Mathematical Properties of the JPEG2000 Wavelet Filters

Michael Unser* and Thierry Blu

accepted for publication in the *IEEE Trans. Image Processing*

EDICS: 2-WAVP,1-STAN

Biomedical Imaging Group, BIO-E, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne EPFL, Switzerland.
(e-mail:michael.unser@epfl.ch, thierry.blu@epfl.ch)

Abstract

The LeGall $5/3$ and Daubechies $9/7$ filters have risen to special prominence because they were selected for inclusion in the JPEG2000 standard. Here, we determine their key mathematical features: Riesz bounds, order of approximation, and regularity (Hölder and Sobolev). We give approximation theoretic quantities such as the asymptotic constant for the $\mathbf{L}^2$ error and the angle between the analysis and synthesis spaces which characterizes the loss of performance with respect to an orthogonal projection. We also derive new asymptotic error formulæ that exhibit bound constants that are proportional to the magnitude of the first non-vanishing moment of the wavelet. The Daubechies $9/7$ stands out because it is very close to orthonormal, but this turns out to be slightly detrimental to its asymptotic performance when compared to other wavelets with four vanishing moments.

I. INTRODUCTION

JPEG2000 has recently been approved as an international standard for the compression of still digital pictures [1]. Unlike JPEG which uses the DCT, the new standard is entirely wavelet-based; it incorporates many of the recent advances[1] that have occurred in the field [2], [3], [4], [5]. For the same subjective image quality, JPEG2000 image files tend to be 20 to 50% smaller than their JPEG counterparts depending on the kind of pictures [6]. The new format has also gotten rid of JPEG's main problem: the infamous blocking artifacts. The JPEG2000 format is rich and rather flexible. It gives a very strict specification of the decoding algorithm which uses an inverse separable wavelet transform. On the other hand, it only provides relatively general guidelines for the encoding algorithm leaving it up to designers to come up with the most efficient solutions. JPEG2000 restricts the user's choice to two wavelet transforms: Daubechies $9/7$ for lossy compression [7], and the $5/3$ LeGall wavelet [8], which has rational coefficients, for reversible or lossless compression. It also specifies that these should be implemented using the lifting scheme [9].

The fact that these two wavelets are an integral part of the international standard gives them a very special status. They are now much more likely to be used in applications than any other wavelets. Thus, the purpose of this note is to evaluate and document some of their mathematical properties within the framework of wavelet theory. What singles them out over others is that they are extremely short, symmetric (to avoid boundary artifacts), with a maximum number of regularity factors (or vanishing moments). This latter concern is at the very heart of wavelet theory and it is significant that it was given more weight

[1]Interestingly, the JPEG2000 coder is not directly based on the popular EZW or SPIHT algorithms; instead, it utilizes a versatile form of block/bitplane coding inspired by the work of Taubman.

by the committee than the purely spectral considerations that used to prevail in the early days of perfect reconstruction filter banks. Because of the minimum support requirement, both wavelets can be obtained by factorizing a maximally flat Daubechies or Dubuc-Deslaurier half-band filter [10]. The $5/3$ LeGall is the shortest symmetrical biorthogonal wavelet with two regularity factors; its synthesis scaling function is a linear B-spline. The $9/7$ is a variant of Cohen-Daubechies-Feauveau's biorthogonal cubic B-spline construction (shortest scaling of order four) [11] with residual factors that have been divided up on both sides in a way that makes the basis functions more nearly orthogonal. Also note that the order in which the filters are applied (analysis versus synthesis) is important: it is such that the shortest and most regular basis functions are placed on the synthesis side; this is consistent with the principle of maximizing the approximation power of the representation [12]. Intuitively, you want the shortest and smoothest basis functions to be on the synthesis side to minimize perceptual artifacts.

The quantities derived in this paper essentially fall into two categories. The first are descriptors of the mathematical properties of the underlying basis functions that are defined in the continuous domain (cf. Section II). Features of interest are the order of approximation, the L_2-stability of the representation (Riesz bounds), and the smoothness of the basis functions. The second category is more directly concerned with approximation theory, as explained in Section III. Our contribution is to present numerical indicators of the approximation power of the transforms (asymptotic error formulas and bounds). Of course, there are also limitations to our analysis because it does not take into account the non-linear aspect of JPEG2000 quantization nor the fact that the visual quality of compressed images is not solely determined by the L_2-norm of the error. It has the merit, however, of being quantitative and transform-specific, which is not the case for most other results available in this area—most notably, non-linear approximation theory, which does not distinguish between wavelet bases [13], [14], [15].

II. Wavelets and their mathematical properties

The wavelet decomposition algorithm uses two analysis filters $\tilde{H}(z)$ (lowpass) and $\tilde{G}(z)$ (highpass). The reconstruction algorithm applies the complementary synthesis filters $H(z)$ (refinement filter) and $G(z)$ (wavelet filter) [16]. These four filters constitute a perfect reconstruction filter bank [17], [18], [19], [20]. In the present case, the system is entirely specified by the lowpass filters $H(z)$ and $\tilde{H}(z)$, which form a biorthogonal pair; the highpass operators are obtained by simple shift and modulation: $\tilde{G}(z) = zH(-z)$ and $G(z) = z^{-1}\tilde{H}(-z)$.

The wavelet transform has a continuous-time domain interpretation that involves the scaling functions $\tilde{\varphi}(x)$ and $\varphi(x)$, which are solutions of two-scale relations with filters $\tilde{H}(z)$ and $H(z)$, respectively.

Definition 1—Two-scale relation: The scaling function $\varphi(x)$ associated with the filter $H(z)$ is the $\mathbf{L}^2$-solution (if it exists) of the two-scale relation

$$\varphi(x) = \frac{2}{H(1)} \sum_{k \in Z} h_k \varphi(2x - k) \tag{1}$$

While it is usually difficult to obtain an explicit characterization of $\varphi(x)$ in the time domain, one can express its Fourier transform as a convergent infinite product (cf. [21]):

$$\hat{\varphi}(\omega) = \prod_{k=1}^{\infty} \frac{H(e^{j\frac{\omega}{2^k}})}{H(1)} \tag{2}$$

A simple way to generate a scaling function is to run the synthesis part of the wavelet transform algorithm starting with an impulse—this is often referred to as the cascade algorithm (cf. [21], [17], [19], [20]).

Much of the early work in wavelet theory has been devoted to working out the mathematical properties (convergence, regularity, order, etc...) of these scaling functions. The wavelets themselves do usually not pose a problem because they are linear combination of the scaling functions; i.e.,

$$\psi(x) = \frac{2}{H(1)} \sum_{k} g_k \varphi(2x - k) \quad \text{and} \quad \tilde{\psi}(x) = \frac{2}{\tilde{H}(1)} \sum_{k} \tilde{g}_k \tilde{\varphi}(2x - k).$$

The corresponding analysis and synthesis wavelet basis functions are $\tilde{\psi}_{i,k} = 2^{-i/2}\tilde{\psi}(x/2^i - k)$ and $\psi_{i,k} = 2^{-i/2}\psi(x/2^i - k)$, respectively, where i and k are the translation and scale indices.

A necessary condition for the convergence of (2) to an $\mathbf{L}^2$-stable function $\varphi(x)$, is that the filter $H(z)$ has a zero at $z = -1$. More generally, the refinement filters will have a specified number of "regularity" factors[2], which determine their order of approximation.

Definition 2—Approximation order: The number L of factors $(1 + z^{-1})$ that divide $H(z)$ is called approximation order.

These factors play a crucial role in wavelet theory [19]. They imply that the scaling function $\varphi(x)$ reproduces all polynomials of degree lesser or equal to $n = L - 1$; in particular, it satisfies the *partition of unity* $\sum_k \varphi(x - k) = 1$. They are also directly responsible for the vanishing moments of the analysis wavelet [21], [20]: $\int x^n \tilde{\psi}(x) \, \mathrm{d}x = 0$ for $n = 0, \dots, L - 1$. Finally, the order L also corresponds to the rate of decay of the projection error as scale a goes to zero [22], [12].

