:=\widetilde{\phi}$, $\tau_A:=\tau_e$ and $\tau_B:=\tau_r$), the refinable function ϕ_0 associated with the mask $\tau_e(2\cdot)\tau_r$ satisfies $\phi_0\in\mathcal{R}^u$. From the assumptions (3), (5) and from the identity $\sum_{\nu\in\mathbf{E}} e_\nu \boldsymbol{\tau}^\nu = 2^n\sum_{\gamma\in\{0,\pi\}^n}(\overline{\tau}_c\tau_r)(\cdot+\gamma)$, we see that, near the origin,

$$\begin{aligned}\zeta &:= 1-\tau_e(2\cdot)\frac{1}{2^n}\sum_{\nu\in\mathbf{E}} e_\nu\boldsymbol{\tau}^\nu\\ &= (1-\tau_e(2\cdot)\overline{\tau}_c)+\tau_e(2\cdot)\overline{\tau}_c(1-\tau_r)\\ &\qquad -\tau_e(2\cdot)\sum_{\gamma\in\{0,\pi\}^n\backslash 0}(\overline{\tau}_c\tau_r)(\cdot+\gamma)\\ &= O(|\cdot|^s).\end{aligned}$$

(The first term above has a zero of order s directly from our assumption on the enhancement filter. It is easy to see that the other terms have zero of order N at the origin, with N the accuracy of the main filter. Since we assume here $s\ge N$, we obtain the order above, as stated). Since $s\ge 2$, the function ζ has a zero of order 2 (or more) at the origin. With $\tau_0:=\tau_e(2\cdot)\tau_r$, we use Result 2 to conclude that there exists a suitable ξ for which the refinable function ϕ^d lies in $\mathcal{R}^\alpha$.

Now we argue that the dual wavelet system $X(\Psi^d)$ determined by the above ξ is Bessel. For that, it suffices to show that $X(\psi_\nu^d)$ is Bessel, for each $\nu\in\mathbf{E}$, where $\widehat{\psi}_\nu^d(2\cdot):=t_\nu^d\widehat{\phi}^d$ with t_ν^d as in (21). Note that $\phi^d\in\mathcal{R}^\alpha$ implies $\psi_\nu^d\in\mathcal{R}^\alpha$ since every dual wavelet mask t_ν^d is a trigonometric polynomial. The condition $\widehat{\psi}_\nu^d(0)=0$ is equivalent to $t_\nu^d(0)=0$, which is trivially satisfied from the assumption $\tau(0)=\tau_e(0)=\xi(0)=1$. Since we verified that ψ_ν^d satisfies all the assumptions in Result 3, we see that $X(\psi_\nu^d)$ is Bessel, for each $\nu\in\mathbf{E}$.

Combining the above with Lemma 1, we see that all the requirements for $(X(\Psi),X(\Psi^d))$ to be a UEP bi-framelet are satisfied, [19]. Furthermore, we showed that $\Psi^d\subset\mathcal{R}^\alpha$. □

To prove Theorems 1 and 2, we introduce a sequence space.

For $\alpha>0$, $\ell_2(\alpha)$ is defined to be the space of all sequences $c:=(c(j,k):j\in\mathbb{Z},k\in\mathbb{Z}^n)$ such that

$$\|c\|_{\ell_2(\alpha)}:=\left(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^n}\left(2^{j\alpha}|c(j,k)|\right)^2\right)^{1/2}<\infty.$$

We denote by $\mathcal{S}:=\mathcal{S}(\mathbb{R}^n)$ the Schwartz space of test functions. Also, we use the notation $a\lesssim b$ to mean that there is a constant $c>0$ such that $a\le cb$. We further use the notation $a\approx b$ to denote two quantities that satisfy $c_1a\le b\le c_2a$, for some positive constants c_1 and c_2.

We now recall the following result (see, e.g., [10]).

