Local Convergence for Wavelet Expansions*

Susan E. Kelly † Mark A. Kon ‡ Louise A. Raphael §

Abstract

Wavelets provide a new class of orthogonal expansions in $L^2(\mathbb{R}^d)$ with good time/frequency localization and regularity/approximation properties [Da2]. They have been successfully applied to signal processing, numerical analysis, and quantum mechanics [Ru].

We study pointwise convergence properties of wavelet expansions and show that such expansions (and more generally, multiscale expansions) of L^p functions ($1 \leq p \leq \infty$) converge pointwise almost everywhere, and more precisely everywhere on the Lebesgue set of the function being expanded. We show that such convergence is partially insensitive to the order of summation of the expansion. It is shown that unlike Fourier series, a wavelet expansion has a summation kernel which is absolutely bounded by dilations of a radial decreasing L^1 convolution kernel $H(|x-y|)$. This fact provides another proof of L^p convergence. These results hold in all dimensions, and apply to related multiscale expansions, including best approximations using spline functions.

1 Introduction and definitions

The purpose of this paper is to study convergence properties of multiresolution expansions, and in particular wavelet expansions. For $L^p(\mathbb{R}^d)$ functions ($1 \leq p \leq \infty$), we show that such expansions converge pointwise almost everywhere, and more specifically, on the entire Lebesgue set of a function f on $\mathbb{R}^d$. In addition, convergence for wavelet expansions holds under various orders of summation, which might only partly respect the ordering of wavelets by their dilation factors. Pointwise convergence properties are also determined for expansions of L^∞ functions. These results also apply to different forms of multiscale expansions, including expansions in Daubechies, Haar, and other orthogonal wavelets, nonorthogonal wavelet expansions, and best approximations using spline functions. The results given here are multidimensional.

*1991 *Mathematics Subject Classification*: Primary 42C15; Secondary 40A30.

†Research partially supported by the National Science Foundation

‡Research partially supported by the Air Force Office of Scientific Research and the National Science Foundation

§Research partially supported by the Air Force Office of Scientific Research

Our technique is to study bounds on summation kernels of the above expansions. We study the pointwise (as well as the function space) relationship between the action $P_j f$ of multiresolution projections P_j on functions f, and the functions' partial sums in wavelets, yielding pointwise convergence (Theorem 1.3 (*ii*),(*iii*)). We remark that the present results are partially insensitive to order of summation. Specifically, they hold for any order in which the sum over translations of a wavelet can remain incomplete in a collection of scales whose range is bounded (see Theorem 1.3 (*iv*)).

The proofs herein center on bounds for the kernels $P_j(x, y)$ of the partial sum operators, and it is shown that these kernels are bounded by rescalings of L^1 radial convolution kernels. We should add that such bounds for wavelet expansions are nontrivial and arise from cancellations which occur in the sum representations of the partial sum kernels. Naive bounding of the summation kernels by using absolute values in their sum representations fails to yield the needed radial bounds for any class of wavelets. This cancellation must be dealt with only in the examination of wavelet expansions (in which there is no assumption on bounds for the scaling function). This cancellation is implicitly taken into account in the bounding of kernels given in the proofs of Lemma 2.9 and Theorem 1.3 (*iii*). There, bounds on the positive scale part of the kernel are exploited to derive bounds on the negative scale portion, which would otherwise be unobtainable through naive absolute value-type bounds.

Wavelets with local support in time and frequency were defined by A. Grossman and J. Morlet [GM] in 1984 in order to analyze seismic data. However, prototypes of wavelets are found in the work of A. Haar [Ha] and the modified Franklin systems of J.-O. Strömberg [Stro]. In order to identify the underlying structure and to generate interesting examples of orthonormal bases for $L^2(\mathbb{R})$, S. Mallat [Ma] and Y. Meyer [Me1] developed an optimal approach to constructing wavelet bases through multiresolution analysis. P.G. Lemarié and Y. Meyer [LM] constructed wavelets in $\mathcal{S}(\mathbb{R}^d)$, the space of rapidly decreasing infinitely differentiable functions. Strömberg [Strö] developed spline wavelets while looking for unconditional bases for Hardy spaces, and G. Battle [Ba] and P.G. Lemarié [Le1] developed these bases. These spline wavelets have exponential decay, but only C^N smoothness. I. Daubechies [Da1] constructed compactly supported wavelets with C^N smoothness. According to the construction, the support of these wavelets increased with the smoothness; in fact, to have infinite smoothness, wavelets must have infinite support.

Meyer [Me1] was among the first to study convergence results for wavelet expansions. He showed that regular wavelet expansions converge in L^p, $1 \leq p < \infty$, and also in L^∞ for expansions of uniformly continuous functions; thus expansions of continuous functions converge everywhere. The results in [Me1] were based on the assumption of so-called regularity for the basic wavelets, which assumed certain minimal decay properties for wavelets and their derivatives. In addition, Walter [Wa1], [Wa2] established pointwise convergence results for regular wavelet expansions of continuous functions. In contrast, the present results assume only that the wavelets being used be bounded by radial decreasing L^1 functions (modulo a logarithmic factor), or that the corresponding scaling function be bounded by such a function (without such a logarithmic factor). There are no regularity assumptions made in terms of differentiability. Our classes of wavelets include the classes of so-called r-regular wavelets defined by Meyer [Me1]. In this context we prove L^p convergence, and convergence on the Lebesgue set (and hence a.e. convergence) for all L^p functions which have wavelet expansions, and in particular for all L^2 functions.

The pointwise and L^p results in this paper were obtained independently by the first author [Ke1] and the second two authors. Results on Gibbs phenomena obtained by the first author [Ke2] and necessary and sufficient conditions for rates of sup-norm convergence of wavelet expansions [KR] by the second two authors will appear elsewhere.

Following Meyer [Me1], by a multiresolution analysis on $\mathbb{R}^d$ $(d \geq 1)$ we mean a decomposition of the space $L^2(\mathbb{R}^d)$ into an increasing sequence of closed subspaces V_n, i.e.

$$...V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \ldots, \tag{1}$$

with the property that the space V_{j+1} is a "rescaling" of the space V_j. By this we mean that

$$f(2x) \in V_{j+1} \text{ if and only if } f(x) \in V_j \text{ for all } j. \tag{2}$$

It is also assumed that

$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\}; \tag{3}$$

$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}^d), \tag{4}$$

where the overline denotes closure, and that V_0 is closed under integer translations, i.e.,

$$f(x) \in V_0 \quad \Rightarrow \quad f(x-k) \in V_0 \tag{5}$$

for all $k \in \mathbb{Z}^d$. Finally, it is assumed that there exists a function $\phi \in L^2(\mathbb{R}^d)$ such that

$$\{\phi_k(x) \equiv \phi(x-k)\}_{k\in\mathbb{Z}^d} \text{ form an orthonormal basis for } V_0. \tag{6}$$

Such a function ϕ is a *scaling function.* We remark that some of our results use only conditions (1-5) (e.g., in Theorems 1.6 and 2.6), and we will state this when it is the case. Let the space W_j denote the orthogonal complement of V_j in V_{j+1}, i.e., $W_j = V_{j+1} \ominus V_j$. From existence of ϕ it follows (see, e.g., [D2]) that there is a family $\{\psi^\lambda(x)\}_{\lambda\in\Lambda}$ of basic wavelets (whose cardinality depends on the dimension d) such that $\psi^\lambda_{jk}(x) \equiv 2^{jd/2}\psi^\lambda(2^j x - k)$ $(j \in \mathbb{Z},\ k \in \mathbb{Z}^d,\ \lambda \in \Lambda)$ form an orthonormal basis for W_j for fixed j, and form an orthonormal basis for $L^2(\mathbb{R}^d)$ as $j,\ k$ vary. In one dimension the cardinality of Λ is one, so that there is one basic wavelet $\psi(x)$. Further, the class of wavelets $\{\psi^\lambda(x)\}_\lambda$ forms a basis for the space W_0, and in general $\{\psi^\lambda_{jk}(x)\}_{\lambda,k}$ forms an orthonormal basis for W_j.

Our results hold in arbitrary dimension. The most direct construction of multi-dimensional wavelets is through tensor products of one dimensional multiresolution analysis (see, e.g., [Da2],[Me1]). For example, in two dimensions a basis for W_j is of the form

$$\{\psi^h_{jk}, \psi^v_{jk}, \psi^d_{jk}\}_{k\in\mathbb{Z}^2} = \{\psi^\lambda_{jk}\}_{k\in\mathbb{Z}^2, \lambda\in\{h,v,d\}},$$

where

$$\psi^h(x,y) = \phi(x)\psi(y); \quad \psi^v(x,y) = \psi(x)\phi(y); \quad \psi^d(x,y) = \psi(x)\psi(y) \tag{7}$$

and $\psi^\lambda_{jk}(x,y) = 2^j\psi^\lambda(2^jx - k_1, 2^jy - k_2)$ (here $(k_1,k_2) = k$). This orthonormal basis generates a multiresolution analysis, and analogous bases can be constructed in higher dimensions. Thus in general we will write, as a wavelet basis for $L^2(\mathbb{R}^d)$, the collection $\{\psi^\lambda_{jk}\}_{j\in\mathbb{Z},k\in\mathbb{Z}^d,\lambda\in\Lambda}$, with Λ an indexing set containing $2^d - 1$ elements. The following results will hold for any set $\{\psi^\lambda\}_\lambda$ whose translations and dilations form an orthonormal basis for $L^2(\mathbb{R}^d)$, regardless of whether they are constructed with tensor products as above (see [D2], Ch. 10).

For the following, we will require the definition of a Lebesgue point of a function f on $\mathbb{R}^d$. Essentially, it is a point x near which the values of f do not deviate too far on the average from the value $f(x)$, and thus can be considered a generalized continuity point.

Definition 1.1 *The point x is a* Lebesgue point *of the function $f(x)$ on $\mathbb{R}^d$ if f is integrable in some neighborhood of x and*

$$\lim_{\epsilon\to 0}\frac{1}{V(B_\epsilon)}\int_{B_\epsilon} |f(x) - f(x+y)|\, dy = 0,$$

where B_ϵ denotes the ball of radius ϵ about the origin, and V denotes volume.

We remark (see, e.g., [SW]) that this set of points has full measure in $\mathbb{R}^d$, i.e., its complement has measure 0, so that convergence of a series on the Lebesgue set implies almost everywhere convergence. Furthermore, all continuity points are also Lebesgue points. Since the set of continuity points of a function can have measure 0 (as for example in the characteristic function of the rational numbers), the Lebesgue set can clearly in some instances be much larger than the continuity set of a function.

The results we use depend on previous results in Fourier analysis, and we will require a notion which will help us exploit these results:

Definition 1.2 *A function $f(x)$ on $\mathbb{R}^d$ is* radial *if f depends on $|x|$ only. A real valued radial function is* radial decreasing *if $f(x) \leq f(y)$ whenever $|x| \geq |y|$. A function $f(x)$ is in the class $\mathcal{RB}$ if it is absolutely bounded by an L^1 radial decreasing function $\eta(x)$, i.e., with $\eta(x_1) = \eta(x_2)$ whenever $|x_1| = |x_2|$, and with $\eta(x_1) \leq \eta(x_2)$ whenever $|x_1| \geq |x_2|$, and $\eta(x) \in L^1(\mathbb{R}^d)$. We define P_j and Q_j, respectively to be the* *orthogonal projections onto the spaces V_j and W_j, with kernels $P_j(x,y)$ and $Q_j(x,y)$. We define the projection Q_{jk}^{λ} to be the orthogonal projection onto the span of $\psi_{jk}^{\lambda}(x)$ with kernel*

$$Q_{jk}^{\lambda}(x,y) = \psi_{jk}^{\lambda}(x)\overline{\psi_{jk}^{\lambda}}(y).$$

We will say that given $f \in L^2$,
(i) the *multiresolution expansion* of f is defined by the sequence $\{P_n f\}_n$.
(ii) the *wavelet expansion* of f is

$$\sum_{j;k,\lambda} a_{jk}^{\lambda}\psi_{jk}^{\lambda}(x) \sim f, \tag{8}$$

where the a_{jk}^{λ} are the L^2 expansion coefficients of f.
(iii) the *scaling expansion* of f is

$$\sum_k b_k \phi_k(x) + \sum_{j\geq 0;k,\lambda} a_{jk}^{\lambda}\psi_{jk}^{\lambda}(x) \sim f, \tag{9}$$

where the b_k, a_{jk}^{λ} are L^2 expansion coefficients of f.