Another important characteristic of a scaling function is its smoothness in the sense of degree of differentiability. This notion comes in many flavors, a unified treatment of which is provided by the notion of

[2]This widely used nomenclature is often a source of confusion. In wavelet theory, regularity usually means smoothness in the sense of Hölder (r) or Sobolev (s) (cf. Definitions 3 and 4). A minimum number of factors is necessary for regularity; i.e., $L \geq s$, but it is generally not sufficient.

Besov space [23]. The Besov regularity essentially specifies the fractional degree of differentiability of the function in the $\mathbf{L}^p$-sense. The most stringent measure corresponds to the case $p = \infty$ and coincides with the Hölder regularity which is a classical measure of pointwise continuity [24], [25]. The other most commonly-used measure is the Sobolev regularity which corresponds to the intermediate Besov case $p = 2$; it is a more global indicator of smoothness that is entirely specified in the Fourier domain [24], [25].

Definition 3—Hölder regularity: The smallest real number r_c such that: 1) φ is $N = \lceil r \rceil - 1$ times continuously differentiable for $r < r_c$; and 2)

$$\sup_{x,y\in\mathbb{R}} \frac{|\varphi^{(N)}(y) - \varphi^{(N)}(x)|}{|y-x|^{r-N}} \text{ is finite for all } r < r_c;$$

is called Hölder critical regularity exponent.

Definition 4—Sobolev regularity: The smallest real number s_c such that

$$\int |\hat{\varphi}(\omega)|^2 (1+\omega^2)^s \,\mathrm{d}\omega \text{ is finite for all } s < s_c,$$

is called Sobolev critical regularity exponent.

Both measures are linked through the following inequality: $r_c \leq s_c \leq r_c + \frac{1}{2}$ (under slight conditions, see [25]). For compactly supported functions, there is also a similar inequality that holds in the more general $\mathbf{L}^p$-case: $r_c \leq s_c(p) \leq r_c + \frac{1}{p}$ where the upper bound comes as a consequence of the Besov embedding theorems. While the regularity of wavelets has been a subject of major concern for mathematicians, it is still not clear how important it is in practical applications (cf. [26], [27]). However, most researchers agree that a minimum of regularity (typically, continuity) is desirable for a good convergence behavior of the iterated filterbank.

The next point concerns the stability of the wavelet representation and of its underlying multiresolution bases. The crucial mathematical property is that the translates of the scaling functions and wavelets form Riesz bases [28]. Thus, one needs to characterize their Riesz bounds and other related quantities.

Definition 5—Cross-correlation: The 2π-periodic function $a_{\varphi_1,\varphi_2}(\omega)$

$$\begin{aligned} a_{\varphi_1,\varphi_2}(\omega) &= \sum_{k\in\mathbb{Z}} \hat{\varphi}_2(\omega + 2k\pi)^* \hat{\varphi}_1(\omega + 2k\pi) \\ &= \sum_{k\in\mathbb{Z}} e^{-kj\omega} \varphi_{12}(k) \end{aligned}$$

is the cross-correlation filter associated with the pair $\{\varphi_1, \varphi_2\}$; $\varphi_{12}(x) = \int \varphi_2(\xi)\varphi_1(\xi + x)\,\mathrm{d}\xi$ is the corresponding cross-correlation function.

Thus, one can define a biorthogonal pair $\{\varphi, \tilde{\varphi}\}$ as a set of scaling functions for which the cross-correlation filter is identity; i.e., $a_{\varphi,\tilde{\varphi}}(\omega) = 1$. Here, we will mostly consider the autocorrelation filter $a_{\varphi,\varphi}(\omega)$ also denoted by $a_\varphi(\omega)$.

Definition 6—Riesz bounds: The tightest upper and lower bounds, $B < \infty$ and $A > 0$, of the autocorrelation filter of $\varphi(x)$ are the Riesz bounds of $\varphi(x)$; i.e., $A^2 = \inf_{\omega\in[0,2\pi]} a_\varphi(\omega)$ and $B^2 = \sup_{\omega\in[0,2\pi]} a_\varphi(\omega)$. Equivalently, they satisfy

$$A = \inf_{c\in\ell^2} \frac{\left\| \sum_{k\in\mathbb{Z}} c_k \varphi(x-k) \right\|_{\mathbf{L}^2}}{\|c\|_{\ell^2}} \quad \text{and} \quad B = \sup_{c\in\ell^2} \frac{\left\| \sum_{k\in\mathbb{Z}} c_k \varphi(x-k) \right\|_{\mathbf{L}^2}}{\|c\|_{\ell^2}}$$

The existence of the Riesz bounds ensures that the underlying basis functions are in $\mathbf{L}^2$ and that they are linearly ℓ^2-independent. The Riesz basis property expresses an equivalence between the $\mathbf{L}^2$-norm of the expanded functions and the ℓ^2-norm of their coefficients in the wavelet or scaling function basis. There is a perfect norm equivalence (Parseval's relation) if and only if $A = B = 1$, in which case the basis is orthonormal.

Many key concepts in wavelet theory are best explained in terms of projection operators. Given a set of biorthogonal functions $\tilde{\varphi}$ and φ, we define the projection operator at scale $a = 2^i$

$$\forall f \in \mathbf{L}^2,\ \ \tilde{\mathcal{P}}_a f(x) = \sum_{k\in\mathbb{Z}} \langle f(a\cdot), \tilde{\varphi}(\cdot - k)\rangle\, \varphi\Big(\frac{x}{a} - k\Big) \tag{3}$$

which produces an approximation of a function $f \in \mathbf{L}^2$ within the subspace $\mathcal{V}_a = \mathrm{span}_{k\in\mathbb{Z}}\{\varphi(\frac{x}{a} - k)\}$ at resolution a. Interestingly, this approximation can be interpreted geometrically as a projection into $\mathcal{V}_a$ perpendicular to the analysis space $\tilde{\mathcal{V}}_a = \mathrm{span}_{k\in\mathbb{Z}}\{\tilde{\varphi}(\frac{x}{a} - k)\}$ [11]. In general, $\tilde{\mathcal{P}}_a f$ will not provide the orthogonal projection which we denote $\mathcal{P}_a$, unless the "angle" between the analysis and synthesis spaces is zero; i.e., unless $\mathcal{V}_a = \tilde{\mathcal{V}}_a$.

Definition 7—Projection angle: The (generalized) projection angle θ between the synthesis and analysis subspaces $\mathcal{V}_a$ and $\tilde{\mathcal{V}}_a$ is defined as [29]

$$\cos\theta \stackrel{\text{def}}{=} \inf_{f\in\tilde{\mathcal{V}}_a} \frac{\|\mathcal{P}_a f\|_{\mathbf{L}^2}}{\|f\|_{\mathbf{L}^2}} = \frac{1}{\sup\limits_{\omega\in[0,2\pi]} \sqrt{a_\varphi(\omega)\, a_{\tilde{\varphi}}(\omega)}} \tag{4}$$

This fundamental quantity is scale-independent; it allows us to compare the performance of the biorthogonal projection $\tilde{\mathcal{P}}_a$ with that of the optimal least squares solution $\mathcal{P}_a$ for a given approximation space $\mathcal{V}_a$. Specifically, we have the following sharp error bound (cf. [29])

$$\forall f \in \mathbf{L}^2,\ \ \|f - \mathcal{P}_a f\|_{\mathbf{L}^2} \le \|f - \tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2} \le \frac{1}{\cos\theta} \|f - \mathcal{P}_a f\|_{\mathbf{L}^2} \tag{5}$$

In other words, the biorthogonal projector $\tilde{\mathcal{P}}_a$ will be essentially as good as the optimal one (orthogonal projector onto the same space) provided that $\cos\theta$ is close to one. Moreover, the angle is directly related to the norm of the underlying operator: $\|\tilde{\mathcal{P}}_a\| \stackrel{\text{def}}{=} \sup_f \frac{\|\tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2}}{\|f\|_{\mathbf{L}^2}} = \frac{1}{\cos\theta}$ (see [30]). The above results are also directly transposable to the corresponding wavelet spaces $\tilde{W}_i = \mathrm{span}_{k\in\mathbb{Z}}\{\tilde{\psi}_{i,k}\}$ and $W_i = \mathrm{span}_{k\in\mathbb{Z}}\{\psi_{i,k}\}$ because the geometry and the angle remain the same [31].

