# Quantitative Fourier Analysis of Approximation Techniques: Part II—Wavelets

Thierry Blu, *Member, IEEE*, and Michael Unser, *Fellow, IEEE*

***Abstract*—In a previous paper, we proposed a general Fourier method that provides an accurate prediction of the approximation error, irrespective of the scaling properties of the approximating functions. Here, we apply our results when these functions satisfy the usual two-scale relation encountered in dyadic multiresolution analysis. As a consequence of this additional constraint, the quantities introduced in our previous paper can be computed explicitly as a function of the refinement filter. This is, in particular, true for the asymptotic expansion of the approximation error for biorthonormal wavelets as the scale tends to zero.**

**One of the contributions of this paper is the computation of sharp, asymptotically optimal upper bounds for the least-squares approximation error. Another contribution is the application of these results to B-splines and Daubechies scaling functions, which yields explicit asymptotic developments and upper bounds. Thanks to these explicit expressions, we can quantify the improvement that can be obtained by using B-splines instead of Daubechies wavelets. In other words, we can use a coarser spline sampling and achieve the same reconstruction accuracy as Daubechies: Specifically, we show that this sampling gain converges to $\pi$ as the order tends to infinity.**

## I. Introduction

The notion of order is at the heart of wavelet theory. The standard requirement for a wavelet transform of order $L$ if that the refinement filter $H(z)$ on the synthesis side has a built-in factor $(1+z^{-1})^L$ [1]–[3]. For filter designers, this imposes a multiple-zeros constraint at $\omega = \pi$, which is the only property that distinguishes wavelet filters from the more conventional perfect reconstruction filterbanks [4], [5]. This order constraint has some remarkable consequences, such as the vanishing moments of the analysis wavelet, the ability of the scaling function to reproduce polynomials of degree $n \leq L-1$ (approximation property), and the special eigenstructure of the two-scale transition operator [3, ch. 7]. The order has also a strong influence on the smoothness of the underlying basis functions: Most wavelet families exhibit a regularity index $r$ that is roughly proportional to $L$ (typically, $r = \alpha L$ with $\alpha < 1$). It is therefore quite natural to index common families of wavelets (Daubechies [6], orthogonal splines [7], [8], semi-orthogonal splines [9], [10], biorthonormal splines [11], coiflets, etc.) by the order parameter $L$.

The other remarkable consequence of the order constraint is that the residual error of a scale-truncated wavelet expansion will decrease like the $L$th power of that scale [12]–[15]. This is a basic property well known in approximation theory (Strang–Fix conditions) [16]. To make the connection explicit, we observe that a scale-truncated biorthonormal wavelet expansion can also be expressed as a single scale expansion using the scaling functions at the next finer scale

$$s_i = \sum_{j>i}\sum_{k\in Z}\langle s, \tilde{\psi}_{j,k}\rangle \psi_{j,k} = \sum_{k\in Z}\langle s, \tilde{\varphi}_{i,k}\rangle \varphi_{i,k} \qquad (1)$$

with the short-form notation $\varphi_{j,k} = 2^{-j/2}\varphi(2^{-j}x-k)$. This suggests that we can estimate the wavelet approximation error $\|s - s_i\|_{\mathbf{L}^2}$ by studying the properties of the projector operator

$$\tilde{\mathcal{P}}_T s(x) = \sum_n \left[\int s(\xi)\tilde{\varphi}\left(\frac{\xi}{T}-n\right) d\frac{\xi}{T}\right]\varphi\left(\frac{x}{T}-n\right) \qquad (2)$$

simply because $s_i = \tilde{\mathcal{P}}_{2^i} s$. This problem therefore clearly falls into the general framework of the companion paper [17], except that the present situation is more constrained: The functions $\varphi$ and $\tilde{\varphi}$ are biorthonormal and both satisfy a two-scale relation. From what is known in approximation theory, we would expect higher order wavelets to provide better approximations of piecewise smooth functions, at least in the $\mathbf{L}^2$-sense. The price to pay is that higher order basis functions tend to be less localized: They require more computations and can induce Gibbs-like oscillations around sharp signal transitions. Those limitations notwithstanding, it is of great interest to compare wavelet transforms from the point of view of their approximation properties.