Note that any function in $\mathcal{RB}$ must be bounded, since it must be bounded by a radial decreasing η function which is defined at the origin. We remark that the L^2 expansion coefficients in (*ii*) and (*iii*) (defined by integration against f) are defined and uniformly bounded for any $f \in L^p$, $1 \leq p \leq \infty$.

The three expansions above converge pointwise and in L^p under various hypotheses, as is seen in Theorem 1.3 below. A priori, they are not everywhere or even almost everywhere equal to each other at corresponding intermediate stages of summation.

When a function $f \in L^2(\mathbb{R}^d)$ is expanded in a wavelet or multiresolution expansion (using the wavelet basis above), it is known that the series converges in L^2. Holschneider and Tchamitchian have proved [HT] that the wavelet integral transform of an $L^2(\mathbb{R}^d)$ function converges pointwise at continuity points. Further results leading toward questions of simultaneous L^p convergence have been dealt with in [Me1] under certain hypotheses on the basic wavelet ψ^{λ} or the scaling function ϕ. Certainly it is not possible for any expansion of a function to converge correctly at all points, since functions can be redefined on sets of measure 0 without changing their expansions. So the question arises, what can be said about the pointwise convergence of wavelet expansions? Do they converge pointwise almost everywhere for $L^2(\mathbb{R}^d)$ functions? At all Lebesgue points? The answer is yes to both of these questions for a general class of multiscale expansions.

Theorem 1.3 *(i) Assume only that the scaling function ϕ of a given multiresolution analysis is in $\mathcal{RB}$, i.e. that it is bounded by an L^1 radial decreasing function. Then for an $f \in L^p(\mathbb{R}^d)$ $(1 \leq p \leq \infty)$, its multiresolution expansion converges to f pointwise almost everywhere.*

(ii) If $\phi, \psi^\lambda \in \mathcal{RB}$ for all λ, then also the scaling (9) (if $1 \le p \le \infty$) and wavelet (8) (if $1 \le p < \infty$) expansions of any $f \in L^p(\mathbb{R}^d)$ converge to f pointwise almost everywhere. If further ϕ and ψ^λ are (partially) continuous, then both of these expansions additionally converge to f on its Lebesgue set.

(iii) If we assume only $\psi^\lambda(x)\ln(2+|x|) \in \mathcal{RB}$ for all λ, then the wavelet (for $1 \le p < \infty$) and multiresolution (for $1 \le p \le \infty$) expansions of any $f \in L^p(\mathbb{R}^d)$ converge to f pointwise almost everywhere; if further the ψ^λ are (partially) continuous, then the wavelet and multiresolution expansions converge to f on its Lebesgue set.

(iv) The last two statements hold for orders of summation where, at any stage, the range of the values of j for which the sum over k and λ is partially complete always remains bounded.

In the last condition, a summation over k and λ with a fixed j is *partially complete* if it contains some terms, but not all with the given value of j. By the *range of values of* j for which the k, λ sum is partially complete we mean the difference of the largest and smallest values of j for which the sum is partially complete. Statement (*iv*) requires that this range always be smaller than some constant M.

We emphasize that statement (*iii*) of Theorem 1.3 makes no assumptions on the scaling function ϕ. If such bounds could be assumed they would make the proof (in section 2) less complicated. A radial bound on the scaling function ϕ is not necessarily guaranteed by the fact that one exists for the basic wavelet ψ. We mention however that it has been proved that in d dimensions certain classes of wavelet bases come from a multiresolution analysis, i.e., are associated to a scaling function ϕ. This has been recently proved by Auscher [Au2] under general conditions on the Fourier transform of the wavelet ψ which do not require compactness or r-regularity for ψ. Lemarié-Riensset [Le2] has also proved this under different assumptions.

We also remark that the pointwise convergence results in Theorem 1.3 also hold for functions $f \in L^\infty(\mathbb{R}^d)$ for the case of multiresolution and scaling expansions. That this fails to hold for wavelet expansions is easily seen by considering the expansion of the function $f(x) = 1$. In this case the wavelet expansion is identically 0, since $\int \psi(x)dx = 1$. It is interesting to point out, however, that for an L^∞ function $f(x)$ whose average value is 0 (in the sense that average values on certain rescaled sets tend to 0 as the sets increase in size), it can be shown that the convergence of the wavelet expansion to $f(x)$ again occurs almost everywhere and on the Lebesgue set of f if ψ is partially continuous, using small modifications of the techniques of the proofs of (*ii*) and (*iii*) of Theorem 1.3.

The following result is a consequence of the proof of Theorem 1.3, and has been proved earlier under somewhat stronger hypotheses, yielding stronger conclusions in [Me1].

Proposition 1.4 *Under the hypotheses of cases (i) to (iii) of Theorem 1.3 L^p convergence of the expansions of L^p functions also follows for $1 \leq p < \infty$. This remains true for any order of summation as in (iv) above.*

Thus, for one and multidimensional multiresolution expansions (including wavelet series), essentially all hoped for convergence properties hold. Questions involving rates of convergence are considered in [KR].

We remark that the proofs that we give of Theorem 1.3 and Proposition 1.4 effectively use maximal function techniques through their dependence on Theorem 2.2, which is a variation on a standard result in harmonic analysis.

The convergence issues arising in multiresolution expansions parallel similar ones which have come up in Fourier series and more general eigenfunction expansions. Pointwise convergence (almost everywhere) of Fourier series for L^2 functions was not established until Carleson's work was published in 1965 [C]. In one dimension, L^p convergence of Fourier series holds for $1 < p < \infty$, but fails in general for $p = 1$ and ∞. Hunt [Hu] proved that for $f \in L^p$, $p > 1$, convergence almost everywhere of Fourier series also holds. Multidimensional results have been harder to come by, and it is known for example that multidimensional Fourier series of L^p functions for certain ranges of p fail to converge almost everywhere (see, e.g., [SW], Corollary VII.4.5, and [Fe]). Using equisummability results [KRY], this implies similar facts for Sturm-Liouville series as perturbations of Fourier series. Fortunately various summability techniques (e.g., Abel or Cèsaro) do guarantee convergence of multidimensional Fourier series to be almost everywhere and in L^p [SW] [GK2]. More general multidimensional eigenfunction expansions have the same properties because of equisummation results (see, e.g., [Ko], [KRY]). It has been shown however [KST] that even Riesz summation does not guarantee convergence of higher dimensional harmonic expansions in all L^p spaces.

Regarding L^p convergence of wavelet expansions, Strichartz [Stri] has also studied L^p properties of the projections P_n on larger function spaces (which include measures). In addition, uniform convergence of spline and therefore associated wavelet expansions is known to occur for continuous functions [deB] [Me1]. Similar results are known for wavelet expansions of functions in Sobolev spaces H^s, given sufficient regularity of the basic wavelet ψ forming the basis, or equivalently of the projection kernel $Q_0(x, y)$ onto the span of the translations $\psi^\lambda(x - k)$ of the basic wavelets ([Me1],[Wa2]).

The basis for Theorem 1.3 is the following bound on the kernel of the projection P_n onto the scaling space V_n. Unlike the case of Fourier series, these wavelet summation kernels $P_n(x, y)$ fall into the class $\mathcal{RB}$ of kernels which are absolutely bounded by dilations of an L^1 radial decreasing convolution kernel $H(|x - y|)$. Previous results on this [GK1], are utilized to show the above-mentioned convergence properties. Under the assumption that the scaling function $\phi \in \mathcal{RB}$, this proposition is easy to prove, though it is more technical and difficult when assumptions only on ψ are made. The following proposition together with fairly well known techniques in Fourier analysis gives the theorem above, and assists in various cases of other theorems. It is related to other properties of projections onto the basic subspaces V_j given, say in [Me1], which has stronger hypotheses on ψ^λ and stronger conclusions.

Proposition 1.5 *(i) Under the assumption that $\phi \in \mathcal{RB}$ or that $\psi^\lambda(x)\ln(2+|x|) \in \mathcal{RB}$ for all λ, the kernels $P_m(x,y)$ of the projections onto V_m satisfy the convolution bound:*

$$|P_m(x,y)| \leq C2^{md}H(2^m|x-y|)\,, \tag{10}$$

where $H \in \mathcal{RB}$, i.e. H is in L^1 and radially decreasing.
(ii) This convolution bound continues to hold if P_m is replaced by the partially complete sum

$$P_m + \sum_{\substack{m\leq j\leq m+M \\ (k,\lambda)\in K_j}} Q^\lambda_{jk}, \tag{11}$$

where the set K_j is for each j an arbitrary collection of k and λ, and M is a fixed constant.

The above convergence statement for multiresolution expansions applies directly to spline expansions as well. Given a grid K in $\mathbb{R}^d$ and a family of splines $\{\phi_{jk} \equiv 2^{jd/2}\phi(2^jx-k)\}_{j\in\mathbb{Z},k\in K}$ of fixed polynomial order which spans $L^2(\mathbb{R}^d)$, let P_j denote the orthogonal projection onto the closed span $sp\{\phi_{jk}\}_{k\in K}$. Given such a family of splines, what is the behavior of L^2 projections of functions onto it? Do they converge pointwise as the mesh becomes small? In L^p? Of interest in approximation theory has been the conjecture [deB] that for arbitrary sequences of meshes in one dimension, the L^2 projections of a continuous function f onto spline spaces (of given order) on these meshes converge in L^∞ norm to f.

We give almost everywhere pointwise convergence results for best L^2 approximations of functions in $L^2(\mathbb{R}^d)$ on a uniform grid.

Theorem 1.6 *Let $f \in L^2(\mathbb{R}^d)$. The best L^2 spline approximations P_jf of f in the space of splines of order ℓ on a uniform grid converge to f almost everywhere and in L^p, for $1 \leq p < \infty$.*

Remark: Though almost everywhere convergence is shown to occur for multiresolution (e.g., wavelet or spline) expansions, such convergence can be arbitrarily slow even for expansions of continuous functions. This is seen from the example of functions in one dimension which have behavior of the form x^ϵ near the origin, for small ϵ, whose convergence can be made arbitrarily slow at the origin.

We remark that since our results are for multiscale expansions in general, the results as applied to wavelet expansions are not sensitive to such issues as orthogonality of wavelets. Essentially any basis ψ_{jk}^{λ} will do as long as this family conforms to a multiscale analysis, i.e., as long as there exists a family of closed subspaces $V_j \subset V_{j+1}$, with $\phi(x) \in V_j \Rightarrow \phi(2x) \in V_{j+1}$, such that $\cup_j V_j$ is dense in $L^2(\mathbb{R}^d)$, and the functions $\{\psi_{jk}^{\lambda}\}_{k,\lambda}$ form a basis for $W_j \equiv V_{j+1} \ominus V_j$. This is because the partial sums of expansions in the wavelets ψ_{jk}^{λ} will have kernels $P_j(x,y)$ given by projections onto the spaces V_j, regardless of the particular wavelet basis. The required conditions on the kernels $P_j(x,y)$ (that they be bounded as in Proposition 1.5) will involve L^1 convolution bounds. These can be tested through the identity $P_j(x,y) = \int P_j(x,y')\delta(y-y')dy'$, where δ denotes the delta distribution. That is, if the best approximation of a highly peaked function (such as $\delta(y-y')$) in wavelets of order $j' \geq j$ decays essentially in an L^1 fashion, then we can expect the corresponding wavelet expansions to converge to functions they are approximating almost everywhere, in $L^p(\mathbb{R}^d)$ $(1 \leq p < \infty)$, and everywhere for functions in $C(\mathbb{R}^d)$. This type of condition can be directly tested (in various ways) for wavelet, spline, and any other multiscale expansions. In particular, the bounds we find below for kernels of scaling space projections P_j of orthonormal wavelet expansions carry over quite easily to more general non-orthogonal multiscale expansions (see [Da2]). For these reasons of generality, we begin this paper with an analysis of convergence properties of general multiscale expansions.