III. Wavelets and approximation theory

The JPEG2000 algorithm takes advantage of two fundamental properties of wavelet expansions: (1) the magnitudes of the wavelet coefficients are strongly correlated across scales (persistence across-scale property), and (2) the wavelet coefficients of a piecewise smooth image essentially fall into two categories: the ones with large amplitudes that are located near edges, and the smaller ones, for the most part included in the smooth regions of the image. To understand why this is the case, one has to turn to approximation theory which also provides us with quantitative indicators (bound constants) that characterize the approximation power of a given transform.

A. Linear approximation theory

Linear approximation theory is concerned with the characterization of the approximation error as a function of the scale. It is possible, for instance, to predict the projection error $\|f - \mathcal{P}_a f\|_{\mathbf{L}^2}$ quite precisely by integration of the spectrum of the function $f \in \mathbf{L}^2$ to approximate against the so-called Fourier approximation kernel [32], [33]:

$$\|f - \tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2} = \sqrt{\frac{1}{2\pi}\int |\hat{f}(\omega)|^2 E_{\tilde{\varphi},\varphi}(\omega a)\,\mathrm{d}\omega}. \tag{6}$$

Definition 8—Fourier approximation kernel: The Fourier approximation kernel associated with the functions φ and $\tilde{\varphi}$ is given by

$$E_{\tilde{\varphi},\varphi}(\omega) = |1 - \hat{\tilde{\varphi}}(\omega)^*\hat{\varphi}(\omega)|^2 + |\hat{\tilde{\varphi}}(\omega)|^2 \sum_{n\neq 0} |\hat{\varphi}(\omega + 2n\pi)|^2 \tag{7}$$

Note that the equality in (6) is rigourously exact only if the function f is bandlimited or if the error is averaged over all possible shifts of the input: $f(x-\tau)$, with $0 \le \tau \le a$. Otherwise, the formula calls for a small correction term that is bounded in terms of the smoothness of the input signal (cf. [33, Thm.1]). Based on the approximation result above, it is possible to derive the following inequality (proof in Appendix A)

$$\|f - \tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2} \le C\, a^r\, \|f^{(r)}\|_{\mathbf{L}^2} \quad \text{where } r = \min\{L, s\} \tag{8}$$

where L is the order of the transform and where s is a differentiation depth for which $\|f^{(s)}\|_{\mathbf{L}^2}$, the norm of the s^{th} (fractional) derivative of f, is still finite (thus s may be arbitrarily close to the Sobolev critical regularity exponent of f). This result indicates that the approximation error decays like the r^{th} power of the scale. There are essentially two regimes: the "rough" one ($s < L$), where the rate of decay is specified by the Sobolev smoothness (or lack thereof) of the input signal, and the "smooth" one ($s \geq L$) where it is limited by the order of approximation of the transform (cf. Definition 2). For $a = 2^i$, the quantity on the left hand side of (8) also corresponds to a scale-truncated wavelet expansion with a scale index running from $+\infty$ down to i (non-inclusive). Likewise, the wavelet component at scale i is given by $r_i = \sum_{k\in\mathbb{Z}} \langle f, \tilde{\psi}_{i,k}\rangle\, \psi_{i,k} = \tilde{\mathcal{P}}_{2^{i-1}} f - \tilde{\mathcal{P}}_{2^i} f$ which we may also write as $r_i = (f - \tilde{\mathcal{P}}_{2^i} f) - (f - \tilde{\mathcal{P}}_{2^{i-1}} f)$. This allows us to carry over the bound to the wavelet residual by making use of the triangle inequality:

$$\Big\| \sum_{k\in\mathbb{Z}} \langle f, \tilde{\psi}_{i,k}\rangle\, \psi_{i,k} \Big\| = \|(f - \tilde{\mathcal{P}}_{2^i} f) - (f - \tilde{\mathcal{P}}_{2^{i-1}} f)\|_{\mathbf{L}^2} \leq D\, 2^{ir}\, \|f^{(r)}\|_{\mathbf{L}^2} \tag{9}$$

Finally, we invoke the Riesz norm equivalence to get a comparable $O(a^r)$ decay for the ℓ^2-norm of the wavelet coefficients at scale i

$$\left(\sum_{k\in\mathbb{Z}} |\langle f, \tilde{\psi}_{i,k}\rangle|^2 \right)^{\frac{1}{2}} \leq D'\, 2^{ir}\, \|f^{(r)}\|_{\mathbf{L}^2} \tag{10}$$

Qualitatively, the results that have just been described are also valid locally. For instance, Jaffard proved that the wavelet coefficients decay like $O(a^{\min\{L,\alpha\}+\frac{1}{2}})$ where α is the Hölder regularity of the signal within the wavelet cone of influence [34], [20]. Thus, a high order has the beneficial effect of producing very small coefficients—hence lots of zeros—in the smooth regions of the image. On the other hand, one would like the influence cone to be reasonably small which calls for a compromise in terms of filter length.

B. Asymptotics for the approximation of smooth functions

The decay results that have been mentioned so far are qualitative and do not distinguish between wavelet types, except for their order property. While it is difficult to say much more for the "rough" mode where the wavelet decay is entirely driven by the signal discontinuities, one can get more quantitative for the "smooth" regime which is relevant for the regions where the signal is slowly varying. Specifically, it is possible to determine the exact asymptotic form of the approximation error in (8):

$$\|f - \tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2} \underset{a\to 0}{=} C_\varphi^-\, \|f^{(L)}\|_{\mathbf{L}^2}\, a^L, \tag{11}$$

an expression that is valid provided that the sampling step a is small compared to the smoothness scale of the signal f as measured by $\|f^{(L)}\|_{\mathbf{L}^2}$.

Definition 9—Asymptotic approximation constant: The constant C_φ^- that rules the decay of the approximation error (11) when the scale a tends to 0 is called "asymptotic approximation constant".
This constant is independent of $\tilde{\varphi}$ whenever $\tilde{\varphi}$ has order 1; i.e., oblique and orthogonal projection are asymptotically equivalent [12].

Similarly, we can obtain the exact asymptotic forms corresponding to (9) and (10) (cf. proof in Appendix B). In particular, we show that there is an exact proportionality between the asymptotic approximation error at scale i and its corresponding wavelet residual:

$$\Big\| \sum_{k\in\mathbb{Z}} \langle f, \tilde{\psi}_{i,k}\rangle \, \psi_{i,k} \Big\| = \sqrt{1-4^{-L}}\, \|(f - \tilde{\mathcal{P}}_{2^i} f)\|_{\mathbf{L}^2}, \quad \text{as } i \to -\infty \tag{12}$$

irrespective of the type of projector (orthogonal or not). The ℓ^2-norm of the wavelet coefficients follows the same asymptotic pattern

$$\left(\sum_{k\in\mathbb{Z}} |\langle f, \tilde{\psi}_{i,k}\rangle|^2 \right)^{\frac{1}{2}} = C_{\tilde{\psi}}\, 2^{iL}\, \|f^{(L)}\|_{\mathbf{L}^2}, \quad \text{as } i \to -\infty \tag{13}$$

with a constant that is proportional to the magnitude of the first non-vanishing moment of the analysis wavelet: $C_{\tilde{\psi}} = \frac{1}{L!}\int x^L \tilde{\psi}(x)\,\mathrm{d}x$.

Thus, we may compare the approximation power of wavelets of the same order L by looking at the magnitude of their asymptotic approximation constants. This naturally leads to the definition of the asymptotic sampling gain of one system (φ) against a reference one (φ_0) with the same order: $\gamma = \sqrt[L]{\frac{C_{\varphi_0}^-}{C_\varphi^-}}$. In other words, by using φ instead of φ_0, we can increase the sampling step by γ and keep the asymptotic error within the same bound [35].

C. Non-linear approximation theory

For completeness sake, we mention the existence of theoretical results for non-linear wavelet approximation which have been shown to be very relevant for coding [13], [14], [15]. The non-linear aspect here is that the signal is reconstructed from its n largest expansion coefficients, which is equivalent to applying a signal-dependent threshold in the wavelet domain. DeVore and others proved that wavelet bases provide the best n-terms approximation for certain types of smoothness spaces (Hölder, Sobolev, and Besov) [36]. These results rest on the fact that wavelets are unconditional bases for Besov spaces [28], [37]. Of course, this can only be the case if the basis functions themselves are included in the corresponding smoothness space. Thus, wavelets with higher degrees of smoothness extend the range of applicability of these approximation theorems and give access to better rates of decay for the n-term approximation error, provided of course that the function to approximate is sufficiently smooth as well. This theory, however, does not provide any hard bound estimates; it is mainly concerned with achievable rates of decay and does not distinguish between wavelet bases.