Result 4: Let $\varphi\in\mathcal{S}$ be such that

$$\begin{aligned}&\operatorname{supp}\widehat{\varphi}\subset\{\frac{1}{2}\le|\omega|\le 2\},\\ &|\widehat{\varphi}(\omega)|\ge c>0,\quad \frac{3}{5}\le|\omega|\le\frac{5}{3},\\ &|\widehat{\varphi}(\omega)|^2+|\widehat{\varphi}(\frac{\omega}{2})|^2=1,\quad 1<|\omega|<2,\end{aligned}\tag{24}$$

for some constant c. Then, we have

$$f=\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^n}\langle f,\varphi_{j,k}\rangle\varphi_{j,k},\quad f\in L_2(\mathbb{R}^n).\tag{25}$$

Let $\alpha>0$. Then, for every $f\in L_2(\mathbb{R}^n)$,

$$|f|_{W_2^\alpha(\mathbb{R}^n)}\approx\|T^*_{X(\varphi)}f\|_{\ell_2(\alpha)}.\tag{26}$$

We also recall two pertinent results from [13]:

Result 5: Let $\alpha>0$. Let $\mathbf{A}$ be a complex-valued matrix whose rows and columns are indexed by $\mathbb{Z}\times\mathbb{Z}^n$:

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

Suppose that there exists a constant $\varepsilon:=\varepsilon(\alpha)>0$ such that, for all j,l,

$$\|\mathbf{A}_{j,l}\|_2:=\|\mathbf{A}_{j,l}\|_{\ell_2(\mathbb{Z}^n)\to\ell_2(\mathbb{Z}^n)}\lesssim 2^{(l-j)\alpha}2^{-|l-j|\varepsilon}.$$

Then $\mathbf{A}$ is a bounded endomorphism of $\ell_2(\alpha)$.

Result 6: Let $j,l\in\mathbb{Z}$, $\alpha>0$. Suppose that there exist constants $\gamma>n$ and $\beta\in\mathbb{R}$ such that for all $k,m\in\mathbb{Z}^n$,

$$|\mathbf{A}_{j,l}(k,m)|\lesssim\frac{2^{(l-j)(\alpha+\frac{n}{2})}2^{-|l-j|\beta}}{2_+^{(l-j)n}}\left(1+\frac{|2^{l-j}k-m|}{2_+^{l-j}}\right)^{-\gamma},$$

where $2_+^a:=\max\{2^a,1\}$. Then we have

$$\|\mathbf{A}_{j,l}\|_2\lesssim 2^{(l-j)\alpha}2^{-|l-j|\beta}.$$

Another relevant result from [14] is also in order. For its statement, we need the following definition:

Definition 2: Let $\lambda \geq 0$ and $\gamma > 0$. Let $\mathcal{M}_\gamma^\lambda := \mathcal{M}_\gamma^\lambda(\mathbb{R}^n)$ be the set of functions f such that

$$|f(t)| \lesssim (1+|t|)^{-\gamma},$$

and, for all $\beta \in \mathbb{N}_0^n$ with $|\beta| \leq \lambda$,

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

Result 7: Let ξ be a linear combination of a finite number of translates of $\chi(2\cdot)$, and let $\eta \in \mathcal{M}_\gamma^0$ for all $\gamma \in \mathbb{N}$. Let

$$\mathbf{A}_{j,l}(k,m) := \langle \xi_{j,k}, \eta_{l,m} \rangle, \quad j,l \in \mathbb{Z},\ k,m \in \mathbb{Z}^n.$$

Then for $l > j$ and $\alpha > -\frac{1}{2}$, there exists $\varepsilon := \varepsilon(\alpha) > 0$ such that

$$\|\mathbf{A}_{j,l}\|_2 \lesssim 2^{(l-j)\alpha} 2^{-|l-j|\varepsilon}.$$

We further need a result from [11].

Result 8: Let $\beta > 0$ be a non-integer. Let $\gamma > n+\beta$. Then, for $j \leq l$,

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

provided $\theta \in \mathcal{R}_\gamma^\beta$ and $\zeta \in \mathcal{M}_\gamma^\lambda$, with λ satisfying $\lfloor \lambda \rfloor + 1 > \beta$.

We are ready to prove the Jackson-type performance result:

Proof of Theorem 1: We first note that, thanks to Lemma 2, the L-CAMP system $X(\Psi)$ is a frame. For each L-CAMP mother wavelet $\psi \in \Psi \subset L_2(\mathbb{R}^n)$, the following identity

$$T_{X(\psi)}^* f = (T_{X(\psi)}^* T_{X(\varphi)}) T_{X(\varphi)}^* f, \quad f \in L_2(\mathbb{R}^n)$$

is valid by (25), where $\varphi \in \mathcal{S}$ is any function satisfying (24).