This kind of investigation was initiated by Sweldens *et al.* [13], [14]. They derived some upper bound constants for the asymptotic error and used them to compare various wavelet transforms. They also proved that the asymptotic error depends on the order properties of the primary representation space only (synthesis) and not on how the complementary wavelet spaces are chosen [13]. Their main conclusion was that spline wavelets (irrespective of their kind) had by far the best approximation properties. The main problem with Sweldens' analysis was its complexity and its lack of numerical efficiency, mainly because it was entirely done with wavelets [i.e., the $\tilde{\psi}$-expression in (1)]. Some progress was achieved by reformulating the problem, as has just been done above, and studying the error behavior of the more general projection operator $\tilde{\mathcal{P}}_T$ [15]. This work resulted in an exact asymptotic error formula as well as a computational method for obtaining the leading constant in the wavelet case.

Thanks to the general results that have been presented in our companion paper [17], we are now in the position to go further. The main problem with wavelets, however, is

Manuscript received August 5, 1998; revised March 21, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Xiang-Gen Xia.

The authors are with the Biomedical Imaging Group, EPFL—Swiss Federal Institute of Technology, Lausanne, Switzerland (e-mail: thierry.blu@epfl.ch; michael.unser@epfl.ch).

Publisher Item Identifier S 1053-587X(99)07566-2.

that they generally have no closed-form representation, which implies that we cannot simply apply the error formulæ as they appear in [17]. While it is conceivable to evaluate the approximation kernel $E(\omega)$ (cf. Section II–C) using infinite products, obtaining the various derivatives and extrema that are required by the theory is less obvious. The purpose of this paper is to show how we can circumvent this difficulty and use the two-scale relation to our full advantage to derive much simpler equations for the bound constants. This leads to much more direct and explicit wavelet computations than those found in the general theory. For instance, we will be able to exhibit closed-form error formulæ for spline and Daubechies wavelets of any order $L$.

The paper is organized as follows. In Section II, we start with a brief review of the approximation results that were presented in [17]. All these results make use of a new quantity—the approximation kernel—which was defined in [18]. In Section III, we review some not-so-well known results in wavelet theory. These will be needed to develop computational solutions for the exact evaluation of some of the basic quantities (inner products, moments) that are required by our formulation. In Section IV, we enunciate theorems that provide the asymptotic development of the error as well as some upper bounds. Finally, we illustrate the theory in Section V, with the useful examples of splines and Daubechies wavelets and provide general formulæ for all quantities of interest.

In this respect, our most notable finding concerns the superiority of splines over Daubechies wavelets: Asymptotically, when the approximation order tends to infinity, we obtain the same approximation quality as the latter by using spline wavelets of the same order with a sampling step that is *coarser by a factor* $\pi$ exactly.

### A. Notations

The notations are the same as in [17]. We recall them below for the sake of self consistency.

The conventional inner product $\int s_1(x)s_2(x)\,dx$ between two $\mathbf{L}^2$ functions $s_1$, $s_2$ is denoted $\langle s_1, s_2\rangle$, and the associated Euclidean norm is $\|\cdot\|_{\mathbf{L}^2}$.

The Fourier transform of $s(x)$ is $\hat{s}(\omega)$. Let $r$ be a positive real number; the Sobolev space $\mathbf{W}_2^r$ is defined as the collection of functions satisfying $\int(1+\omega^2)^r|\hat{s}(\omega)|^2\,d\omega < \infty$. In line with this definition of regularity, we extend $\|s^{(r)}\|_{\mathbf{L}^2}$ to noninteger values of $r$ by equating it to the square root of $\frac{1}{2\pi}\int|\omega|^{2r}|\hat{s}(\omega)|^2\,d\omega$. The smoothness of a function $s(x)$ can thus be characterized by the maximum $r$ such that $s \in \mathbf{W}_2^r$; this regularity exponent $r_{\max}$ indicates that $s(x)$ has $\lfloor r\rfloor$ derivatives in $\mathbf{L}^2$ for all $r < r_{\max}$. There is also a direct connection with *pointwise* smoothness: If $s \in \mathbf{W}_2^r$ with $r > \frac{1}{2}$, then $s(x)$ has at least $\lfloor r - \frac{1}{2}\rfloor$ continuous derivatives [19].