IV. METHODS

We now briefly indicate how to compute the mathematical parameters that have been identified starting from the wavelet filters.

A. *Exact computation of the scaling function*

There is a well-known procedure for determining the values of a scaling function at the dyadic rationals [24]. Here, we describe a matrix approach that allows an exact determination of $\varphi(n)$ with n integer. Clearly, $\varphi(n) \neq 0$ only when $l_H \leq n \leq L_H$, where $\{l_H \dots L_H\}$ is the support of the filter H. Thus, we can consider the two scale relation (1) for $x = n$ and write it as an equivalent matrix equation $\mathbf{A}\mathrm{y} = \mathrm{y}$, where the vector y is defined by $[\mathrm{y}]_n = \varphi(n)$ for $n = l_H \dots L_H$, and the matrix coefficients are given by $[\mathbf{A}]_{k,l} = \frac{2}{H(1)} h_{2k-l}$. Next, we use the partition of unity, $\sum_k \varphi(k) = 1$ because φ is at least of order 1. Hence, the eigen-solution of $\mathbf{A}\mathrm{y} = \mathrm{y}$ that we are seeking is the one that satisfies $\mathrm{y}_{l_H} + \mathrm{y}_{l_H+1} + \dots \mathrm{y}_{L_H} = 1$ [19], [35]. Thus, the solution is a normalized version of the eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 1. Here, we propose to determine this vector explicitly from the solution of a linear system of equations that is a slightly modified version of $(\mathbf{A} - \mathbf{I})\mathrm{y} = 0$ with one row being substituted by the partition of unity constraint.

B. *Exact computation of the Hölder and Sobolev exponents*

In general, there is no systematic algorithm to compute the Hölder regularity exponent for an arbitrary scaling function φ, and one usually has to turn to numerical methods to obtain estimates [24], [38], [25], [23]. Fortunately for us, an exact procedure is available when the filters are *symmetric* and *positive on the unit circle* (i.e., $H(e^{j\omega}) \geq 0$ for $\omega \in [-\pi, \pi]$), which is precisely the case for JPEG2000 filters. It is as follows:

- Eliminate all the order factors $(1 + z^{-1})$ from $H(z)$; i.e., find $Q(z)$ such that $H(z) = \frac{H(1)}{2} z^{\frac{L}{2}} (1 + z^{-1})^L Q(z)$. Note that, because of the positivity on the unit circle, L is necessarily even.

- Build the matrix $[\mathbf{Q}]_{k,l} = q_{2k-l}$ where $-L_Q \leq k, l \leq L_Q$ and $\{-L_Q \ldots L_Q\}$ is the support of the filter Q.
- Find the largest eigenvalue, λ, of $\mathbf{Q}$. Then $r = -\log_2 |\lambda|$.

Computing the Sobolev exponent is an easier theoretical task; unlike the Hölder exponent, it can *always* be achieved through the determination of the spectral radius of a reduced transition operator [39], [25], without any symmetry or positivity assumption on the filter. It turns out that the algorithm is essentially the same as the one outlined above with the filter now being $H_2(z) = H(z)H(z^{-1})$. This is simply because the Sobolev exponent is one half the Hölder exponent of the autocorrelation function $\varphi_2(x) = \int \varphi(\xi)\varphi(\xi + x)\,\mathrm{d}\xi$ which is the scaling function corresponding to the refinement filter $H_2(z)$. Thus the method consists in computing the Hölder exponent r_2 of φ_2, which provides $s = \frac{1}{2}r_2$.

C. Exact calculation of Riesz bounds and projection angles

These determinations are straightforward once we know the auto-correlation filters $a_\varphi(\omega)$ and $a_{\tilde{\varphi}}(\omega)$, which are trigonometric polynomial in $e^{-j\omega}$. More generally, we can consider the definition 5 of the cross-correlation filter $a_{\varphi_1,\varphi_2}(\omega)$ which is entirely specified by the integer samples of the cross-correlation function φ_{12}. Indeed, if φ_1 and φ_2 are both scaling functions with filters $H_1(z)$ and $H_2(z)$, then $\varphi_{12}(x)$ is a scaling function as well with refinement filter $\tilde{H}_1(z^{-1})H_2(z)$. Thus, we can determine the $\varphi_{12}(n)$'s exactly using the computational procedure in IV-A.

D. Determination of the Fourier approximation kernel

The development of the expression (7) yields $E_{\tilde{\varphi},\varphi}(\omega) = 1 - 2\,\mathscr{R}\left\{\hat{\tilde{\varphi}}(\omega)^* \hat{\varphi}(\omega)\right\} + |\hat{\tilde{\varphi}}(\omega)|^2 a_\varphi(\omega)$. The functions $\hat{\varphi}(\omega)$ and $\hat{\tilde{\varphi}}(\omega)$ are computed to any desired precision using the infinite product (2) while the autocorrelation filter $a_\varphi(\omega)$ is obtained by the method described above.

E. Exact formula for the asymptotic constant

The initial infinite sum formula proposed in [12] was latter simplified in [33]. A concise version of this formula is

$$C_\varphi^- = \frac{\sqrt{a_\varphi(\pi)}}{L!\,\sqrt{4^L - 1}} \left| \frac{H^{(L)}(-1)}{H(1)} \right| \tag{14}$$

where L is the approximation order of φ; $H^{(L)}(-1)$ denotes the L^{th} derivative of $H(z)$ evaluated at $z = -1$.

V. RESULTS

A. *LeGall* $5/3$ *filters*

The JPEG2000 LeGall $5/3$ scaling filters are given by

$$\tilde{H}(z) = \frac{1}{8}\, z(1+z^{-1})^2(-z-z^{-1}+4) \quad \text{and} \quad H(z) = \frac{1}{2}\, z(1+z^{-1})^2$$

These filters form the shortest symmetrical biorthogonal pair of order 2. Their main characteristics are summarized in Table I. The synthesis function (Fig. 2) is the linear B-spline which is boundedly differentiable. This is a classical function whose smoothness in the general $\mathbf{L}^p$ sense has also been characterized: $s_c(p) = 1 + 1/p$ [23]; interestingly, this corresponds to the upper bound specified by the Besov embedding theorems. The analysis function (Fig. 1) is also in $\mathbf{L}^2$ but has no smoothness at all; it is merely bounded.

The Riesz bounds are not very tight indicating that the functions are far from orthogonal. The cosine is reasonably close to one; this is good news for the underlying projector which should essentially perform as well as a least squares piecewise linear approximation. Our results also show that the approximation properties of the LeGall $5/3$ are much better than those of the Daubechies filter of the same order ($\mathrm{Daubechies}_4$). The corresponding sampling gain is $\gamma = \frac{1}{\sqrt{1/3}} \approx 1.73$. By comparing the Fourier approximation kernels in Fig. 3, we also see that the LeGall filters offer a near 10 dB improvement over $\mathrm{Daubechies}_4$ over the first half of the Nyquist band.

A less favorable aspect of this transform is the relative magnitude of the wavelet constant which comes as a consequence of the lack of tightness of the Riesz bounds. While this may compromize the efficacy of bit-allocation in the wavelet domain, it is not really a problem here because this wavelet transform is intended to be used for lossless coding.