In addition, the scaling of the spaces V_j by factors of 2 is also not crucial; the arguments in this paper hold just as well for scalings by other constant factors, as long as a multiresolution analysis of the function space results (see [Au1]).

Our approach here is to attempt to be as general as possible, since multiscale expansions do not necessarily have to take the form of wavelet expansions. For years much of the work in approximation theory has been based on the notion that multiscale expansions are useful and interesting. The connection of spline expansions to wavelet expansions from the approximation theoretic viewpoint can be found in [BM], [BDR], [Ch] and [CW].

The extension of this result to pointwise convergence of best L^2 approximations by splines is essentially a consequence of the fact that such approximations can be framed in the context of multiresolution expansions, for which in fact there exist orthonormal wavelets with the same convergence properties.

Our approach in this paper is to look at the kernel of the partial sums of wavelet expansions. Namely, if P_m denotes the orthogonal projection onto the scale space V_m (in this case in one dimension), then we will show that under the present general hypotheses, we can write its kernel $P_m(x,y)$ in the two forms

$$P_m(x,y) = \sum_{j<m;k,\lambda} \psi_{jk}^{\lambda}(x)\overline{\psi_{jk}^{\lambda}}(y) = \sum_k \phi_{mk}(x)\overline{\phi_{mk}}(y)\,, \tag{12}$$

with pointwise absolute convergence of both the above sums for fixed j. This kernel of course converges in some sense to the delta function $\delta(x-y)$ as $m \to \infty$. We will show that this occurs in such a way that

$$(13) \qquad |P_m(x,y)| \le 2^{md} H(2^m|x-y|)\,,$$

i.e., that $P_m(x,y)$ is bounded by dilations of a convolution kernel given by a radial decreasing L^1 function $H(|x|)$.

This fact can be easily illustrated in the simple case of one-dimensional Haar wavelets. It is most easily seen from the fact that in this case the scaling function $\phi = \chi_{[0,1]}$ is the characteristic function of the unit interval, and the basic Haar wavelet is given by:

$$\psi(x) = \begin{cases} 1, & \text{if } 0 \le x < 1/2 \\ -1, & \text{if } 1/2 \le x < 1 \end{cases}$$

with $\psi = 0$ elsewhere. In this case the projection onto the basic subspace V_0 can be written in the two forms

$$(14) \qquad P_0(x,y) = \sum_k \phi(x-k)\phi(y-k) = \sum_{j<0;k} \psi_{jk}(x)\psi_{jk}(y)\,,$$

modulo questions of pointwise convergence of the above expansions, which in this simple case are not difficult to verify. Using the first representation in (14) it is easy to show that P_0 is radially bounded as indicated above. First, note that the sum $\sum_k \phi(x-k)\phi(y-k)$ is supported in a diagonal band of width $\sqrt{2}$ in the $x-y$ plane:

INSERT FIGURE

Figure 1

Clearly the function $P_m(x,y)$ (being also uniformly bounded) satisfies the desired radial bound (10) for $m=0$, with, say, $H(x) = \chi_{[-1,1]}(x)$. The fact that (10) holds for all m then follows immediately from the scaling properties of the spaces V_m. Namely, since the V_m satisfy

$$f(x) \in V_m \quad \Leftrightarrow \quad f(2^{-m}x) \in V_0\,,$$

it follows from this scaling property that the projection kernels satisfy the L^2 scaling property

$$P_m(x,y) = 2^m P_0(2^m x, 2^m y)\,,$$

from which (13) for all m follows immediately. For the proof of convergence it is also necessary to prove that the integral $\int P_m(x,y)dy$ converges to 1 a.e.[x], but this is also easy to prove using the first representation in (14).

Since one is sometimes given only properties of the basic wavelet ψ, it is also of interest to derive the bound from the second representation (14) of $P_m(x,y)$. In this case, the bound (say when $m = 0$) of course still holds, but by virtue of a more interesting phenomenon. Namely, if one naively tries to bound the second representation in (14) by bounding the absolute values of the summands, the best bound obtainable is of the form

$$|P_m(x,y)| \leq \sum_{j<m;k} |\psi^\lambda_{jk}(x)\overline{\psi^\lambda_{jk}}(y)| \leq \frac{C}{|x-y|}$$

(the problems with getting a better bound occur in this case at ∞ rather than the origin). The improved L^1 bound only follows from the fact that there is a lot of cancellation going on in this sum, as there is (in a more complicated way) for general sums of the form (12). In this case, the cancellation can be followed more or less explicitly because of the piecewise constant summands, with the result that $P_0(x,y)$ is again zero almost everywhere (or everywhere depending on how the Haar wavelet is defined at its points of discontinuity) outside of the band in Figure 1. Thus the kernel here not only is bounded by an L^1 decreasing convolution kernel at ∞, but is again identically 0 outside of a finite band.

This cancellative phenomenon occurs for all wavelets under minimal hypotheses detailed later. It is interesting that for compactly supported wavelets, the cancellation results in zero values of the kernel outside diagonal bands, (see the above example) at finite stages in the summation (12). Thus, for example, in the case of Daubechies wavelets, this cancellation also occurs, leading to support for the kernel $P_m(x,y)$ as in Figure 1. In this case the cancellation also occurs, leading to support for the kernel $P_m(x,y)$ as in Figure 1. In this case the cancellation is much harder to see explicitly however.

Once the above radial bounds are established, the approach in this paper is to use a variation on a standard result in harmonic analysis (see Theorem 2.2) to obtain a.e. and L^p convergence. It is possible also to obtain a.e. convergence, for example, in the case of Haar expansions, by using the above L^1 bound on the summation kernel P_m to bound it with the Hardy-Littlewood maximal operator.

We also remark that these results imply some interesting observations regarding expansions in discontinuous wavelets such as Haar wavelets. In such expansions, unlike the Fourier case, everywhere pointwise convergence for representations say of continuous functions may depend critically on the definition of the basic wavelet at the points of discontinuity. In one dimension, the Haar wavelet can be defined to be partially continuous, namely either left- or right-continuous. By Theorem 1.3 either of these definitions will result in everywhere convergent expansions of continuous functions. Correspondingly (see Lemma 2.11) the kernels $P_m(x, y)$ in this situation will converge to 0 *everywhere* off the diagonal.

This type of reasoning also extends to higher dimensional expansions (say by tensored Haar wavelets), but with many more interesting possibilities. For example, the condition that the higher dimensional Haar wavelets be partially continuous in several dimensions results in several ways of defining them at their discontinuity points, depending on the choice of the set A of directions along which the basic wavelets are required to be continuous. It is easy to see that defining the value of a Haar wavelet at its points of discontinuity is crucial in determining whether it will accurately represent continuous functions everywhere. According to Theorem 1.3 there are at least two ways of defining a one-dimensional Haar wavelet in order that Haar expansions of continuous functions converge everywhere. The analogous process in two dimensions actually leads to a larger number of possible definitions for Haar wavelets, which nevertheless lead to expansions which converge everywhere for continuous functions.

We finally remark here that basic wavelets ψ which are not partially continuous (see below) can often be redefined on a set of measure 0 so as to have this property. In any case, essentially all wavelets which have been constructed so far (including the discontinuous Haar wavelets) have the property of being redefinable to satisfy the condition of partial continuity, so that, for example, expansions by them of continuous functions can be made to converge everywhere, and the corresponding summation kernels $P_m(x, y)$ converge pointwise everywhere to 0 off the diagonal $D = \{(x, y) \colon x = y\}$.

We add that in any case the kernel $P_m(x, y)$ diverges on the diagonal D as $n \to \infty$, as is standard for summation kernels of orthonormal expansions.

2 Proofs of Theorems in Section 1

In order to maintain generality, we will state the aspects of our results that apply to all multiscale expansions, sometimes including situations where there is no scaling function ϕ whose translates form an orthonormal basis for V_0. We will specialize to the case of standard wavelet multiresolution analyses (which include the existence of a scaling function) at the end. Let P_j denote the orthogonal projection onto V_j, and $Q_j = P_{j+1} - P_j$ the projection onto W_j.

Conventions: The word decreasing is synonymous with non-increasing. For notational convenience, we assume that any index labeled k is in $\mathbb{Z}^d$. We will follow the notation of the multiresolution analysis in the introduction, and use the abbreviation $\psi_{jk}^\lambda \equiv 2^{jd/2}\psi^\lambda(2^j x - k)$; $\phi_k(x) \equiv \phi(x-k)$. Whenever they exist, we will assume that the functions ϕ and ψ^λ are in the class $\mathcal{RB}$ defined in the introduction.

We remark that at any finite stage (say at level j in the scaling) of the summation of a function in wavelets, the number of different values of j in the sum is already infinite. These infinite sums (for fixed j) are shown below to converge absolutely, and hence are fully order independent.

Given a function $f \in L^2(\mathbb{R}^d)$, let

$$f \sim \sum_{j;k,\lambda} a_{jk}^\lambda \psi_{jk}^\lambda(x) \tag{15}$$

be its multiresolution expansion.

Definition 2.1 *We denote by*

$$f_m \equiv \sum_{j<m;k,\lambda} a_{jk}^\lambda \psi_{jk}^\lambda(x) \tag{16}$$

the partial sum of the expansion of the function f.

Note that in definition 2.1, the order independent convergence of the sum over k for fixed j follows from the assumption $\psi \in \mathcal{RB}$.

Our strategy will be to study the representation of $f_m(x)$ given by

$$f_m(x) = \int_{\mathbb{R}^d} P_m(x,y) f(y) dy, \tag{17}$$

where

$$P_m(x,y) = \sum_{j<m;k,\lambda} \psi_{jk}^\lambda(x)\overline{\psi_{jk}^\lambda}(y) = \sum_k \phi_{mk}(x)\overline{\phi_{mk}}(y)\,. \tag{18}$$

The second sum converges absolutely under our assumptions, and the fact that the first sum also does will be proved later (see Lemma 2.9). The integral (17) is also absolutely convergent, which will follow from the fact that $P_m(x,y)$ is bounded by a convolution kernel H as in Proposition 1.5 of the introduction, which will also be proved later.

Note that $P_m(x,y)$ is just the kernel of P_m, the projection onto V_m. We will focus on the kernels $P_m(x,y)$ and $Q_m(x,y)$ of the projections P_m and Q_m, respectively, i.e., the reproducing kernels of V_m and W_m.

Our strategy is to use the following theorem [GK] regarding properties of kernels bounded as in Proposition 1.5. It is a variant of theorems on scaled convolution kernels from Fourier analysis [SW].

Theorem 2.2 (GK) *Assume there is a kernel* $K_n(x,y)$ *with* $\int K_n(x,y)dy \underset{n\to\infty}{\to} C$, *with* C *a constant, and such that* $|K_n(x,y)| \le c_n^d H(c_n|x-y|)$, *where* $H(|x|)$ *is an* L^1*-radially bounded decreasing function and* d *is dimension. Then for any* $f \in L^p(\mathbb{R}^d)$ $(1 \le p \le \infty)$, *we have* $\int K_n(x,y)f(y)dy \underset{n\to\infty}{\to} Cf(x)$ *at any* x *which is a Lebesgue point of* f, *and hence almost everywhere. Further, the convergence to* f *above also occurs in* L^p $(1 \le p < \infty)$.

Our main result here (Theorem 1.3 in the introduction) is that wavelet expansions converge pointwise almost everywhere for all $f \in L^p(\mathbb{R}^d)$. In addition, if the projection kernel P_0 of the basic subspace V_0 (or, say, the basic wavelets ψ^λ or scaling function ϕ) is (are) partially continuous (see Def. 2.3), we will show that corresponding wavelet expansions in addition converge at Lebesgue points of the expanded function. On the other hand, continuity of the reproducing kernel (or ψ^λ or ϕ) is not the most general condition to guarantee convergence at Lebesgue points. We will need the following definition:

Definition 2.3 *The function* ψ *is partially continuous if there exists a set* A *of vectors* $a \in \mathbb{R}^d$ *with positive measure such that* $\lim_{\epsilon\to 0}\psi(x+\epsilon a) = \psi(x)$ *for* $a \in A$. A *is called the set of* continuity directions *of* ψ. *If* $\psi(x,y)$ *is a function of two variables, then* ψ *is partially continuous in* x *if for each fixed* y *it is partially continuous as a function of* x, *with set of continuity directions* A_y *such that* $A = \cap_y A_y$ *has positive measure.*

Remark: In one dimension, any function ψ of bounded variation can be redefined on a set of measure 0 to be partially continuous (e.g., right-continuous). This follows from the fact that ψ can be written as a sum of monotone functions, for which right limits exist, and which can be redefined on sets of measure 0 to be right continuous.