Let $0 < \alpha < s$ be fixed.

Once we show that $T_{X(\psi)}^* T_{X(\varphi)}$ is bounded on $\ell_2(\alpha)$, then by invoking (26) in Result 4, we obtain, for each $\psi \in \Psi$,

$$\|T_{X(\psi)}^* f\|_{\ell_2(\alpha)} \lesssim \|T_{X(\varphi)}^* f\|_{\ell_2(\alpha)} \approx |f|_{W_2^\alpha(\mathbb{R}^n)}.$$

Thus, it remains to show that $T_{X(\psi)}^* T_{X(\varphi)}$ is bounded on $\ell_2(\alpha)$, for each $\psi \in \Psi$. This is equivalent to proving that

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

with $\mathbf{M}(j,k;l,m) := \langle \psi_{j,k}, \varphi_{l,m} \rangle$, is a bounded endomorphism of $\ell_2(\alpha)$.

When $l > j$, we can apply Result 7 to our matrix $\mathbf{M}$ since it satisfies all the assumptions there. Thus, there exists $\varepsilon_1 > 0$ such that

$$\|\mathbf{M}_{j,l}\|_2 \lesssim 2^{(l-j)\alpha} 2^{-|l-j|\varepsilon_1}. \tag{27}$$

When $j \geq l$, we choose a non-integer u so that $\alpha < u < s$. Using the fact that ψ has at least s vanishing moments and using Result 8 (for $\theta := \varphi$, $\zeta := \psi$, $\beta := u$, $\lambda := s-1$), we get, with $\varepsilon_2 := u - \alpha > 0$,

$$\begin{aligned} |\mathbf{M}_{j,l}(k,m)| &\lesssim 2^{-(j-l)(u+\frac{n}{2})} \left(1 + \frac{|2^{j-l}m - k|}{2^{j-l}}\right)^{-\gamma} \\ &= 2^{(l-j)(\alpha+\frac{n}{2})} 2^{-|l-j|\varepsilon_2} \left(1 + |2^{l-j}k - m|\right)^{-\gamma}. \end{aligned}$$

From Result 6, we obtain

$$\|\mathbf{M}_{j,l}\|_2 \lesssim 2^{(l-j)\alpha} 2^{-|l-j|\varepsilon_2}. \tag{28}$$

Thus, by combining (27) with (28) and by invoking Result 5, we obtain that $\mathbf{M}$ is a bounded endomorphism of $\ell_2(\alpha)$. Therefore for any $\alpha < s$, the L-CAMP frame $X(\Psi)$ satisfies (16). That is, $s_J \geq s$. □

Finally we present the proof of the Bernstein-type performance result.

Proof of Theorem 2: The Jackson-type performance result, $s_J \geq s$, is already proved in Theorem 1. Let $0 < \alpha < \eta$ be fixed.

We let u be a non-integer such that $\alpha < u < \eta$. Then, by Lemma 2, there exists a wavelet system $X(\Psi^d)$ associated with a refinable function ϕ^d (whose mask is the corresponding τ^d from (20)), so that the pair $(X(\Psi), X(\Psi^d))$ is a bi-framelet and $\Psi^d \subset \mathcal{R}^u$. In particular, the L-CAMP wavelet system $X(\Psi)$ satisfies

$$\sum_{\psi \in \Psi} \sum_{l,m} \langle f, \psi_{l,m} \rangle\, \psi_{l,m}^d = f, \quad f \in L_2(\mathbb{R}^n).$$

This implies that

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

where $\varphi \in \mathcal{S}$ is any function satisfying (24). That is,

$$T_{X(\varphi)}^* f = \sum_{\psi \in \Psi} (T_{X(\varphi)}^* T_{X(\psi^d)}) T_{X(\psi)}^* f.$$

Thus, once we show that for each $\psi^d \in \Psi^d$ the operator $T_{X(\varphi)}^* T_{X(\psi^d)}$ is bounded on $\ell_2(\alpha)$, we will obtain that

$$\|T_{X(\varphi)}^* f\|_{\ell_2(\alpha)} \lesssim \sum_{\psi \in \Psi} \|T_{X(\psi)}^* f\|_{\ell_2(\alpha)}, \quad f \in L_2(\mathbb{R}^n).$$

Then by invoking (26) of Result 4, we will reach the stated Bernstein-type performance result.