The Riemann zeta function is defined as $\zeta(s) = \sum_{n\geq 1} n^{-s}$.

Discrete filters are either described by their impulse response $h_n$ (lowercase letters) or by their $z$-transform $H(z) = \sum_n h_n z^{-n}$ (uppercase letters).

Most of the asymptotic expansions are presented with "$o(\cdot)$" and "$O(\cdot)$" terms, which allows a more compact and understandable form to the results. Writing $f(x) = o(x^n)$ is equivalent to writing $\limsup_{x\to 0}|f(x)/x^n| = 0$; in the same spirit, writing $f(x) = O(x^n)$ is equivalent to $\limsup_{x\to 0}|f(x)/x^n| < \infty$ (not necessarily 0).

## II. Sampling, Approximation, and Interpolation

In this section, we summarize the results obtained in the companion paper [17]; refer either to this paper or to [18] for the more technical developments and proofs.

### A. The Approximation Scheme

We want to approximate a given $\mathbf{L}^2$ function by a linear combination of uniformly shifted functions $\varphi(.-n)$ at a given scale $T$. One of the most general approximating operators $\mathcal{Q}_T$ satisfying these conditions takes the form

$$\mathcal{Q}_T s(x) = \sum_n c_n \varphi\left(\frac{x}{T} - n\right) \tag{3}$$

with

$$c_n = \int s(\xi)\tilde{\varphi}\left(\frac{\xi}{T} - n\right) d\frac{\xi}{T} \tag{4}$$

where $\tilde{\varphi}$ is a distribution that we shall term "sampling distribution" or "sampling function." It can be shown [18] that this expression has a meaning (i.e., it converges in a stable way toward an $\mathbf{L}^2$ function) when the following standard hypotheses are made on the functions involved.

1) $\{\varphi(.-n)\}_{n\in\mathbb{Z}}$ is a Riesz basis, i.e., there exists two positive finite constants $B \geq A > 0$ such that

$$A\|c\|_{\ell^2} \leq \left\| \sum_n c_n \varphi(.-n) \right\|_{\mathbf{L}^2} \leq B\|c\|_{\ell^2}. \tag{5}$$

2) $\sup_{\omega\in\mathbb{R}} |\hat{\tilde{\varphi}}(\omega)| < \infty$, which includes the Dirac delta distribution but not its derivatives.
3) $s \in \mathbf{W}_2^r$, where $r > \frac{1}{2}$ (this implies that $s$ is continuous but not much more).

Although there is no *a priori* requirement on the support of $\varphi$, we will consider only compactly supported functions because they are easier to handle in the multiresolution framework.

It is known [20] that (5) is satisfied iff $A \leq \hat{a}_\varphi(\omega) \leq B$ for almost every $\omega \in [-\pi, \pi]$, where $\hat{a}_\varphi$ is the Fourier transform of the autocorrelation sequence $\{\langle\varphi(x), \varphi(x-n)\rangle\}_{n\in\mathbb{Z}}$

$$\hat{a}_\varphi(\omega) = \sum_n \langle\varphi(x), \varphi(x-n)\rangle e^{-in\omega} = \sum_n |\hat{\varphi}(\omega + 2n\pi)|^2. \tag{6}$$

This function will play a central role in our argument; it has the nice feature of being easy to compute for compactly supported refinable wavelets (see Subsection III–C).

In this paper, we consider the $\mathbf{L}^2$ approximation error $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$. In addition, we will essentially limit our investigations to biorthonormal schemes, i.e., $\langle\tilde{\varphi}(x-k), \varphi(x-l)\rangle = \delta_{k-l}$ for all $k, l \in \mathbb{Z}$. Although the *quasibiorthonormal* framework considered in [18] is much more general, and although we shall briefly deal with quasiinterpolation (see Section IV–C), the present restriction contains most cases of interest for wavelets.