Lemma 2.4 *Let* $\{\psi_n\}_n$ *be a sequence of partially continuous functions, and* A_n *be the set of continuity directions of* ψ_n. *If* $A = \cap_n A_n$ *has positive measure, then a uniformly convergent sum* $\psi = \sum_0^\infty \psi_n$ *is also partially continuous.*

Proof. This follows from the fact that for $y \in A$ the limit $\lim_{\epsilon\to 0}\psi(x+\epsilon y)$ and the uniformly convergent sum commute, which is proved in the same way as the fact that a uniformly convergent sum of continuous functions is continuous. $\square$

Lemma 2.5 *Suppose that $f(x,y)$, defined on $\mathbb{R}^{d+\nu}$, depends on the two variables $x \in \mathbb{R}^d$, $y \in \mathbb{R}^\nu$. Then if $f(x,y)$ is partially continuous in x and if for every x there is a δ such that*

$$\int_{\mathbb{R}^\nu} \max_{|z-x|\le\delta} |f(z,y)|\, dy < \infty\,, \tag{19}$$

then $\int_{\mathbb{R}^\nu} f(x,y)dy$ is also partially continuous.

Proof.

Note that if $a \in A$ (see Definition 2.3) and, (without loss) $|a| \le 1$, then

$$\left|\lim_{\epsilon\to 0}\int f(x+\epsilon a, y) - f(x,y)\, dy\right| \le \lim_{\epsilon\to 0}\int |f(x+\epsilon a,y) - f(x,y)|\, dy \underset{n\to\infty}{\longrightarrow} 0\,, \tag{20}$$

where we have used the Dominated Convergence Theorem, with bounding function $2\cdot \max\limits_{\epsilon\le\delta}|f(x+\epsilon a, y)|$, which is in L^1. □

We emphasize that in the following theorem, in the interest of generality, we have assumed only that we have a multiscale analysis satisfying conditions (1-5), i.e, the existence of a scaling function ϕ is not to be assumed. The proof can be simplified if we assume the existence of a scaling function.

Theorem 2.6 *Suppose only that the kernel $P_0(x,y)$ of the projection operator P_0 satisfies a convolution bound of the form*

$$|P_0(x,y)| \le H(|x-y|)\,, \tag{21}$$

where $H(|x|)$ is an L^1 radial decreasing function (without any assumption on existence of a scaling function). Then

(i) $P_m \underset{n\to\infty}{\longrightarrow} I$ strongly in L^p, for $1 \le p < \infty$. Thus for $f \in L^p$, the approximations $P_m f$ converge to f in L^p.

(ii) For $f \in L^p(\mathbb{R}^d)$ $(1 \le p \le \infty)$,

$$P_m f(x) \underset{n\to\infty}{\longrightarrow} f(x) \qquad \text{pointwise almost everywhere}. \tag{22}$$

(iii) $P_m(x,y)$ can be redefined on a set of Lebesgue measure 0 so that (22) holds for every Lebesgue point x of f (for $1 \le p \le \infty$).

(iv) If furthermore $P_m(x,y)$ is continuous (or more generally partially continuous) in x, then the convergence in (ii) holds for all points x which are Lebesgue points of f.

Proof. We will show that statements *(i)* and *(ii)* follow directly from Theorem 2.2, if we can check the hypothesis that $\int P_m(x,y) \underset{n\to\infty}{\longrightarrow} 1$ for almost every x. Note that by the scaling properties of the spaces V_m, we have $V_{m+1} = SV_m$, where for $f \in L^p(\mathbb{R}^d)$,

$$Sf(x) = 2^{d/2} f(2x) \tag{23}$$

where the normalization $2^{d/2}$ makes S a unitary operator. Thus it follows that $P_m = SP_{m-1}S^{-1}$, so that the kernel

$$P_m(x,y) = 2^d P_{m-1}(2x,2y) = 2^{md}P_0(2^m x, 2^m y). \tag{24}$$

Therefore by Theorem 2.2, (i) and (ii) will follow if we can show

$$\int_{\mathbb{R}^d} P_0(x,y)\,dy = 1 \tag{25}$$

for almost every x, which is proved below. For by (21) and (24) it will follow then that $|P_m(x,y)| \le 2^{md}H(2^m|x-y|)$, and that $\int_{\mathbb{R}^d} P_m(x,y)\,dy = 1$ for almost every x.

Before continuing we note that the kernel P_0 is translation invariant, in that

$$P_0(x+k, y+k) = P_0(x,y) \tag{26}$$

for $x, y \in \mathbb{R}^d$, $k \in \mathbb{Z}^d$. This follows easily from the fact that if T_k denotes translation by $k \in \mathbb{Z}^d$, i.e., $T_k f(x) = f(x-k)$, then

$$\begin{aligned} T_k P_0 T_k^{-1} f &= T_k P_0 f(x+k) \\ &= T_k\{g(x) \in V_0 : \|f(x+k) - g(x)\|_2 = \min\} \\ &= T_k\{g(x+k) \in V_0 : \|f(x+k) - g(x+k)\|_2 = \min\} \\ &= T_k\{g(x+k) \in V_0 : \|f(x) - g(x)\|_2 = \min\} \\ &= \{g(x) \in V_0 : \|f(x) - g(x)\|_2 = \min\} \\ &= P_0 f(x). \end{aligned} \tag{27}$$

Thus $T_k P_0 T_k^{-1} = P_0$ or equivalently, (26) holds.

To show (25), assume it is false. Then there exists a set $E_0 \subset \mathbb{R}^d$ of positive measure such that for $x \in E_0$,

$$\left|\int P_0(x,y)\,dy - 1\right| > \epsilon \tag{28}$$

for some $\epsilon > 0$. Note that the set E_0 is invariant under integer translations, since P_0 is integer translation invariant. Further, defining E_m by

$$\left|\int P_m(x,y)\,dy - 1\right| > \epsilon \tag{29}$$

for $x \in E_m$, we see by the scale invariance of the kernels P_m that modulo sets of measure 0,

$$E_m \equiv \{2^{-m}x : x \in E_0\}, \tag{30}$$

Since the sets E_m are invariant under translations by $2^{-m}k$ for $k \in \mathbb{Z}^d$, and each E_m is a rescaling by 2^{-m} of E_0, and E_0 is an integer translation invariant set of positive (and hence infinite) measure, it follows without much difficulty that for any ball B_r of radius r about the origin,

$$\mu(E_m \cap B_r) \underset{n\to\infty}{\longrightarrow} C \tag{31}$$

for some constant $C > 0$, where μ is Lebesque measure.

Consider the characteristic function $\chi_r(x)$ of B_r. We have

$$\int P_m(x,y)\,dy - \int P_m(x,y)\chi_r(y)\,dy = \int_{|y|>r} P_m(x,y)dy \underset{n\to\infty}{\longrightarrow} 0 \tag{32}$$

for $x \in B_r$, by (21). On the other hand,

$$\int P_m(x,y)\chi_r(y)\,dy \underset{n\to\infty}{\longrightarrow} \chi_r(x) \tag{33}$$

in L^2, and hence in measure. Therefore, for any $\epsilon > 0$, the set of x in B_r for which

$$\left|\int P_m(x,y)\chi_r(y)dy - 1\right| > \epsilon \tag{34}$$

goes to 0 in measure as $m \to \infty$. Therefore, the set E_m of x in B_r for which

$$\left|\int P_m(x,y)\,dy - 1\right| > \epsilon \tag{35}$$

also goes to 0 in measure as $m \to \infty$ by (32). This contradicts (31), showing that indeed, (25) holds. This, together with Theorem 2.2, implies (i) and (ii).

Statement (iii) will hold by Theorem 2.2 if (25) holds for all x. This holds a.e., and clearly by a redefinition of $P_m(x,y)$ on a set of measure 0, it will hold everywhere. Furthermore, this redefinition can be accomplished so that the radial bound (21) still holds. Thus, *(iii)* follows.

If $P_m(x,y)$ is partially continuous in x, its integral in y is also partially continuous. Indeed, to apply Lemma 2.5 we note that

$$\begin{aligned} \int_{\mathbb{R}^\nu} \max_{|z-x|\le\delta} |P_m(z,y)|\,dy &\le \int_{\mathbb{R}^\nu} \max_{|z-x|\le\delta} H(|z-y|)\,dy \\ &= \int_{\mathbb{R}^\nu} \max_{|z-x+y|\le\delta} H(|z|)\,dy \\ &= \int_{\mathbb{R}^\nu} \max_{|z+y|\le\delta} H(|z|)\,dy < \infty. \end{aligned} \tag{36}$$

The last inequality follows without difficulty from the fact that $H \in \mathcal{RB}$. Since the integral in (25) is partially continuous and 1 a.e., it must follow that it is 1 everywhere, as desired, proving *(iv)* for $P_m(x,y)$ partially continuous, by Theorem 2.2. $\square$

We will now go on to prove that under the assumptions $\phi(x) \in \mathcal{RB}$ or $\psi(x)\ln(2+|x|) \in \mathcal{RB}$, the hypotheses of Theorem 2.6 are satisfied.

Lemma 2.7 *On $\mathbb{R}^d$, let $f(x)\ln(2+|x|)$ be absolutely bounded by an integrable radial decreasing function $\eta(|x|)$. Then if F is a closed set not containing 0,*

$$\sum_{j=0}^{\infty}\int_F 2^{jd}\left|f(2^jx)\right|dx<\infty\,. \tag{37}$$

PROOF. We can assume without loss that $F=\mathbb{R}^d\sim B_\rho(0)$, where $\rho>0$, and $B_\rho(0)$ is the open ball of radius ρ centered at 0. We then have (letting Ω denote angular variables and ω_d be the surface volume of the unit sphere in d dimensions)

$$\begin{aligned}
\sum_{j=0}^{\infty}\int_F 2^{jd}\left|f(2^jx)\right|dx &\leq \sum_{j=0}^{\infty}\int d\Omega\int_\rho^\infty dr\, r^{d-1}2^{jd}\eta(2^jr)/\ln(2+2^jr)\\
&= \omega_d\sum_{j=0}^{\infty}2^{jd}\int_\rho^\infty dr\, r^{d-1}\eta(2^jr)/\ln(2+2^jr)\\
&= \omega_d\sum_{j=0}^{\infty}\int_{2^j\rho}^\infty dr\, r^{d-1}\eta(r)/\ln(2+r)\\
&= \omega_d\int_0^\infty dr\sum_{j=0}^{\infty}\chi_{[2^j\rho,\infty]}(r)r^{d-1}\eta(r)/\ln(2+r)\\
&\leq C\int_0^\infty dr\,\ln(2+r)r^{d-1}\eta(r)/\ln(2+r)\\
&< \infty\,;
\end{aligned} \tag{38}$$

note that last integral in the last sum converges by the hypothesis that $\eta(|x|)\in L^1$. The next to last inequality follows from the fact that $\sum_{j=0}^{\infty}\chi_{[2^j\rho,\infty]}(r)$ equals the cardinality of the collection of nonnegative integers j such that $2^j\rho$ is less than or equal to r, which is bounded by $C\ln(2+r)$. □

For any two sets A and B, we define the distance $d(A,B)$ between them as $d(A,B)=\inf_{\substack{x\in A\\ y\in B}}|x-y|$. We define the *diagonal* of $\mathbb{R}^d\times\mathbb{R}^d$ as $D=\{(x,x)\colon x\in\mathbb{R}^d\}$.