B. *Daubechies* $9/7$ *filters*

We have derived the exact form of the JPEG2000 Daubechies $9/7$ scaling filters:

$$\tilde{H}(z) = \frac{2^{-5}}{\frac{64}{5\rho} - 6 + \rho}\, z^2(1+z^{-1})^4\Big(z^2 + z^{-2} - (8-\rho)(z+z^{-1}) + \frac{128}{5\rho} + 2\Big) \tag{15}$$

$$H(z) = \frac{2^{-3}}{\rho - 2}\, z^2(1+z^{-1})^4(-z-z^{-1}+\rho) \tag{16}$$

where $\rho \approx 3.3695$ is the unique real root of $128 - 116x + 40x^2 - 5x^3$, i.e.,

$$\rho = \frac{8}{3} + \frac{2}{3}\sqrt[3]{\frac{7}{25}}\left(\sqrt[3]{10+3\sqrt{15}} + \sqrt[3]{10-3\sqrt{15}}\right) \tag{17}$$

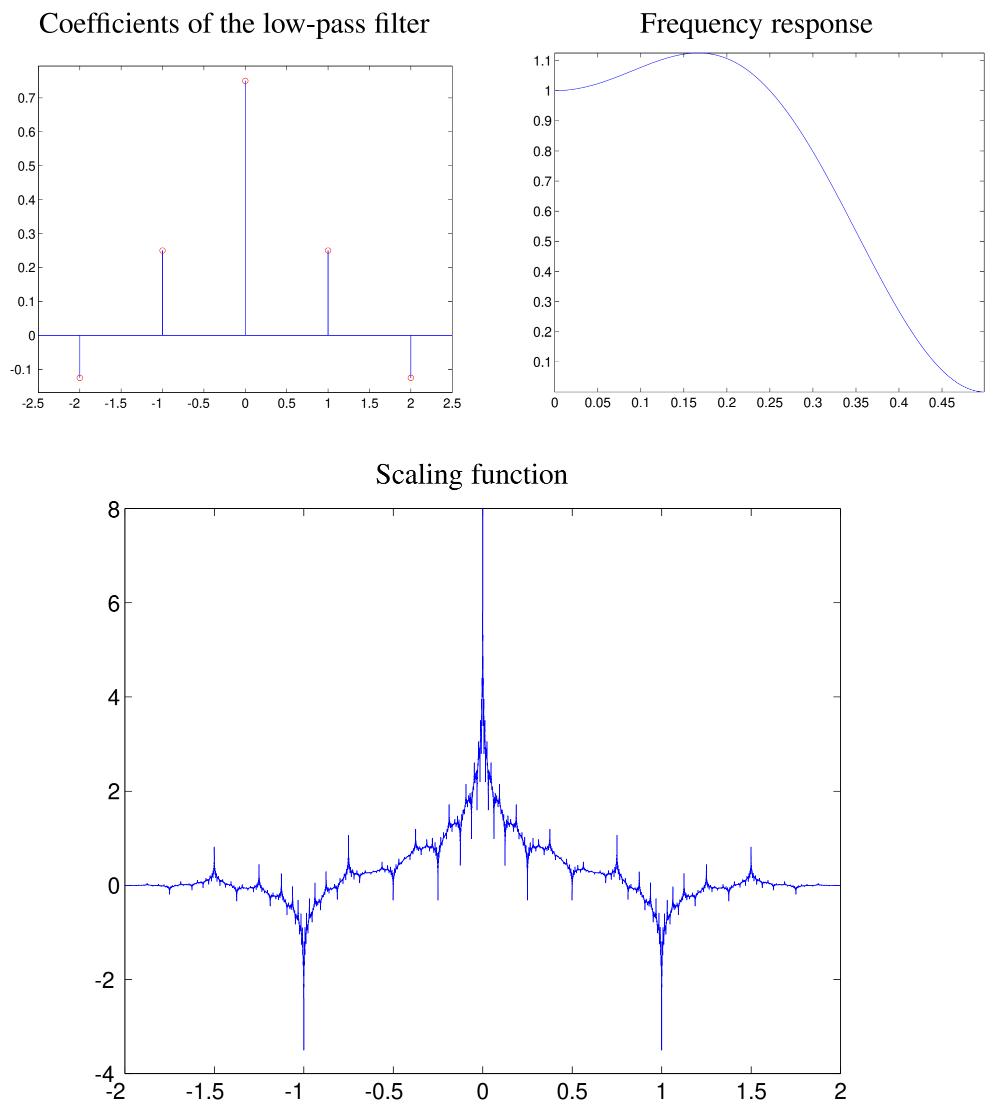

Fig. 1. Analysis LeGall JPEG filter of length 5 and corresponding scaling function.

These filters result from the factorization of the same polynomial as $\mathrm{Daubechies}_8$ [10]. The main difference is that the $9/7$ filters are symmetric. Moreover, unlike the biorthogonal splines of Cohen-Daubechies-Feauveau [11], the non-regular part of the polynomial has been divided among both sides, and as evenly as possible.

The mathematical properties of the $9/7$ filters are summarized in Table II. These are 4^{th} order filters meaning that the wavelets have 4 vanishing moments. Here too, the shortest basis functions are placed on the synthesis side. All functions are at least once continuously differentiable with a fair amount of extra smoothness on the synthesis side—this is quite visible when comparing Figs. 4 and 5.

Perhaps the most important property that is truly specific to the $9/7$ is the tightness of Riesz bounds,

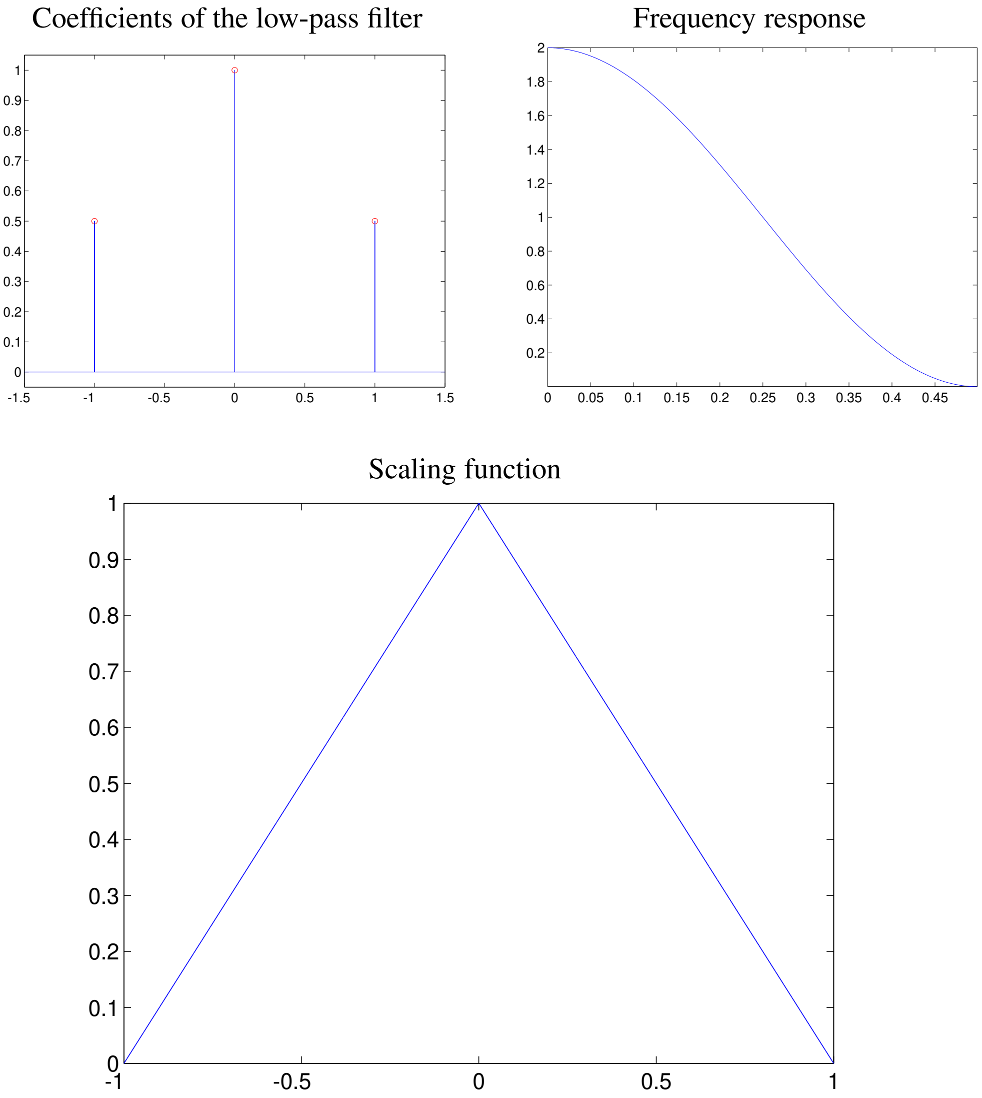

Fig. 2. Synthesis LeGall JPEG filter of length 3 and corresponding scaling function (i.e., β^1, B-spline of degree 1).

indicating that the basis functions are very nearly orthogonal. The cosine is also very close to one, reflecting the fact that the system essentially behaves like an orthogonal projector. A nice surprise is that the approximation properties of the $9/7$ wavelets are better than those of $\mathrm{Daubechies}_8$, albeit not quite as dramatically as in the previous case. Here, the sampling gain over $\mathrm{Daubechies}_8$ is $\gamma = \frac{1}{\sqrt[4]{0.389}} \approx 1.27$. The plot of the Fourier approximation kernel in Fig. 6 also indicates a near 5 dB improvement over a good portion of the frequency axis. The wavelet constant $C_{\tilde{\psi}}$ is also very favorable—close to C_{JPEG}^{-}—reflecting the property that the transform is nearly orthogonal. This is all the more remarkable given the strong design constraints: symmetry and compact support. Thus the $9/7$ filterbank appears to win over $\mathrm{Daubechies}_8$ in all respects.