So, it remains to show that $T_{X(\varphi)}^* T_{X(\psi^d)}$ is bounded on $\ell_2(\alpha)$, for each $\psi^d \in \Psi^d$. This is equivalent to proving that

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

with $\mathbf{N}(j,k;l,m) := \langle \varphi_{j,k}, \psi_{l,m}^d \rangle$, is a bounded endomorphism of $\ell_2(\alpha)$.

When $l > j$, we use the facts that ψ^d has at least one vanishing moment and is of compact support. Result 8 (for $\theta := \varphi$ and $\zeta := \psi^d$) implies then that for any $0 < \beta < 1$, with $\varepsilon_1 := \beta + \alpha > 0$,

$$\begin{aligned} |\mathbf{N}_{j,l}(k,m)| &\lesssim 2^{-(l-j)(\beta+\frac{n}{2})} \left(1 + \frac{|2^{l-j}k - m|}{2^{l-j}}\right)^{-\gamma} \\ &= \frac{2^{(l-j)(\alpha+\frac{n}{2})} 2^{-|l-j|\varepsilon_1}}{2^{(l-j)n}} \left(1 + \frac{|2^{l-j}k - m|}{2^{l-j}}\right)^{-\gamma}. \end{aligned}$$

When $j \geq l$, we use the fact that $\psi^d \in \mathcal{R}^u$. Result 8 (for $\theta := \psi^d$ and $\zeta := \varphi$) implies then that, with $\varepsilon_2 := u - \alpha > 0$,

$$\begin{aligned} |\mathbf{N}_{j,l}(k,m)| &\lesssim 2^{-(j-l)(u+\frac{n}{2})} \left(1 + \frac{|2^{j-l}m - k|}{2^{j-l}}\right)^{-\gamma} \\ &= 2^{(l-j)(\alpha+\frac{n}{2})} 2^{-|l-j|\varepsilon_2} \left(1 + |2^{l-j}k - m|\right)^{-\gamma}. \end{aligned}$$

By invoking Result 5 and Result 6, we obtain that $\mathbf{N}$ is a bounded endomorphism of $\ell_2(\alpha)$. Thus, for any $\alpha < \eta$, the L-CAMP frame $X(\Psi)$ satisfies (18). Therefore, we get $s_B \geq \min\{s, \eta\}$. □

TABLE V
THE L-CAMP SYSTEMS IN 1 DIMENSION FOR $s = 3$ AND $s = 4$.

L-CAMP SYSTEMS	accuracy of h	order of h_{sm}	$\widetilde{h}_e$ ($h_e =: \widetilde{h}_e * h_{sm}$) -3	-2	-1	0	1	2	λ	tap-size of h_e	vol(Ψ)	s_J	s_B
III-a	4	0				$\frac{5}{8}$	$\frac{4}{8}$	-$\frac{1}{8}$	5	3	5.5	3	$\geq$ 2.3561
IV-a	4	1		-$\frac{1}{4}$	$\frac{3}{4}$	$\frac{2}{4}$			5	4	6.5	3	3
VI-a	4	0			-$\frac{5}{64}$	$\frac{55}{64}$	$\frac{17}{64}$	-$\frac{3}{64}$	5	4	6.5	4	$\geq$ 2.0342
VIII-a	6	2		-$\frac{17}{64}$	$\frac{75}{64}$	$\frac{13}{64}$	-$\frac{7}{64}$		7	6	9.5	4	$\geq$ 3.7604

TABLE VI
THE L-CAMP SYSTEMS IN 2 DIMENSIONS FOR $s = 4$.