Lemma 2.8 *(i) If $\phi\in\mathcal{RB}$, then the kernel $P(x,y)\equiv\sum_{k\in\mathbb{Z}^d}\phi(x-k)\overline{\phi}(y-k)$ satisfies*

$$|P(x,y)|\leq H_0(|x-y|)\,, \tag{39}$$

where H_0 is a bounded radial decreasing L^1 function. Further, the convergence of this sum is uniform on $\mathbb{R}^{2d}$. This sum forms the L^2 kernel of the projection P_0 onto V_0.

(ii) If $\psi^\lambda(x)\in\mathcal{RB}$, then $Q(x,y)\equiv\sum_{k\in\mathbb{Z}^d;\lambda}\psi^\lambda(x-k)\overline{\psi^\lambda}(y-k)$ converges uniformly and absolutely on $\mathbb{R}^{2d}$, and is bounded. Further, if $\psi^\lambda(x)\ln(2+|x|)\in\mathcal{RB}$, then

$$|Q(x,y)|\leq H_1(|x-y|)/\ln(2+|x-y|) \tag{40}$$

where $H_1(|x|)$ is a bounded radial L^1 decreasing function. This sum $Q(x,y)$ is the kernel of the orthogonal projection Q_0 onto W_0.

PROOF. (i) To prove uniform convergence, let us write $|\phi(x)| \leq \eta(|x|)$, where $\eta(|x|)$ is a radial decreasing function with $\eta(|x|) \in L^1(\mathbb{R}^d)$. Then, defining

$$M = \sup_{x\in\mathbb{R}^d} |\phi(x)| \, , \tag{41}$$

we can write

$$\begin{aligned} \sum_{k\in\mathbb{Z}^d} \left|\phi(x-k)\overline{\phi}(y-k)\right| &\leq M \sum_{k\in\mathbb{Z}^d} |\phi(x-k)| \\ &< CM \int \eta(|x|)\, dx < \infty. \end{aligned} \tag{42}$$

The last inequality following from the fact that the sum can easily be estimated by the integral since η is bounded, radial, and decreasing. Thus the sum is clearly uniform in x and y.

To show the bound (39), note first that since $P(x,y)$ is invariant under the translation $(x,y) \to (x+\ell, y+\ell)$ for $\ell \in \mathbb{Z}^d$, we can assume that

$$x \in E_0 \equiv \{x : 0 \leq x_i \leq 1,\ 1 \leq i \leq d\}\,, \tag{43}$$

with E_0 the positive lattice cube with the origin at one vertex. We then have for $x \in E_0$ and $y/4 >$ diam $E_0 \equiv$ diameter of $E_0 = \sqrt{d}$:

$$\begin{aligned} |P(x,y)| &\leq \sum_{k\in\mathbb{Z}^d} \left|\phi(x-k)\overline{\phi}(y-k)\right| \\ &\leq \sum_{|k|\leq\left|\frac{y}{2}\right|} \eta(|x-k|)\eta(|y-k|) + \sum_{|k|>\left|\frac{y}{2}\right|} \eta(|x-k|)\eta(|y-k|) \\ &\leq \sum_{|k|\leq\left|\frac{y}{2}\right|} \eta(|x-k|)\eta\left(|\frac{y}{2}|\right) + \sum_{|k|>\left|\frac{y}{2}\right|} \eta\left(\left|\frac{y}{2}\right| - \mathrm{diam}E_0\right)\eta(|y-k|) \\ &\leq \eta\left(|\frac{y}{2}|\right) \sum_{|k|\leq\left|\frac{y}{2}\right|} \eta(|x-k|) + \sum_{|k|>\left|\frac{y}{2}\right|} \eta\left(\left|\frac{y}{2}\right| - \mathrm{diam}\ E_0\right)\eta(|y-k|) \\ &\leq \eta\left(\left|\frac{y}{2}\right|\right)\cdot C_1 + \eta\left(\left|\frac{y}{2}\right| - \sqrt{d}\right) \sum_{|k|>\left|\frac{y}{2}\right|} \eta(|y-k|) \\ &\leq \eta\left(\left|\frac{y}{2}\right|\right)\cdot C_1 + \eta\left(\left|\frac{y}{2}\right| - \sqrt{d}\right)\cdot C_2 \\ &\leq C_3\eta\left(\left|\frac{y}{4}\right|\right) \\ &\leq C_3\eta\left(\left|\frac{y-x}{5}\right|\right), \\ &\equiv H_0(|x-y|), \end{aligned} \tag{44}$$

where in the third inequality we have used $x \in E_0$. In the last two inequalities we have again used that $x \in E_0$, and hence that $|x| \leq \sqrt{d}$, together with $|y| > 4\sqrt{d}$, which was assumed above.

To show that this infinite sum indeed represents the kernel of the projection P_0, we note simply that the operators defined by finite partial sums of the series for $P(x, y)$ converge strongly to P_0, and using standard arguments it follows that the pointwise uniform limit $P(x, y)$ of these partial sums is the kernel of the operator P_0. This completes the proof of *(i)*.

(ii) The proof of uniform convergence is identical to that in part *(i)*. The proof of the bound (40) follows exactly as the proof of (39), except that $\eta(|x|)$ is replaced by $\eta(|x|)/\ln(2 + |x|)$. □

In the following two lemmas, we separate our summation kernel into two parts, one corresponding to negative and one to positive scales. We first bound the positive scaled part, and from this we bound the negative scaled part.

Lemma 2.9 *If $\psi^\lambda(x) \cdot \ln(2 + |x|) \in \mathcal{RB}$ for each λ, then in any set $F \subset \mathbb{R}^d \times \mathbb{R}^d$ with $F \cap D = \emptyset$, the positive scale kernel $M_m(x, y) \equiv \sum_{0 \leq j < m; k, \lambda} \psi^\lambda_{jk}(x)\overline{\psi^\lambda_{jk}}(y)$ satisfies $M_m(x, y) \underset{n\to\infty}{\rightarrow} M(x, y)$, where $M(x, y)$ is bounded by*

$$|M(x, y)| \leq H_2(|x - y|)\,, \tag{45}$$

with $H_2(|x|)$ a radial decreasing L^1 function (possibly infinite at 0). Further, the convergence of M_m to M is uniform and absolute in F if F has a positive distance from the diagonal $D = \{(x, y) : x = y\}$. Furthermore, the absolute sum satisfies

$$\sum_{0 \leq j < m; k, \lambda} \left|\psi^\lambda_{jk}(x)\overline{\psi^\lambda_{jk}}(y)\right| \leq H_3(|x - y|) \tag{46}$$

for some other function H_3 with the same properties as H_2.

Proof. We have

$$M_m(x, y) = \sum_{0 \leq j < m} 2^{jd} Q_0(2^j x, 2^j y). \tag{47}$$

Thus, by Lemma 2.8,

$$|M_m(x, y)| \leq \sum_{0 \leq j < \infty} 2^{jd} H_1(2^j |x - y|)/\ln(2 + (2^j\,|x - y|)) \equiv H_2(|x - y|)\,. \tag{48}$$

We see by Lemma 2.7 that the restriction of $H_2(|x|)$ to F is in L^1, if F has a positive distance from the origin. The fact that this function is radial and decreasing is clear from the same property for the scaled function $H_1(|x|)/\ln(2 + |x|)$ in the above sum. Uniformity of convergence of M_m in F will follow if we can show that convergence of the sum in (48) is uniform. However, this follows from the fact that the functions $2^{jd}H(2^j|x|)$ are decreasing in x, and that therefore if convergence occurs at any point $z = x - y$, it occurs uniformly at all points z_1 with $|z_1| \geq |z|$. Clearly, however the sum converges for arbitrarily positive $|z|$ (by Lemma 2.7 and since the terms are monotonic in z). Therefore it converges uniformly for $|x - y| = |z| \geq \epsilon$ for any positive ϵ, completing the proof of the first statement. The proof of the second statement follows from the fact that all of the above inequalities hold when absolute values are inserted around the terms $\psi_{jk}(x)\overline{\psi_{jk}^{\lambda}}(y)$ in the sums. □

Let us define $Q_j(x, y)$ analogously to $P_j(x, y)$ by

$$Q_j(x, y) \equiv 2^{jd}Q_0(2^jx, 2^jy)$$

Thus Q_j is the projection onto the wavelet subspace W_j. Below, we show that the kernel of P_0 has the expected form in terms of the wavelets ψ_{jk}^{λ} (note that this kernel $P(x, y)$ has already been expressed in terms of the scaling function ϕ in Lemma 2.8.

Lemma 2.10 *Under only the assumption $\psi^{\lambda}(x) \in \mathcal{RB}$ the negative scale part of the kernel $N(x, y) = \sum_{j<0;k,\lambda} \psi_{jk}^{\lambda}(x)\overline{\psi_{jk}^{\lambda}}(y)$ converges uniformly and absolutely on $\mathbb{R}^d \times \mathbb{R}^d$ and is a bounded kernel. $N(x, y)$ is the kernel of P_0, the projection onto V_0.*

Proof. By an argument exactly as in (44), we have

$$\sum_{k,\lambda} |\psi_{jk}^{\lambda}(x)\overline{\psi_{jk}^{\lambda}}(y)| \leq H_1(|x - y|)$$

for some (bounded) radial decreasing L^1 function $H_1(|\cdot|)$. Therefore,

$$\sum_{j<0;k,\lambda} |\psi_{jk}^{\lambda}(x)\overline{\psi_{jk}^{\lambda}}(y)| \leq \sum_{j<0} 2^{jd}H_0(2^j|x - y|) \leq C\sum_{j<0} 2^{jd}, \tag{49}$$

where $C = \sup_x H(|x|)$, which implies that the sum defining $N(x, y)$ converges absolutely and uniformly in x and y, by the Weierstrass M-test.

To show that $N(x, y)$ is the kernel of P_0, note only that the sum $\sum_{j<0;k,\lambda} \psi_{jk}^{\lambda}(x)\overline{\psi_{jk}^{\lambda}}(y) = \sum_{j<0} Q_j(x, y)$ converges uniformly to its limit $N(x, y)$, and that the operators represented by the partial sums converge in the strong operator topology of $L^2(\mathbb{R}^d)$ to P_0, so that by standard arguments it follows that $N(x, y)$ is equal (almost everywhere) to the kernel $P_0(x, y)$ of P_0, as desired. □

Lemma 2.11 *Under only the assumption $\psi^{\lambda}(x)\ln(2 + |x|) \in \mathcal{RB}$ for all λ, the sum*

$$P_{\infty}(x, y) \equiv \sum_{-\infty<j<\infty} Q_j(x, y) \tag{50}$$

converges uniformly and absolutely on any set F with a positive distance from the diagonal D, and is equal to 0 almost everywhere in F. If the kernel Q_0 is continuous (or more generally partially continuous) in x, then $P_{\infty}(x, y) = 0$ everywhere in F.

PROOF.

We write

$$P_\infty(x,y) = \sum_{j<0} Q_j(x,y) + \sum_{j\geq 0} Q_j(x,y). \tag{51}$$

Uniform (and in fact absolute and order independent) convergence of both sums away from the diagonal follows from Lemmas 2.9 and 2.10. Define

$$M_m(x,y) = \sum_{0\leq j<m} Q_j(x,y) \tag{52}$$

and

$$N_{-m}(x,y) = \sum_{-m\leq j<0} Q_j(x,y)\,. \tag{53}$$

Suppose now that it is false that the sum off the diagonal is 0 a.e. Then without loss we may assume there is a set A of positive measure in $\sim D$ on which $P_\infty(x,y) > \epsilon > 0$. Let $B_1, B_2 \subset \mathbb{R}^d$ be closed balls with $B_1 \cap B_2 = \emptyset$, and $\lambda((B_1 \times B_2) \cap A)) > 0$, where λ is Lebesgue measure in $\mathbb{R}^d \times \mathbb{R}^d$. Let $f_1(x)$ and $f_2(x)$ be bounded nonnegative functions supported in B_1 and B_2, respectively. Then letting $P_m(x,y) \equiv M_m(x,y) + N(x,y)$, and denoting by P_m the operator with kernel $P_m(x,y)$, we have

$$P_m f_2 \underset{n\to\infty}{\longrightarrow} f_2,$$

with the convergence in L^2, so that

$$\begin{aligned}
\int f_1(x)P_\infty(x,y)f_2(y)dy\,dx &= \lim_{m\to\infty} \int f_1(x)P_m(x,y)f_2(y)dy\,dx \\
&= \lim_{m\to\infty} \int f_1(x)(P_m f_2)(x)dx \\
&= \int f_1(x)f_2(x)dx = 0.
\end{aligned}$$

The first equality follows from the uniform convergence of the integrand on its right side (on its support, i.e., on $B_1 \times B_2$). This proves that $P_\infty(x,y) = 0$ a.e. on $B_1 \times B_2$, contradicting the assumption that $P_\infty(x,y)$ is not 0 a.e. on A. Thus $P_\infty(x,y) = 0$ a.e. off D.