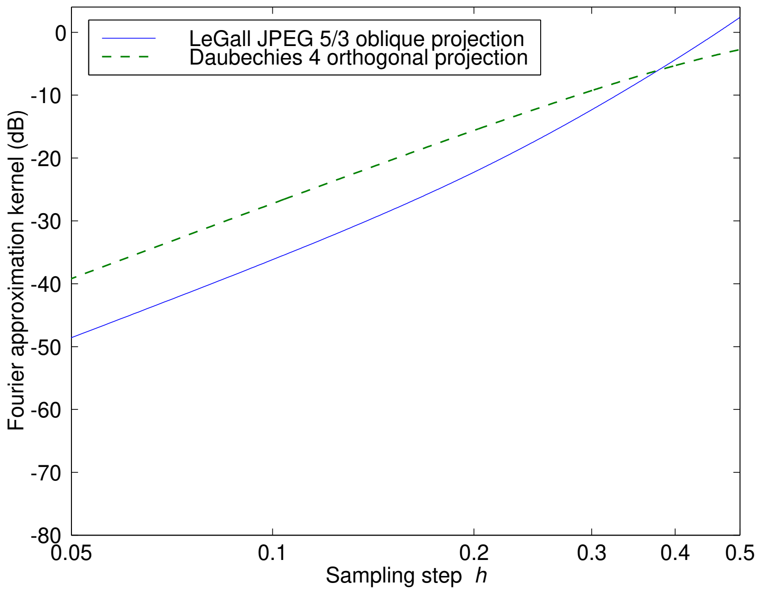

Fig. 3. Fourier approximation kernel corresponding to LeGall JPEG $5/3$, compared to Daubechies of same order.

One should point out, however, that the smoothness and approximation properties of the $9/7$ filterbank are by far not the best achievable for wavelets with four vanishing moments. The top performers in both of these categories are the various brands of cubic spline wavelets which achieve $r = 3$ and offer a sampling gain over $\mathrm{Daubechies}_8$ that is greater than 2 (see [12], [33]). However, these splines wavelets are either infinitely supported—which is perceived as a serious handicap for coding applications— or very far from orthogonal which makes the task of optimal quantization and bit allocation much more difficult [40].

APPENDIX

A. *Link Sobolev regularity—Approximation power*

Let $r < s_c$, the Sobolev critical regularity exponent of f. Then from [33, Thm. 1], we know that

$$\|f - \tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2} \leq \sqrt{\int |\hat{f}(\omega)|^2 E_{\tilde{\varphi},\varphi}(\omega a)\,\frac{\mathrm{d}\omega}{2\pi}} \;+\; K\,a^r\,\|f^{(r)}\|_{\mathbf{L}^2} \tag{18}$$

where K is a constant independent of f and a. Let now $r = \min\{L, s\}$ with $s < s_c$ and $C = \sup_{\omega\in\mathbb{R}} \frac{E_{\tilde{\varphi},\varphi}(\omega)}{|\omega|^{2r}}$; this latter quantity is bounded because $E_{\tilde{\varphi},\varphi}(\omega)$ has a $2L^{\mathrm{th}}$ degree of flatness at the origin as a consequence of the order property. Then we have

$$\begin{aligned}
\|f - \tilde{\mathcal{P}}_a f\|_{\mathbf{L}^2} &\le a^r \sqrt{\int |\omega^r \hat{f}(\omega)|^2 \, \frac{E_{\tilde{\varphi},\varphi}(\omega a)}{|\omega a|^{2r}} \, \frac{\mathrm{d}\omega}{2\pi}} \; + \; K\, a^r \, \|f^{(r)}\|_{\mathbf{L}^2} \\
&\le C\, a^r \sqrt{\int |\omega^r \hat{f}(\omega)|^2 \, \frac{\mathrm{d}\omega}{2\pi}} \; + \; K\, a^r \, \|f^{(r)}\|_{\mathbf{L}^2} \\
&\le (C + K)\, a^r \, \|f^{(r)}\|_{\mathbf{L}^2}.
\end{aligned}$$

■

B. Asymptotics of single scale wavelet expansions

Let us define the general linear approximation operator at scale a

$$\forall f \in \mathbf{L}^2, \;\; \mathcal{Q}_a f(x) = \sum_{k \in \mathbb{Z}} \langle f(a\cdot), \varphi_1(\cdot - k) \rangle \, \varphi_2\Big(\frac{x}{a} - k\Big), \tag{19}$$

which is associated with the pair of analysis/synthesis functions (φ_1, φ_2)—not necessarily biorthogonal.

By using the exact same proof as for the derivation of (18), we establish the formula

$$\|\mathcal{Q}_a f\|_{\mathbf{L}^2} = \sqrt{\int |\hat{f}(\omega)|^2 F_{\varphi_1,\varphi_2}(\omega a) \, \frac{\mathrm{d}\omega}{2\pi}} \; + \; o(a^s) \tag{20}$$

where $F_{\varphi_1,\varphi_2}(\omega) = |\hat{\varphi}_1(\omega)|^2 a_{\varphi_2}(\omega)$. This result holds for any pair (φ_1, φ_2); in particular, for $\varphi_1 = \tilde{\psi}$ and $\varphi_2 = \psi$, which corresponds to the wavelet projector at scale $a = 2^i$. The crucial step is then to compute the Taylor series of $F_{\tilde{\psi},\psi}(\omega a) = |\hat{\tilde{\psi}}(\omega a)|^2 a_\psi(\omega a)$ around the origin, which yields

$$F_{\tilde{\psi},\psi}(\omega a) = (a\omega)^{2L} \Big| \frac{\hat{\tilde{\psi}}^{(L)}(0)}{L!} \Big|^2 a_\psi(0) \; + \; o(a^{2L}) \tag{21}$$

where we have used the fact that $\tilde{\psi}$ has L vanishing moments. By computing the autocorrelation of the wavelet relation for ψ, we show that $a_\psi(0) = a_\varphi(\pi)$. Next, we consider the Fourier domain version of the wavelet relation: $\hat{\tilde{\psi}}(\omega) = \frac{\tilde{G}(e^{j\omega/2})}{H(1)} \hat{\tilde{\varphi}}(\omega/2)$. Based on the fact that $\tilde{G}(e^{j\omega}) = e^{j\omega} H(-e^{j\omega})$ has an Lth order zero at $\omega = 0$, we evaluate $\hat{\tilde{\psi}}^{(L)}(0)$, which gives

$$F_{\tilde{\psi},\psi}(\omega a) = (a\omega)^{2L} \Big| \frac{1}{L!\, 2^L} \frac{H^{(L)}(-1)}{H(1)} \Big|^2 a_\varphi(\pi) \; + \; o(a^{2L}).$$

Using (20) where we let $a = 2^i$, we can thus claim that

$$\begin{aligned}
\Big\| \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle \psi_{i,k} \Big\|_{\mathbf{L}^2} &= \frac{\sqrt{a_\varphi(\pi)}}{L!\, 2^L} \Big| \frac{H^{(L)}(-1)}{H(1)} \Big| \, 2^{iL} \, \|f^{(L)}\|_{\mathbf{L}^2} \; + \; o(2^{iL}) \\
&= \sqrt{1 - 4^{-L}} \, C_\varphi^- \, 2^{iL} \, \|f^{(L)}\|_{\mathbf{L}^2}, \quad \text{as } i \to -\infty.
\end{aligned}$$