- oblique (or biorthonormal) projection $\mathcal{Q}_T = \tilde{\mathcal{P}}_T$ [11], [21], [15], [22]. This includes the most general wavelets generated by perfect reconstruction filter banks.
- least-squares approximation $\mathcal{Q}_T = \mathcal{P}_T$. That is, orthogonal projection, for which $\tilde{\varphi} = \tilde{\varphi}_d$ such that

$$\hat{\varphi}_d(\omega) = \frac{\hat{\varphi}(\omega)}{\hat{a}_\varphi(\omega)}. \tag{7}$$

  Using this function—also called "dual"—on the analysis side yields the smallest $\mathbf{L}^2$ approximation error. This corresponds to the case of orthogonal [6], [23] and semi-orthogonal wavelets [9], [10].
- interpolation $\mathcal{Q}_T = \mathcal{I}_T$. That is, with the property that $\mathcal{I}_T s(kT) = s(kT)$ by choosing $\hat{\tilde{\varphi}} = \hat{\varphi}_I$, where

$$\hat{\varphi}_I(\omega) = \frac{1}{\sum_n \hat{\varphi}(\omega + 2n\pi)^*} = \frac{1}{\hat{b}_\varphi(\omega)^*}. \tag{8}$$

  This is the inverse of a digital FIR filter since $\hat{b}_\varphi(\omega) = \sum_n \varphi(n) e^{-ni\omega}$. A particular case is the spline interpolator, which is investigated in [24] and [25]. Computing such an interpolation provides a consistent way of initializing the wavelet transform at the finer resolution level.

### B. Approximation Order and Strang–Fix Theory

A crucial notion in approximation theory is the order of approximation describing the rate of decay of $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$ as the sampling step goes to zero. The fundamental result in this area, due to Strang and Fix [16], is that the minimum error (i.e., for $\tilde{\varphi} = \varphi_d$) has an $L$th-order decay $\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \propto T^L$ if and only if

$$\hat{\varphi}(0) \neq 0 \quad \text{and} \quad \hat{\varphi}^{(k)}(2n\pi) = 0 \quad \text{for} \begin{cases} n \neq 0 \\ k = 0 \ldots L-1 \end{cases}. \tag{9}$$

This powerful equivalence was initially proven for compactly supported $\varphi$; it also holds with fewer restrictions (e.g., polynomial decay, multiple generators) [18], [26], [27].

### C. Approximation Theorem

The main result emphasized in [17] is that the approximation error $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$ can be evaluated in a very accurate way by computing expression

$$\eta_s(T) = \left[\frac{1}{2\pi}\int |\hat{s}(\omega)|^2 E(T\omega)\, d\omega\right]^{\frac{1}{2}} \tag{10}$$

where $E(\omega)$ is a kernel that depends only on $\varphi$ and $\tilde{\varphi}$

$$E(\omega) = |1 - \hat{\tilde{\varphi}}(\omega)^* \hat{\varphi}(\omega)|^2 + |\hat{\tilde{\varphi}}(\omega)|^2 \sum_{n\neq 0} |\hat{\varphi}(\omega + 2n\pi)|^2 \tag{11}$$

$$= \underbrace{1 - \frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}}_{E_{\min}(\omega)} + \underbrace{\hat{a}_\varphi(\omega)|\hat{\tilde{\varphi}}(\omega) - \hat{\varphi}_d(\omega)|^2}_{E_{\rm res}(\omega)}. \tag{12}$$

Obviously, $E_{\min}(\omega)$ is the main expression to consider since it is the least-squares approximation kernel.

The theorem proved in [18] states that $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} = \eta_s(T) + o(T^r)$ whenever $s$ belongs to $\mathbf{W}_2^r$. Some other nice features (average theorem, stochastic theorem [17]) lend confidence in the choice of $\eta_s$ as a faithful estimate of the true approximation error.

Specific applications of this theorem lead to asymptotic expansions and sharp upper bounds of the approximation error as a function of the sampling step $T$. We now show how these theoretical results can be efficiently exploited when the functions $\tilde{\varphi}$ and $\varphi$ satisfy a two-scale relation.