If $Q_0(x,y)$ is partially continuous in x with a set A of continuity directions (for all y) which has positive measure, then the sum defining $P_\infty(x,y)$ is a uniformly convergent sum of partially continuous functions (in x) with a common set A of continuity directions, and hence is itself partially continuous off of D, by Lemma 2.4. Since the sum is 0 a.e., it follows that the sum in fact vanishes everywhere off D. □

We are now ready to prove Theorem 1.3 of the introduction.

PROOF OF THEOREM 1.3 AND PROPOSITION 1.4: We remark that we will assume in the proofs of statements (*ii*) and (*iii*) that all summations are carried out in such a way that for a given scale j, all terms at scale j are added before any terms at the next scale $j+1$ are added. That is, strictly speaking these proofs hold for orders of summation in which all terms at each scale j are summed simultaneously, and the values of j increase by 1 at each successive stage of the summation. This assumption is then relaxed into the form in statement (iv). We now give the proofs of the statements of the Theorem.

Proof of (i):

Here we assume only that the scaling function $\phi(x) \in \mathcal{RB}$. Note that the function $P(x,y) = \sum_k \phi(x-k)\overline{\phi}(y-k)$ is the integral kernel of the projection P_0 onto V_0 (Lemma 2.8). From (39), (25), and Theorem 2.6, we conclude that for $f \in L^p\,(1 \le p \le \infty)$, $P_m f \underset{n\to\infty}{\longrightarrow} f$ almost everywhere. Note in particular that for $f \in L^p$, the kernel $P_m(x,y)$, being radially bounded by an L^1 convolution kernel, can be applied to f via

$$P_m f = \int_{\mathbb{R}^d} P_m(x,y) f(y) dy.$$

By Theorem 2.6 and the properties of P_m, it follows that P_m is a bounded operator and that in addition $P_m f \underset{n\to\infty}{\longrightarrow} f$ in L^p, proving (i) of Proposition 1.4

Proof of (ii) for scaling expansions:

We first consider the claim about the scaling expansion

$$f_m \sim \sum_k b_k \phi(x-k) + \sum_{0\le j<m;k,\lambda} a^\lambda_{jk}\psi^\lambda_{jk}(x)\,. \tag{54}$$

Note first that if $\phi, \psi^\lambda \in \mathcal{RB}$, then the coefficients $b_k = \int \phi(x-k)f(x)dx$ and $a^\lambda_{jk} = \int \psi^\lambda_{jk}(x)f(x)dx$ are uniformly bounded by the Hölder inequality and the fact that $\phi \in L^1 \cap L^\infty$ and $f \in L^p$. We claim that (54) converges absolutely and uniformly for any $f \in L^p$. Indeed the first sum is absolutely bounded by a constant times $\int \phi(y)dy$ (since the coefficients b_k are uniformly bounded), and the second sum, which is summed over a finite set of j, is absolutely convergent by the same type of argument which bounds the first sum, and the fact that the a^λ_{jk} are uniformly bounded.

Note that since P_m has a kernel $P_m(x,y)$ which is bounded by an L^1 convolution kernel $H(|x-y|)$ in $\mathcal{RB}$ (by part (i) of this theorem), P_m is a bounded operator in all L^p spaces, $1 \le p \le \infty$. It is also easy to show in this case that the representation

$$P_m(x,y) = \sum_k \phi(x-k)\overline{\phi}(y-k) + \sum_{0\le j<m;k,\lambda} \psi^\lambda_{jk}(x)\overline{\psi^\lambda_{jk}}(y) \tag{55}$$

also converges absolutely (see the argument of Lemma 2.8).

Further, we claim that for $1 \le p \le \infty$, $P_m(x,y)$ defines a linear projection P_m (see proof of last statement in Lemma 2.8 (i)). For $1 \le p < \infty$, this can be seen by the fact that this is the case in L^2 and that $L^2 \cap L^p$ is dense in L^p. For $p = \infty$, this follows from the fact that it is true in L^1, together with a duality argument.

It is also not difficult to show by using the form (55) of $P_m(x,y)$, together with the uniform absolute convergence of (55), and the dominated convergence theorem that $f_m = P_m f$ almost everywhere. Thus by (i) of Theorem 1.3, the partial sums $f_m = P_m f$ converge to f a.e. $(1 \le p \le \infty)$ and in $L^p(1 \le p < \infty)$, proving the first assertion of (ii) for the scaling expansion, and the corresponding assertion for L^p convergence in Proposition 1.4.

If ψ^λ and ϕ are partially continuous, then by Theorem 2.6 (iv) and the uniform convergence of the sum defining $P_m(x,y)$, we conclude that $P_m(x,y)$ is partially continuous in x, so that $\int P_m(x,y)f(y)dy$ converges to f on its Lebesgue set.

In addition, $\int P_m(x,y)f(y)dy$ is a partially continuous function in x by Lemma 2.5, using the fact that $P_m(x,y)$ is bounded by $2^{md}H(2^m|x-y|)$, so that, e.g., for $m=0$, we have

$$\max_{|z-x|\le\delta} |P_0(z,y)| \le \max_{|z-x|\le\delta} H(|z-y|) = H^*(|x-y|),$$

with $H(|\cdot|)$ and hence also $H^*(|\cdot|)$ contained in $\mathcal{RB} \cap L^\infty(\mathbb{R}^d)$. Further, in this case f_m is also partially continuous in the same set of directions by Lemma 2.4 and the fact that the sum defining f_m converges uniformly, so that $f_m = \int P_m(x,y)f(y)\,dy$ everywhere, and therefore the partial sums f_m also converge to f on its Lebesgue set, as claimed. This proves the claims in (ii) regarding the scaling expansion.

Proof of (ii) for wavelet expansions: To show that under the hypothesis of (ii) the wavelet expansion converges almost everywhere for $f \in L^p(\mathbb{R}^d)$ $(1 \le p < \infty)$, note first that using Lemma 2.10 to prove uniform and absolute convergence of the sum, we can conclude (again using the arguments in the proof of (i)) that

$$P_0(x,y) = \sum_{j<0;k,\lambda} \psi^\lambda_{jk}(x)\overline{\psi^\lambda_{jk}}(y) = \sum_{j<0} Q_j(x,y) \tag{56}$$

is the kernel of a linear projection P_0 in L^p $(1 \le p \le \infty)$. The fact that certain statements in the rest of the proof of (ii) also hold for $p = \infty$ will be pointed out here, but will not be needed for the remainder of the proof of (ii), for which we now only assume that $1 \le p < \infty$.

For $f \in L^p(\mathbb{R}^d)$, the partial sums of (8), given by

$$f_m = \sum_{j<m;k,\lambda} a^\lambda_{jk}\psi^\lambda_{jk}(x), \tag{57}$$

can be shown for $1 \le p \le \infty$ to converge uniformly and order independently for fixed finite m. Indeed, note that the sequence $\{a^\lambda_{jk}\}$ is uniformly bounded, using an argument like in Lemma 2.8 (*ii*), we have (setting $m = 0$ for convenience)

$$\begin{aligned} \sup_x \sum_{j<0;k,\lambda} |a_{jk}^\lambda \psi_{jk}^\lambda(x)| &< C \sup_x \sum_{j<0;k,\lambda} |\psi_{jk}^\lambda(x)| \\ &= C \sup_x \sum_{j<0;k,\lambda} 2^{jd/2} |\psi_{0k}^\lambda(2^j x)| \\ &\leq \; C \sum_{j<0} 2^{jd/2} \sum_\lambda \sup_x \sum_k |\psi^\lambda(x-k)| \\ &\leq \; C' \sum_{j<0} 2^{jd/2} \sum_\lambda \int |\psi^\lambda(y)| dy \\ &< \; \infty. \end{aligned} \tag{58}$$

Since we have absolute convergence, the partial sum (57) is automatically order-independent with respect to all indices of summation.

Next we wish to show that (for $1 \leq p < \infty$)

$$f_m = P_m f \tag{59}$$

almost everywhere, where f_m is now defined as in (57), and P_0 is the L^p projection with kernel $P_0(x, y)$ defined above. The difficulty in this is that we are no longer working in L^2, but with an L^p function f. Thus we first note that

$$P_0(x, y) = \sum_{j<0} Q_j(x, y),$$

where Q_j is the orthogonal projection onto W_j. The absolute convergence of this sum follows from the absolute convergence of (56), together with the representation

$$Q_j(x, y) = \sum_{k,\lambda} \psi_{jk}^\lambda(x) \overline{\psi_{jk}^\lambda}(y). \tag{60}$$

To show that $f_m = P_m f$, it suffices to show that $f_0 = P_0 f$. To do the latter, we write

$$f_0(x) - P_0 f(x) = \lim_{M \to -\infty} \sum_{-M<j<0;k,\lambda} a_{jk}^\lambda \psi_{jk}(x) - \int \lim_{M \to -\infty} \sum_{-M<j<0} Q_j(x, y)\, f(y) dy.$$

It is not difficult to show that for a finite M,

$$\sum_{k,\lambda} a_{jk}^\lambda \psi_{jk}(x) = \int Q_j(x, y) f(y) dy \tag{61}$$

for $f \in L^p$. Indeed, for fixed x we can represent $Q_j(x, y)$ in the form (60), and then use the dominated convergence theorem, with dominating function

$$\sum_{k,\lambda} |\psi_{jk}^\lambda(x) \overline{\psi_{jk}^\lambda}(y)| f(y)| \leq H(|x-y|)|f(y)|$$

where $H(|x|)$ is a radially bounded L^1 function (which is also in L^∞). Clearly, this function is in L^1 in the y variable, since $f \in L^p$ and $H(|\cdot|) \in L^{p'}$ for all $1 \leq p < \infty$, where p' denotes the dual Hölder exponent to p (with $1/p + 1/p' = 1$). The identity (61) can clearly be extended to finite sums over j, so that we have

$$\sum_{-M<j<0;k,\lambda} a_{jk}^\lambda \psi_{jk}(x) = \int \sum_{-M<j<0} Q_j(x,y)\, f(y)dy$$

(in fact, this also holds for functions f in L^∞). Thus

$$\int \sum_{j<0} Q_j(x,y)\, f(y)dy = \int \sum_{j<-M} Q_j(x,y)\, f(y)dy + \int \sum_{-M\leq j<0} Q_j(x,y)\, f(y)dy$$

$$= \int \sum_{j<-M} Q_j(x,y)\, f(y)dy + \sum_{-M\leq j<0;k,\lambda} a_{jk}^\lambda \psi_{jk}(x). \tag{62}$$

Letting $M \to \infty$ in (62), we have first:

$$\|\int \sum_{j<-M} Q_j(x,y)\, f(y)dy\|_p = \|\int P_{-M}(x,y)\, f(y)dy\|_p, \tag{63}$$

where $P_M(x,y)$ is the kernel of P_M as defined above. Now note that since

$$|P_M(x,y)| \leq 2^{Md} H(2^M|x-y|)$$

for a L^1 radial decreasing function $H(|\cdot|)$, we have (again letting p' denote the dual Hölder exponent)

$$\|\int P_M(x,y)\, f(y)dy\|_p \leq \|\int 2^{Md} H(2^M|x-y|)|f(y)|dy\|_p$$

$$\leq \|2^{Md} H(2^M|x|)\|_{p'} \|f(x)\|_p.$$

Note however that for any $p' > 1$,

$$\|2^{Md} H(2^M|x|)\|_{p'} \underset{M\to\infty}{\longrightarrow} 0 \tag{64}$$

by standard scaling properties of functions. Thus we conclude

$$\|\int P_{-M}(x,y)\, f(y)dy\|_p \underset{M\to\infty}{\longrightarrow} 0.$$

Hence letting $M \to \infty$ and taking L^p limits on the right side of (62), using (63), we conclude that

$$\int \sum_{j<0} Q_j(x,y)\, f(y)dy = \sum_{j<0;k,\lambda} a_{jk}^\lambda \psi_{jk}(x), \tag{65}$$

where the right hand sum is interpreted as an L^p ($1 \leq p < \infty$) limit as $j \to -\infty$. However, clearly the L^p limit of the right hand side is the same function as its pointwise limit (which exists, as we have established above). Therefore (65) can be interpreted as a pointwise equality as well. We thus conclude from (65) that for $f \in L^p(\mathbb{R}^d)$,

$$P_0 f \equiv \int P_0(x,y) f(y) dy = \sum_{j<0;k,\lambda} a_{jk}^{\lambda} \psi_{jk}(x) = f_0, \tag{66}$$

where $a_{jk}^{\lambda} = \int \psi_{jk}(y)\, f(y) dy$. By scaling it follows that (59) holds, as desired.