To determine the ℓ^2-norm of the wavelet coefficients, we consider another approximation operator of the form (19) with $\varphi_1 = \tilde{\psi}$ and $\varphi_2 = \psi_o$ where ψ_o is some other wavelet that is chosen to be orthonormal. The asymptotic form of the corresponding $F_{\tilde{\psi},\psi_o}(\omega a)$ is also given by (21) with $a_{\psi_o}(\omega) = 1$. Since the new basis is orthonormal, we have a perfect norm equivalence, and we can use the same argument as before to show that

$$\left(\sum_{k\in\mathbb{Z}} |\langle f, \tilde{\psi}_{i,k}\rangle|^2\right)^{\frac{1}{2}} = \underbrace{\frac{|\hat{\tilde{\psi}}^{(L)}(0)|}{L!}}_{C_{\tilde{\psi}}} 2^{iL} \, \|f^{(L)}\|_{\mathbf{L}^2}, \quad \text{as } i \to -\infty.$$

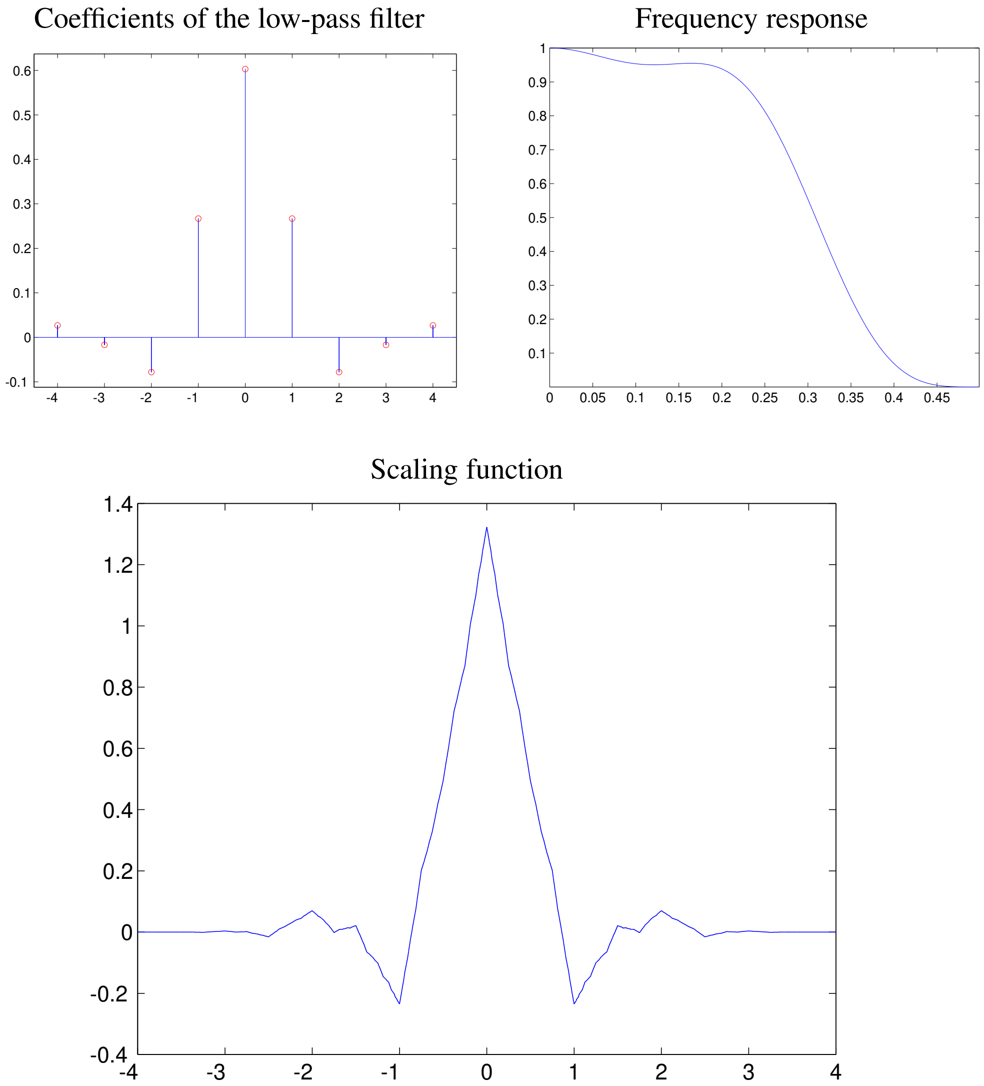

Fig. 4. Analysis Daubechies JPEG filter of length 9 and corresponding scaling function.

REFERENCES

[1] ISO/IEC JTC1 SC29/WG1, "ITU-T recommendation T.800—ISO/IEC FCD15444-1: JPEG2000 image coding system," Tech. Rep., International Organization for Standardization, March 2000.

[2] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3445–3462, December 1993.

[3] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Sys. for Video Tech.*, vol. 6, no. 3, pp. 243–250, 1996.

[4] D. Taubman and A. Zakhor, "A multi-start algorithm for signal adaptive subband systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, San Franscisco, CA, 1992, vol. 3, pp. 213–216.

[5] D. Taubman, "High performance scalable image compression with EBCOTT," *IEEE Trans. Image Process.*, vol. 9, no. 7, pp. 1158–1170, 2000.

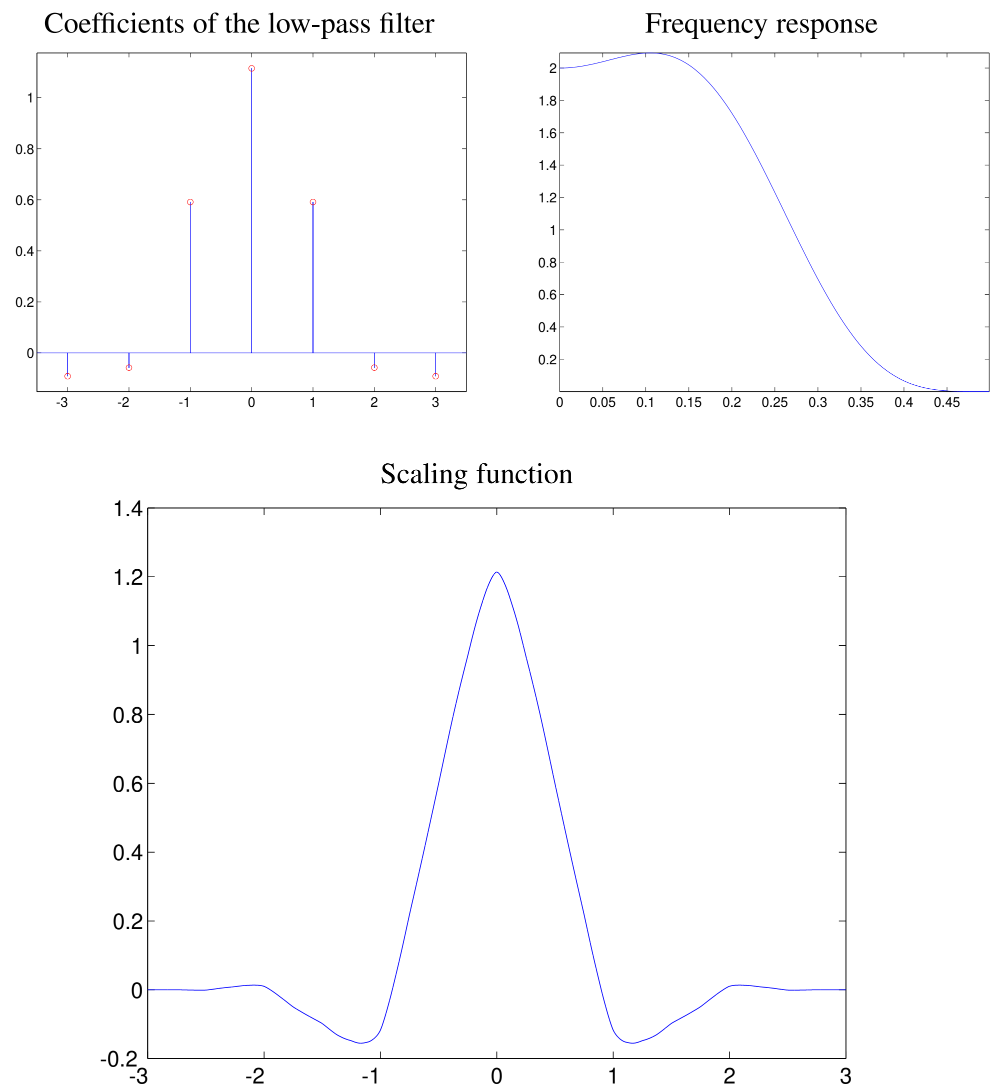

Fig. 5. Synthesis Daubechies JPEG filter of length 7 and corresponding scaling function.