## III. Dyadic Multiresolution Analysis

Our goal in this section is to present selected results on dyadic wavelets that will be needed by our determinations. In particular, we will show how to compute the key quantities required by our formulation (autocorrelation, moments, etc.); however, we will not expose the general multiresolution theory of the wavelet transform. Instead, see [1]–[4] and [28] for a clear and complete exposition.

### A. Two-Scale Difference Equation

A dyadic "father" wavelet (or "scaling function") $\varphi$ satisfies a linear equation relating its values at a given resolution to its values at twice the same resolution [1], [6], [19], [28], [29]

$$\varphi(x) = \sum_k h_k \varphi(2x - k) \tag{13}$$

and is such that $\int \varphi = 1$. Most of the time, the considered filter $h_k$ is FIR, which implies that the solution of (13) is compactly supported: This is our assumption throughout the paper.

Another aspect of (13) is its Fourier equivalent

$$\hat{\varphi}(\omega) = \frac{H(e^{i\frac{\omega}{2}})}{2}\hat{\varphi}\left(\frac{\omega}{2}\right) = \prod_{j\geq 1} \frac{H(e^{i\frac{\omega}{2^j}})}{2}. \tag{14}$$

The convergence of the infinite product on the right-hand side is ensured whenever $H(1) = 2$ [30]; this assumption is always made in practice.

Most often, this kind of implicit equation defines a nonregular function, that is, a function or even a distribution that does not belong to $\mathbf{L}^2$. It is also known that the factor $1 + z^{-1}$ is narrowly linked to the regularity [19], [30], [31]. To be more precise, $\frac{(1+z^{-1})}{2}H(z)$ defines a smoother scaling function, which is the difference $\int_{-\infty}^{x} \varphi(\xi)\,d\xi - \int_{-\infty}^{x-1} \varphi(\xi)\,d\xi$ between the integral of $\varphi$ and its translate. Thus, we can build regular functions by multiplying an initial filter by a sufficient number of regularity factors $\frac{(1+z^{-1})}{2}$. From now on, we shall assume that at least one such factor divides $H(z)$.

### B. Strang–Fix Conditions

The Strang–Fix conditions (9) carry over mechanically from the function $\varphi$ to its generating polynomial $H(z)$. Since $\varphi$ satisfies the Riesz condition (5), we have the following property:

$$\left.\begin{array}{l}\varphi \text{ satisfies the Strang-Fix}\\ \text{conditions of order } L\end{array}\right\}$$
$$\Leftrightarrow \begin{cases} H(z) \text{ is divisible by } (1+z^{-1})^L \\ \text{and} \quad H(1)=2. \end{cases} \tag{15}$$

This property can, for example, be obtained as a special case of [18, Lemma 3] (see also [32] and [33]).

### C. Interpolation Property

The two-scale difference equation (13) implies that it is possible to compute iteratively the value of $\varphi$ at any dyadic points once the samples at the integers are known. Specifically, we have

$$\varphi\left(\frac{n}{2^j}\right) = \sum_k g_{j,k}\varphi(n-k) \tag{16}$$

where $g_{j,n}$ is a discrete sequence defined by induction on $j$: $g_{0,n} = \delta_n$ and $g_{j+1,n} = \sum_k g_{n-2k} g_{j,k}$ (see [19] for a detailed analysis). We now show that it is possible to compute $\varphi(n)$ *exactly* from the FIR filter $H$ since the problem can be written as a linear system of equations [19], [30], [31].

Let $x = n$ in (13), and let $n$ take every integer value on the support of $\varphi$. We can rewrite (see [19], [30], and [31]) the set of corresponding equations into the matrix form

$$\mathrm{Y} = \mathbf{H}\mathrm{Y} \tag{17}$$

where $\mathbf{H}$ is a square matrix with coefficients $\mathbf{H}_{k,l} = h_{2k-l}$, and where $\mathrm{Y}_n = \varphi(n)$. If $H(z) = h_M z^{-M} + \cdots + h_N z^{-N}$, then it is easy to show that the support of $\varphi$ is included in $[M, N]$ so that the indices of the matrix $\mathbf{H}$ run from $M+1$ to $N-1$ (if $\varphi$ is continuous, $\varphi(M) = \varphi(N) = 0$). To ensure the normalization $\int \varphi = 1$, we must add another equation to this system, namely, $\sum_n \varphi(n) = 1$. In vector notation, this becomes $U^t Y = 1$, where $U = (1, 1, \ldots, 1)^t$.