Now it is easy to show that the wavelet expansion $f_m \underset{n\to\infty}{\longrightarrow} f$ a.e., since by (59) it is only required that we show that $P_m f \underset{n\to\infty}{\longrightarrow} f$ a.e., and the latter has already been shown above in the proof of the first statement in (*ii*) (note that the assumptions used in the proof of the first statement in (*ii*) are the same as the present ones).

If ψ and ϕ are partially continuous, then it follows that P_m is partially continuous, since the sum (56) defining P_m converges uniformly (see Lemma 2.10). Therefore by Theorem 2.6, $P_m f$ converges to f on its Lebesgue set. Further, by the same argument as in the proof of the first part of (*ii*), $\int P_m(x,y)\, f(y) dy$ is partially continuous in x (by Lemma 2.5), as is the partial sum f_m (by Lemma 2.4). Since these two coincide a.e., they coincide everywhere, and so we conclude that $f_m \underset{n\to\infty}{\longrightarrow} f$ on the Lebesgue set of f, as desired, for $f \in L^p$.

Proof of (iii):

We remark here that for wavelet expansions, the absolute value type bounds which have implicitly been used in the proofs of the above results do not work here (note we are not assuming anything about bounds for the scaling function). In the formation of the summation kernel $P_j(x,y)$, cancellations must now be taken into account. This is implicitly accomplished below by the invocation of Lemma 2.9, together with the observation (via Lemma 2.11) that the negative scale part $N(x,y)$ of the summation kernel is just the negative of its positive scale part $M(x,y)$, off the diagonal $x = y$.

Assume $\psi^{\lambda}(x) \ln(2 + |x|) \in \mathcal{RB}$ for all λ. Then as in (*ii*), the kernel

$$P_m(x,y) = \sum_{j<m;k,\lambda} \psi_{jk}^{\lambda}(x) \overline{\psi_{jk}^{\lambda}}(y)$$

converges absolutely and uniformly for m finite, as does the partial expansion

$$f_m = \sum_{j<m;k,\lambda} a_{jk}^{\lambda} \psi_{jk}^{\lambda}(x) \tag{67}$$

of a function $f \in L^p$, using the fact that for all λ, $\psi^{\lambda} \in \mathcal{RB}$ as well. Thus it is not difficult to show exactly as in the proof of (59) above that for $f \in L^p(\mathbb{R}^d)$ $(1 \le p < \infty)$,

$$f_m = \int P_m(x,y)\, f(y) dy \tag{68}$$

almost everywhere, with $P_m(x,y)$ as above. Note that according to Lemma 2.9 and the scaling properties of wavelet sums, we have for $m \ge 0$

$$P_m(x,y) = N(x,y) + M_m(x,y). \tag{69}$$

According to Lemmas 2.10 and 2.11, $P_m(x,y) \underset{n\to\infty}{\to} 0$ uniformly a.e. in any set F with positive distance from $D = \{(x,y){:}\, x = y\}$. Because $M_m(x,y) \underset{n\to\infty}{\to} M(x,y)$ by Lemma 2.9, it follows that $N(x,y) = -M(x,y)$ for $x \neq y$. Since by (45) $M(x,y)$ is radially bounded by an L^1 convolution kernel H_1, the same is therefore true of $N(x,y)$, for $x \neq y$. Note however that $M_0(x,y) = 0$ (since its defining sum is empty), so that

$$|P_0(x,y)| \leq H_1(|x-y|). \tag{70}$$

Thus by Theorem 2.6, it follows that the multiresolution expansion of f converges to f a.e. and in L^p. Therefore by (68) the same holds for the wavelet expansion. Further, if ψ is partially continuous, then it is easy to show that as in (*ii*), the partial sum (67) is as well, and that $\int P_m(x,y)\, f(y)dy$ also is. Thus the two coincide a.e. and hence everywhere, so that, by Theorem 2.6 (*iv*), the wavelet expansion (67) converges as $m \to \infty$ on the Lebesgue set of f, as desired. Proof of L^p convergence (for $1 \leq p < \infty$) here and in case (*ii*) follows from the fact that the wavelet partial sums coincide a.e. with $\int P_m(x,y)\, f(y)dy$, and the result in (*i*) of Theorem 2.6.

The bound on $P_0(x,y)$ in (70) then also implies, by Theorem 2.6, that for $1 \leq p \leq \infty$, the multiresolution expansion of f converges to f a.e. If ψ is partially continuous, then so is $P_0(x,y)$ as shown above, so that by Theorem 2.6 the multiresolution expansion of f converges to f on its Lebesgue set for $f \in L^p$ $(1 \leq p \leq \infty)$.

Proof of (iv):

To prove (*iv*), we will show that the wavelet sums in statement (*iii*) converge order independently as stated, since the proof for scaling expansions (*ii*) follows similarly. Thus in the case of a wavelet expansion, assume that at stage t in the summation process $(-\infty < t < \infty)$ there is a finite collection $j_1 < j_2 < \ldots < j_q$ of values of j for which the sum over k and λ is incomplete. Recall that by assumption $j_q - j_1$ remains bounded at all stages in the summation. We separate the partial sum for f at stage t by:

$$f_t = \sum_{j<j_1;k,\lambda}^{\infty} a_{jk}^{\lambda}\psi_{jk}^{\lambda} + \sum_{1\leq i\leq q;(k,\lambda)\in K_i}^{\infty} a_{j_ik}^{\lambda}\psi_{j_ik}^{\lambda}, \tag{71}$$

where the set K_i consists of those pairs (k,λ) which have been summed for $j = j_i$ at the t^{th} stage of the summation.

We then write:

$$f_t = P_{j_1-1}f + P^t f, \tag{72}$$

where $P^t f$ is the linear projection taking f to the second sum in (71). On the other hand, the kernel of the projection $P^t(x,y)$ is bounded by

$$|P^t(x,y)| = \left| \sum_{1\leq i\leq q;k,\lambda\in K_1} \psi_{j_i}^{\lambda}(x)\overline{\psi_{j_i}^{\lambda}}(y) \right|$$

$$
\begin{aligned}
&\leq \sum_{1\leq i\leq q;k,\lambda\in K_i}\left|\psi^{\lambda}_{j_i}(x)\overline{\psi^{\lambda}_{j_i}}(y)\right| \\
&\leq \sum_{j_1\leq j\leq j_1+M;k,\lambda}\left|\psi^{\lambda}_{jk}(x)\overline{\psi^{\lambda}_{jk}}(y)\right|, \\
&= \sum_{j_1\leq j\leq j_1+M} Q^j(x,y),
\end{aligned}
$$

where

$$
Q^j(x,y)\equiv\sum_{k,\lambda}\left|\psi^{\lambda}_{jk}(x)\overline{\psi^{\lambda}_{jk}}(y)\right|,
$$

and M is the maximum value attained by j_1-j_q over all stages t of the summation (recall M is bounded by hypothesis). Unless otherwise specified, the sum over k and λ above is over all k and λ.

On the other hand, by an argument exactly as in (44), we can show that

$$
|Q^j(x,y)|\leq 2^{jd}H(2^j|x-y|),
$$

where $H(|\cdot|)\in\mathcal{RB}$ is a radially bounded L^1 convolution kernel. It thus follows that

$$
|P^t(x,y)|\leq 2^{j_1 d}H_1(2^{j_1}|x-y|),
$$

where $H_1\in\mathcal{RB}$; the above bound on $P^t(x,y)$ follows since each term in the sum defining it is bounded in this way (again we note that the range M of the summation defining P^t remains bounded). Since, by the above, the sum defining $P^t(x,y)$ converges absolutely, we also have

$$
\int P^t(x,y)dy=\sum_{j_1\leq j\leq j_1+M;k,\lambda}\psi^{\lambda}_{jk}(x)\int\overline{\psi^{\lambda}_{jk}}(y)dy=0
$$

by the dominated convergence theorem and the fact that $\int\overline{\psi^{\lambda}_{jk}}(y)dy=0$. Thus by Theorem 2.2, we have that $P^t f \underset{n\to\infty}{\longrightarrow} 0$ almost everywhere. On the other hand, by (iii) we have $P_{j_1-1}f \underset{n\to\infty}{\longrightarrow} f$ a.e. Thus since as $t\to\infty$ we have $j_1\to\infty$, we conclude by (72) that

$$
f_t=P^t f+P_{j_1-1}f \underset{n\to\infty}{\longrightarrow} f
$$

almost everywhere, as desired.

The proof that L^p $(1\leq p<\infty)$ convergence to f also occurs with the orders of summation mentioned above follows similarly, using the fact that for $f\in L^p$, $P^t f \underset{n\to\infty}{\longrightarrow} 0$ in L^p, by Theorem 2.2. □

Proof of Proposition 1.5: All the statements in this proposition follow from the proof above of Theorem 1.3. □

Remark: The condition that ψ satisfy some weak continuity property in order for convergence to 0 to occur everywhere cannot be removed, as can be seen in the simple example of Haar expansions on $\mathbb{R}$, which are discussed in section 1.

3 Special classes of wavelets

The ideas in the last section will now be used to explore properties of specific classes of wavelets on $\mathbb{R}^1$, namely the decay properties of their summation kernels P_j, and convergence properties of expansions. Let $\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k)$ for a scaling function $\phi(x)$ of a multiresolution analysis. Then by the results of the previous section, if $\phi \in \mathcal{RB}$, then

$$P_j(x, y) = \sum_k \phi_{jk}(x)\overline{\phi_{jk}}(y)$$

is the kernel of the orthogonal projection into the multiresolution space V_j. Following Meyer [Me1], ϕ is *regular* if there exists a $c > 0$ such that $|\phi(x)|,\ |\phi'(x)| \leq c/(1+|x|)^2$ for all $x \in \mathbb{R}$. Note that for any regular wavelet ϕ, we have $\phi \in \mathcal{RB}$, so that all results of the previous section apply. We will examine here more specific properties of the kernels of certain classes of regular wavelets.

P.G. Lemarié and Y. Meyer [LM] constructed scaling functions and wavelets in $\mathcal{S}(\mathbb{R})$, the space of functions $\phi(x)$ with rapid decay, satisfying $|D^m\phi(x)| \leq C_{mN}(1 + |x|)^{-N}$ for all m and N, with the proper choice of constants C_{mN}. Here we consider somewhat larger classes of scaling functions with polynomial order of decay N, i.e.,

$$\phi(x) \leq C(1 + |x|)^{-N} \tag{73}$$

for some C (all of which of course include the above class). For such classes, the summation kernels satisfy the following bounds.