L-CAMP SYSTEMS	accuracy of h	order of h_{sm}	h_G					h_{A_1}			h_{A_2}			λ	tap-size of h_e	vol(Ψ)	s_J	$s_B \geq$
			-21	**-1**	0	**1**	**21**	c_1	$\mathbf{p}_1$	l_1	c_2	$\mathbf{p}_2$	l_2					
VI-b	4	0		-$\frac{4}{64}$	$\frac{51}{64}$	$\frac{20}{64}$	-$\frac{3}{64}$	$\frac{1}{64}$	0	-1	$\frac{3}{64}$	0	1	5	8	11.75	4	2.0113
VIII-b	6	2	-$\frac{17}{64}$	$\frac{80}{64}$	$\frac{9}{64}$	-$\frac{8}{64}$		$\frac{5}{64}$	0	-1	-$\frac{1}{64}$	0	1	7	14	19.25	4	3.7604

APPENDIX II
L-CAMP SYSTEMS IN LOWER DIMENSIONS

The L-CAMP systems that are detailed in Table III and Table IV of Section V require the spatial dimension to be "minimally high". Needless to say, constructing low dimensional counterparts of those systems is easier than constructing the general ones in that section. We provide here some of the examples of such low-D L-CAMP constructions. In Table V, we list the 1D systems for $s = 3$ and $s = 4$. In Table VI, we list the 2D systems for $s = 4$.

REFERENCES

[1] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, 1983.

[2] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 45, no. 5, pp. 485–560, 1992.

[3] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: Soc. Ind. Appl. Math., 1992.

[4] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames," *Appl. Comput. Harmon. Anal.*, vol. 14, no. 1, pp. 1–46, 2003. [Online]. Available: ftp://ftp.cs.wisc.edu/Approx/dhrs.ps

[5] C. de Boor, K. Höllig, and S. Riemenschneider, *Box Splines*. Berlin: Springer-Verlag, 1993.

[6] G. Deslauriers and S. Dubuc, "Interpolation dyadique," in *Fractals, Dimensions Non Entières et Applications*, Masson, Paris, 1987, pp. 44–55.

[7] R. DeVore, B. Jawerth, and V. Popov, "Compression of wavelet decompositions," *Amer. J. Math*, vol. 114, no. 4, pp. 737–785, 1992.

[8] M. N. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," in *Proc. IEEE Int. Conf. Image Process.*, Thessaloniki, Greece, Oct. 2001.

[9] ——, "Framing pyramids," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2329–2342, 2003.

[10] M. Frazier and B. Jawerth, "Decomposition of besov spaces," *Indiana Univ. Math. J.*, vol. 34, no. 4, pp. 777–799, 1985.

[11] ——, "A discrete transform and decompositions of distribution spaces," *J. Funct. Anal.*, vol. 93, no. 1, pp. 34–170, 1990.

[12] D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis/synthesis," in *Proc. ACM SIGGRAPH*, 1995, pp. 229–238.

[13] Y. Hur and A. Ron, "CAPlets: wavelet representations without wavelets," 2005, preprint. [Online]. Available: ftp://ftp.cs.wisc.edu/Approx/huron.ps

[14] ——, "New constructions of piecewise-constant wavelets," 2005, preprint. [Online]. Available: ftp://ftp.cs.wisc.edu/Approx/pcf.ps

[15] G. Kyriazis, "Decomposition systems for function spaces," *Studia Math.*, vol. 157, no. 2, pp. 133–169, 2003.

[16] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.

[17] Y. Meyer, *Wavelets and operators*. Cambridge: Cambridge University Press., 1992.

[18] Y. Meyer and R. Coifman, *Wavelets: Calderón-Zygmund and multilinear operators*. Cambridge: Cambridge University Press., 1997.

[19] A. Ron and Z. Shen, "Affine systems in $L_2(\mathbb{R}^d)$ II: dual systems," *J. Fourier Anal. Appl.*, vol. 3, pp. 617–637, 1997. [Online]. Available: ftp://ftp.cs.wisc.edu/Approx/dframe.ps

[20] ——, "Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator," *J. Funct. Anal.*, vol. 148, no. 2, pp. 408–447, 1997. [Online]. Available: ftp://ftp.cs.wisc.edu/Approx/affine.ps

[21] S. Toelg and T. Poggio, "Towards an example-based image compression architecture for video-conferencing," 1994, A.I. Memo No. 1494, M.I.T.

[22] H. Triebel, *Theory of Function Spaces*. Birkhäuser., 1983.

[23] M. Unser, "An improved least squares Laplacian pyramid for image compression," *Signal Process.*, vol. 27, pp. 187–203, 1992.

[24] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[25] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.