Note that even if we do not compute the exact value of $\varphi(n)$, we can use the canonical sequence $g_{j,n}$ to approximate $\varphi(\frac{n}{2^j})$ since this discrete sequence always converges when $\varphi$ is Hölder continuous. This convergence is exponential as shown in [19].

Once $\varphi(n)$ has been computed, we build the polynomial $B(z) = \sum_n \varphi(n) z^n$, which is linked to the interpolation filter $\hat{b}_\varphi$ in (8) through $\hat{b}_\varphi(\omega) = B(e^{i\omega})$. In fact, the matrix equation (17) is equivalent to the polynomial two-scale relation

$$2B(z^2) = H(z)B(z) + H(-z)B(-z). \tag{18}$$

The same technique also works for the exact computation of the autocorrelation filter $\hat{a}_\varphi(\omega)$ by taking note that this filter is built out of $|\hat{\varphi}(\omega)|^2$ (instead of $\hat{\varphi}(\omega)$ for the interpolation filter) and that the corresponding autocorrelation function is solution of a two-scale equation generated by $\frac{1}{2}H(z)H(z^{-1})$ [3].

### D. Moments of $\varphi$

It is also possible to compute exactly the moments of $\varphi$—or, equivalently, the MacLaurin development of $\hat{\varphi}$—from the coefficients of $H$; this is a rather pleasing result, given that we cannot know $\varphi$ exactly at most real points (the case of B-splines is a noteworthy exception) [3, p. 396]. For this purpose, we assume that $\hat{\varphi}$ and $\frac{1}{2}H(e^{i\omega})$ have the MacLaurin development

$$\hat{\varphi}(\omega) = \sum_{k=0}^{K} \alpha_k \omega^k + o(\omega^K) \tag{19}$$

and

$$\frac{H(e^{i\omega})}{2} = \sum_{k=0}^{K} \beta_k \omega^k + o(\omega^K) \tag{20}$$

where we know that $\alpha_0 = 1$ (because $\int \varphi = 1$). Using (14), we end up with the induction equation

$$\alpha_n = \frac{1}{2^n - 1} \sum_{k=0}^{n-1} \beta_{n-k} \alpha_k \tag{21}$$

for $n \geq 1$. The induction on $n$ provides a method for computing the desired moments.

## IV. Approximation Results for Wavelets

We now resume our study of the approximation error and apply the results of [17] to the special case of dyadic wavelets. We consider the most general FIR filters such that $\varphi$ satisfies the Strang–Fix conditions of order $L$. The examples of B-splines and Daubechies wavelets will be treated in Section V.

Here, we will provide two main results. The first one expresses the asymptotic form of the approximation error for biorthonormal wavelets, and the second gives upper bounds for (semi-)orthonormal wavelets (i.e., least-squares approximation). These bounds are not only sharp but also asymptotically optimal when the sampling step tends to zero. They are also easy to compute, especially for orthonormal refinement filters.

It is, in fact, the two-scale relation (13) that makes it practicable to derive results that depend on the generating filter $H(z)$ only. This is not obvious *a priori*, given that the only explicit form of the functions defined by (13) is an infinite product. Moreover, (13) allows the exact computation of quasi-interpolating prefilters (needed to initialize a wavelet transform) or of the shift-invariance error [17] that characterizes the approximation space. Finally, we show how dyadic schemes converge toward their limit functions and how the rate of convergence can be improved by using adequate approximation functions. These results are straightforward by-products of the main approximation theorems [17, Th. 1 and Th. 2]: The extended range of the problems solved shows the power of this approach.

### A. Asymptotic Development of the Error

We gave in [18] the explicit form of the asymptotic development of the least-squares error as $T \to 0$. Here, we refine this result and provide the explicit development of the biorthonormal projection error $\|s - \tilde{\mathcal{P}}_T s\|_{\mathbf{L}^2}$ when the analysis function $\tilde{\varphi}$ is of order $\tilde{L} \leq L$.