Theorem 3.1 *If the scaling function ϕ has polynomial order of decay N with $N > 1$, then for some constant C,*

$$|P_j(x, y)| \leq C\frac{2^j}{(1 + 2^j|x - y|)^N} \leq C2^j. \tag{74}$$

Proof. It suffices by the scaling properties of P proved in section 2 to prove this bound for P_0. By our previous results and the assumptions we have (redefining constants C wherever necessary):

$$|P_0(x, y)| \leq C\sum_{k\in\mathbb{Z}}(1 + |x - k|)^{-N}(1 + |y - k|)^{-N}.$$

Without loss of generality, we can assume that $y \leq x$. The above sum can be bounded by an integral, yielding

$$\begin{aligned} |P_0(x, y)| &\leq C\int_{\mathbb{R}}(1 + |x - z|)^{-N}(1 + |y - z|)^{-N}dz \\ &\leq C\int_{-\infty}^{(x+y)/2}(1 + |x - z|)^{-N}(1 + |y - z|)^{-N}dz \end{aligned}$$

$$
\begin{aligned}
& +C\int_{(x+y)/2}^{\infty}(1+|x-z|)^{-N}(1+|y-z|)^{-N}dz \\
\leq\ & C(1+|(x-y)/2|)^{-N}\int_{-\infty}^{(x+y)/2}(1+|y-z|)^{-N}dz \\
& +C(1+|x-y|)^{-N}\int_{(x+y)/2}^{\infty}(1+|x-z|)^{-N}dz \\
\leq\ & C(1+|(x-y)/2|)^{-N}\int_{-\infty}^{\infty}(1+|z|)^{-N}dz \\
& +C(1+|x-y|)^{-N}\int_{-\infty}^{\infty}(1+|z|)^{-N}dz \\
\leq\ & C(1+|(x-y)|)^{-N},
\end{aligned}
$$

where the value of C has been readjusted above wherever necessary. □

An interesting class of wavelets is the spline wavelets on $L^p(\mathbb{R})$. These wavelets were developed by J.-O Strömberg [Strö], G. Battle [Ba], and P.G. Lemarié [Le1] . The multiresolution space V_j for such wavelets is defined as a spline space:

$V_j = \{f \in L^2(\mathbb{R}) \colon f \in C^{N-1}$ and f is a polynomial of degree N in $[l/2^j, (l+1)/2^j]$, $l \in \mathbb{Z}\}$

Again, a scaling function ϕ can be constructed, and $P_j(x,y) = \sum_k \phi_{jk}(x)\overline{\phi_{jk}}(y)$. Thus it is again useful to obtain bounds on

$$
P_j f(x) - f(x) = \int_{\mathbb{R}} [f(y) - f(x)] P_j(x,y)dy,
$$

via a bound on $P_j(x,y)$.

These wavelets have exponential decay, that is, $|\phi(x)|$ and $|\psi(x)|$ are bounded by $Ce^{-a|x|}$ as $|x| \to \infty$ for some $a > 0$. With this condition, a bound on the kernel can be found.

Theorem 3.2 *Let $P_j(x,y)$ be the projection kernel onto the subspace V_j defined above, for spline wavelets of order N. Then*

$$
|P_j(x,y)| \leq C_a 2^j e^{-a2^j|x-y|}. \tag{75}
$$

The proof of Theorem 3.2 can be completed in a similar way to that of Theorem 3.1, with the replacement of the function $\frac{1}{(1+|x|)^N}$ by $e^{-a|x|}$. It is clear from the proof of Theorem 3.2 that it in fact holds for any scaling function which is exponentially bounded.

We remark that as shown in the previous section, almost everywhere convergence follows directly from bounds such as in Theorems 3.1 and 3.2. Further, results on rates of convergence can also be obtained from these [Ke1].

Finally, similar results are also true for periodic spline wavelets on $L^2[0,1)$. Here, ϕ and ψ are the same functions as those for the splines defined on the line. To periodize these wavelets, we define ψ_{jk} and ϕ_{jk} as follows:

$$\psi_{jk}(x) = 2^{j/2} \sum_{l\in\mathbb{Z}} \psi(2^j(x+l)-k); \qquad \phi_{jk}(x) = 2^{j/2} \sum_{l\in\mathbb{Z}} \phi(2^j(x+l)-k).$$

In the periodic case, the summation kernel $P_m(x,y)$ is slightly different from that on the line, namely,

$$P_m(x,y) = 1 + \sum_{j=0}^{m} \sum_{k=0}^{2^j-1} \psi_{jk}(x)\psi_{jk}(y) = \sum_{k=0}^{2^m-1} \phi_{mk}(x)\phi_{mk}(y). \tag{76}$$

Without proofs, some of the pointwise convergence results for this case will be presented. Proofs can be found in [Ke1].

Theorem 3.3 *Let the kernel $P_m(x,y)$ be defined as in (76), with ψ_{jk} the periodic spline wavelets. Then*

$$|P_m(x,y)| \leq C_a 2^m e^{-a2^m\|x-y\|}, \tag{77}$$

where $\|x\|$ is defined to be the distance from x to the nearest integer.

Using this bound, convergence results can again be obtained.

Theorem 3.4 *For $f \in L^1[0,1)$, x in the Lebesgue set of f, the periodic spline wavelet expansion of f converges to $f(x)$ at x, i.e.,*

$$\lim_{m\to\infty} P_m f(x) = \int_0^1 f(y)dy + \lim_{m\to\infty} \sum_{j=0}^{m} \sum_{k=0}^{2^j-1} < f, \psi_{jk} > \psi_{jk}(x) = f(x). \tag{78}$$

In particular, convergence holds almost everywhere. Furthermore, if f is uniformly continuous, then the convergence is uniform.

The same statement regarding orders of convergence which is given in Theorem 1.3 (*iv*) also extends to this situation, as stated below:

Corollary 3.5 *The result of Theorem 3.4 also holds for more general orders of summation in which the range of values of j for which the sum over k is incomplete remains bounded.*

(See the remarks after the statement of Theorem 1.3 regarding the definitions of complete and incomplete sums).

Acknowledgments: The authors wish to thank Guido Weiss for bringing us together. The first author would like to acknowledge that this work was done under the direction of Richard Rochberg and Mitchell Taibleson, and she thanks them for their guidance. The last two authors wish to thank Ingrid Daubechies, Björn Jawerth, Stephane Mallat, and Robert Strichartz for useful comments.

References

[Au1] Auscher, P., Wavelet bases for $L^2(\mathbf{R})$ with rational dilation factor, in Ruskai. et al, eds., *Wavelets and Their Applications*, Jones and Bartlett, Boston, 1992, 439-452.

[Au2] Auscher, P., Solution of two problems of wavelets, preprint, Universitè de Rennes I, 1993.

[Ba] Battle, G., A block spin construction of ondelettes. Part I: Lemarié functions, *Communictions in Mathematics and Physics* **110** (1987), 601-615.

[deB] deBoor, C., On a Max-Norm Bound for the least-squares spline approximant, *Approximation and Function Spaces*, Z. Ciesielski, ed., Warsaw, 1981, 163-175.

[BDR] deBoor, C., R. Devore and A. Ron, Approximation from shift-invariant subspaces of $L^2(\mathbb{R}^d)$, Center for Mathematical Sciences, Reprort # 92-2, University of Wisconsin, 1992 (preprint).

[BM] Buhmann, M. and C. Micchelli, Spline prewavelets for non-uniform knots, *Numerische Mathematik* **61** (1992), 455-474.

[C] Carleson, L., On convergence and growth of partial sums of Fourier series, *Acta Mathematica* **116** (1966), 135-195.

[Ch] Chui, C.K., *An introduction to wavelets*, Academic Press, 1992.

[CW] Chui, C.K. and J. Wang, A cardinal spline approach to wavelets, *Proc. AMS* **113** (1991), 785-793.

[Da1] Daubechies, I., Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **41** (1988), 909-996.

[Da2] Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NSF Series in Applied Mathematics, SIAM, 1992.

[Fe] Fefferman, C., The new multiplier theorem for the ball, *Ann. Math.* **94** (1971), 330-336.

[GM] Grossman, A., and J. Morlet, Decomposition of Hardy functions into square-integrable wavelets of constant shape, *SIAM J. Math. Anal.* **15** (1984), pp. 723-736.

[GK1] Gurarie, D., and M. Kon, Radial bounds for perturbations of elliptic operators, *J. Functional Analysis* **56** (1984), 99-123.

[GK2] Gurarie, D., and M. Kon, Resolvents and regularity properties of elliptic operators, in *Operator Theory: Advances and Applications*, C. Apostol, Ed., Birkhauser Verlag, 1983.

[Ha] Haar, A., *Zur theorie der orthogonalen Funktionen-Systeme*, Math Ann. **69** (1910), 331-371.

[HT] Holschneider, R. and Ph. Tchamitchian, Régularité locale de la fonction 'non-differentiable' de Riemann, in *Les Ondelettes in* 1989, P. G. LeMarie, ed., Springer-Verlag, Berlin, 1990.

[Hu] Hunt, R.A., On the convergence of Fourier series, in *Proceedings of the Conference on Orthogonal Expansions and Their Continuous Analogues* (Southern Illinois University), 1968, 235-255.

[Ke1] Kelly, S., doctoral thesis, Washington University, St. Louis, 1992.

[Ke2] Kelly, S., Gibbs phenomenon for wavelets, preprint.

[KKR] Kelly, S., M. Kon and L. Raphael, Pointwise convergence for wavelet expansions, *Bulletin of AMS*, **30, #1**, (1994), 87-94.

[KST] Kenig, C.E., R.J. Stanton, and P.A. Tomas, Divergence of eigenfunction expansions, *J. Functional Analysis* **46** (1982), 28-44.

[Ko] Kon, M., Summation and spectral theory of eigenfunction expansions, in *Hermann Weyl: Forty More Years of Ramifications* (Anniversary Volume), Texas A&M University Press, 1991, 49-76.

[KR] Kon, M. and L. Raphael, Convergence rates of wavelet expansions, preprint.

[KRY] Kon, M., L. Raphael and J. Young, *Kernels and equisummation properties of uniformly elliptic operators*, J. Differential Equations **67** (1987), 256-268.

[Le1] Lemarié, P.G., Ondelettes á localisation exponentielle, to appear, *J. Math. Pures et Appl.*

[Le2] Lemarié-Riensset, P.G. Existence de fonctions-pères pour les ondelettes de dimension n, preprint, 1993 Orsay, France.

[LM] Lemarié, P.G. and Y. Meyer, Ondelettes et bases hilbertennes, *Rev. Math Iberamericana* **2** (1986), 1-18.

[Ma] Mallat, S., Multiresolution approximation and wavelets, *Trans. Am. Math. Soc.* **315** (1989), 69-88.

[Me1] Meyer, Y., *Ondelettes*, Hermann, Paris, 1990.

[Me2] Meyer, Y., Wavelets and Operators, *Proc. Special Year in Modern Analysis*, Urbana 1986-87, Cambridge University Press, 1989.

[Ru] Ruskai, M.B., G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, eds., *Wavelets and their applications*, Jones and Bartlett, Boston, 1992.

[Sh] Shapiro, H.S., Mathematics of convergence almost everywhere of convolution integrals with a dilation parameter, in *Multivariate Approximation Theory*, W. Schempp, ed., Springer LNM **vol. 571**, 250-266.

[SW] Stein, E. and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, 1971.

[Ste] Stein, E., *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

[Stra] Strang, G., Wavelets and dilation equations, a brief introduction, *SIAM Review* **31**(1989), 614-627.

[SF] Strang, G. and G. Fix, A Fourier analysis of the finite element variational method, in *Constructive Aspects of Functional Analysis*, Edizioni Cremonese, Rome, 1973.

[Stri] Strichartz, R., Wavelet expansions of fractal measures, *J. Geometric Anal.*, **1** (1991), 269-289.

[Strö] Strömberg, J.-O., A modified Franklin system and higher-order spline system on $\mathbf{R}^n$ as unconditional bases for Hardy spaces, in *Conference on Harmonic Analysis in honor of Antoni Zygmund*, W. Beckner et al., eds., Wadsworth, (1981).

[SP] Sweldens, W. and R. Piessens, Asymptotic error expansions for wavelet approximations of smooth functions, Technical Report TW164, Katholieke Universiteit Leuven, 1992.

[Wa1] Walter, G., Approximation of the delta function by wavelets, *J. Approx. Theo.* **71** (1992), 329-343.

[Wa2] Walter, G., Pointwise convergence of wavelet expansions, *J. Approx Theo.* (to appear).