[6] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: An overview," *IEEE Trans. Consumer Electronics*, vol. 46, no. 4, pp. 1103–1127, 2000.

[7] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Process.*, vol. 1, no. 2, pp. 205–220, April 1992.

[8] D. LeGall and A. Tabatabai, "Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques," in *Int. Conf. on Acoustic, Speech and Signal Processing*, New York, NY, 1988, pp. 761–765, IEEE.

[9] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Appl. Comput. Harmon. Anal.*, vol. 3, no. 2, pp. 186–200, 1996.

[10] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. XLI, pp. 909–996, November 1988.

[11] A. Cohen, I. Daubechies, and J.C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Comm. Pure Appl.*

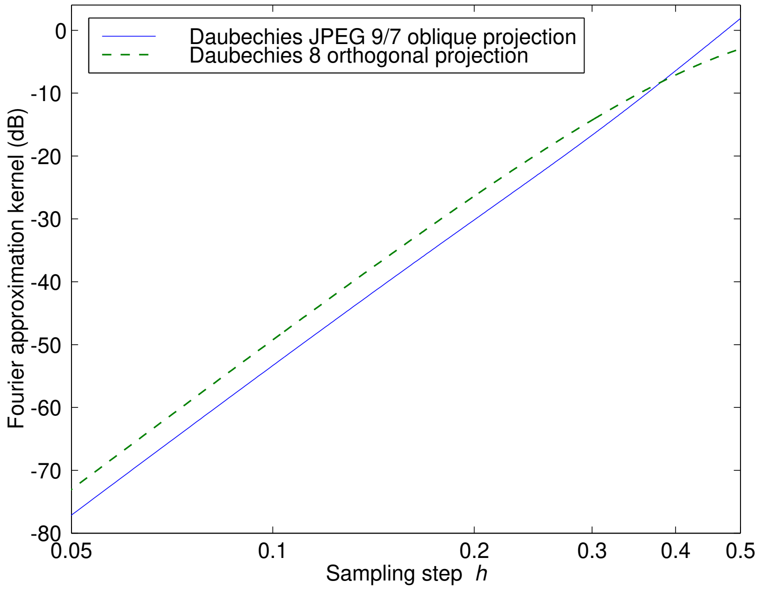

Fig. 6. Fourier approximation kernel corresponding to Daubechies JPEG $9/7$, compared to Daubechies of same order.

Math., vol. 45, no. 5, pp. 485–560, 1992.

[12] M. Unser, "Approximation power of biorthogonal wavelet expansions," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 519–527, March 1996.

[13] R.A. DeVore, B. Jawerth, and B.J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Information Theory*, vol. 38, no. 2, pp. 719–746, 1992.

[14] S.G. Mallat and F. Falzon, "Analysis of low bit rate image transform coding," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 1027–1042, 1998.

[15] D.L. Donoho, M. Vetterli, R.A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Trans. Inform. Th.*, vol. 44, no. 6, pp. 2435–2476, 1998.

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

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

[18] P.P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE*, vol. 78, no. 1, pp. 56–93, January 1990.

[19] G. Strang and T.Q. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Cambridge MA, 1996.

[20] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.

[21] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[22] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspects of Functional Analysis*, G. Geymonat, Ed., pp. 793–840. Edizioni Cremonese, Rome, 1973.

[23] L. Villemoes, "Wavelet analysis of refinement equations," *SIAM J. Math. Anal.*, vol. 25, no. 5, pp. 1433–1460, September 1994.

[24] I. Daubechies and J. Lagarias, "Two-scale difference equations I. Existence and global regularity of solutions," *SIAM J.*

Math. Anal., vol. 22, no. 5, pp. 1388–1410, September 1991.

[25] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544–1576, November 1992.

[26] O. Rioul, "On the choice of wavelet filters for still image compression," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Minneapolis, MN, April 1993, vol. V, pp. 550–553.

[27] J.D. Villasenor, B. Belzer, and J. Liao, "Wavelet filter evaluation for image compression," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1053–1060, August 1995.

[28] Y. Meyer, *Ondelettes*, Hermann, Paris, France, 1990, In French.

[29] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2915–2925, November 1994.

[30] M. Unser and J. Zerubia, "Generalized sampling: Stability and performance analysis," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 2941–2950, 1997.

[31] A. Aldroubi, P. Abry, and M. Unser, "Construction of biorthogonal wavelets starting from any two given multiresolution analyses," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 1130–1133, 1998.

[32] T. Blu and M. Unser, "Approximation error for quasi-interpolators and (multi-) wavelet expansions," *Appl. Comput. Harmon. Anal.*, vol. 6, no. 2, pp. 219–251, March 1999.

[33] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2783–2795, October 1999.

[34] S. Jaffard, "Pointwise smoothness, two-microlocalization and wavelet coefficients," *Publicacions Matemàtiques*, vol. 35, pp. 155–168, 1991.

[35] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part II—Wavelets," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2796–2806, October 1999.

[36] R.A. DeVore, "Nonlinear approximation," *Acta Numerica*, vol. 7, pp. 51–150, 1998.

[37] D.L. Donoho, "Unconditional bases are optimal bases for data compression and estimation," *Appl. Comput. Harmon. Anal.*, vol. 1, no. 1, pp. 100–105, 1993.

[38] I. Daubechies and J. Lagarias, "Two-scale difference equations II. Local regularity, infinite products of matrices and fractals," *SIAM J. Math. Anal.*, vol. 23, no. 4, pp. 1031–1079, July 1992.

[39] G. Strang, "Eigenvalues of $(\downarrow 2)H$ and convergence of the cascade algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 233 – 238, 1996.

[40] P. Moulin, "A multiscale relaxation algorithm for SNR maximization in nonorthogonal subband coding," *IEEE Trans. Image Process.*, vol. 4, no. 9, pp. 1269–1281, 1995.

TABLE I

MATHEMATICAL DATA FOR LEGALL JPEG $5/3$ FILTERS

	Analysis	Synthesis
Approximation order	2	2
Hölder regularity ($\dot{C}^{r_c}$)	0	1
Sobolev regularity ($\mathbf{W}_2^{s_c}$)	0.44	1.5
Approximation constant $\frac{C^-_{\mathrm{JPEG}}}{C^-_{\mathrm{Daubechies}_4}}$	-	$\boxed{1/3}$
Wavelet constant $\frac{C_{\tilde{\psi}}}{C^-_{\mathrm{JPEG}}}$	-	1.677
Riesz bounds	$1 < 1.944$	$0.577 < 1$
Projection cosinus	0.85736	

TABLE II

MATHEMATICAL DATA FOR DAUBECHIES JPEG $9/7$ FILTERS

	Analysis	Synthesis
Approximation order	4	4
Hölder regularity ($\dot{C}^{r_c}$)	1.07	1.70
Sobolev regularity ($\mathbf{W}_2^{s_c}$)	1.41	2.12
Approximation constant $\frac{C^-_{\mathrm{JPEG}}}{C^-_{\mathrm{Daubechies}_8}}$	-	$\boxed{0.389}$
Wavelet constant $\frac{C_{\tilde{\psi}}}{C^-_{\mathrm{JPEG}}}$	-	1.058
Riesz bounds	$0.926 < 1.065$	$0.943 < 1.084$
Projection cosinus	0.98387	