*Theorem 1:* Assume that $s$ and $s^{(2L)}$ are in $\mathbf{L}^2$ and that the scaling functions $(\varphi, \tilde{\varphi})$ are generated by two FIR biorthonormal filters $(H(z), \tilde{H}(z))$. Moreover, assume that $(\varphi, \tilde{\varphi})$ are of order $L$ and $\tilde{L}$, respectively, with $\tilde{L} \geq \frac{L}{3}$. We then have the asymptotic development

$$\|s - \tilde{\mathcal{P}}_T s\|_{\mathbf{L}^2}^2 = \sum_{k=L}^{2L-1} \frac{\gamma_k}{4^k - 1} \|s^{(k)}\|_{\mathbf{L}^2}^2 T^{2k} + \sum_{k=L+\tilde{L}}^{2L-1} \lambda_k \|s^{(k)}\|_{\mathbf{L}^2}^2 T^{2k} + O(T^{4L}) \tag{22}$$

where $\gamma_k$ and $\theta_k$ are defined by the MacLaurin developments

$$\overbrace{\frac{\hat{a}_\varphi(\omega+\pi)}{\hat{a}_\varphi(2\omega)} \left| \frac{H(-e^{i\omega})}{2} \right|^2}^{\rho(\omega)} = \sum_{k\geq 0} \gamma_k \omega^{2k} \tag{23}$$

$$\frac{H(-e^{i\omega})^* \tilde{H}(-e^{i\omega})}{4} = \sum_{k \geq 0} \theta_k \omega^k \tag{24}$$

and where $\lambda_k$ are determined by $\theta_k$ as in

$$\lambda_k = \sum_{n=0}^{2k} \frac{\theta_{2k-n} \theta_n^*}{(2^{2k-n} - 1)(2^n - 1)}. \tag{25}$$

Note that the order hypotheses imply that $\gamma_k = 0$ for all $k < L$ and that $\theta_k = \lambda_k = 0$ for all $k < L + \tilde{L}$. If $\tilde{L} < \frac{L}{3}$, then (22) still holds with a remainder of lower power, namely, $O(T^{3(L+\tilde{L})})$.

*Proof:* We start with (12). We have shown in [18] how to develop the first term $E_{\min}$ of the kernel into power series up to order $4L$. The result provides the first sum of the right-hand side of (22). The second term of (12), namely, $E_{\text{res}}$, can be rewritten as $\frac{1}{1-E_{\min}} |1 - \hat{\tilde{\varphi}} \hat{\varphi}^* - E_{\min}|^2$. Since $\varphi$ and $\tilde{\varphi}$ are biorthonormal, $1 - \hat{\tilde{\varphi}} \hat{\varphi}^* = \sum_{n \neq 0} \hat{\tilde{\varphi}}(.-2n\pi) \hat{\varphi}(.-2n\pi)^*$, which is thus $O(\omega^{L+\tilde{L}})$. In addition, we know that $E_{\min} = O(\omega^{2L})$. Thus, we have $E_{\text{res}} = |1 - \hat{\tilde{\varphi}} \hat{\varphi}^*|^2 + O(\omega^{4L})$ if $\tilde{L} \leq L$.

On the other hand, using (14), we find a two-scale induction relation satisfied by $\hat{d}_\varphi = 1 - \hat{\tilde{\varphi}} \hat{\varphi}^*$

$$\begin{aligned} \hat{d}_\varphi(2\omega) &= \hat{d}_\varphi(\omega) + (1 - \hat{d}_\varphi(\omega)) \left( 1 - \frac{\tilde{H}(e^{i\omega}) H(e^{i\omega})^*}{4} \right) \\ &= \hat{d}_\varphi(\omega) + (1 - \hat{d}_\varphi(\omega)) \frac{\tilde{H}(-e^{i\omega}) H(-e^{i\omega})^*}{4} \\ &= \hat{d}_\varphi(\omega) + \frac{\tilde{H}(-e^{i\omega}) H(-e^{i\omega})^*}{4} + O(\omega^{2(L+\tilde{L})}). \end{aligned}$$

Thus, if $\hat{d}_\varphi(\omega) = \sum_{k \geq 0} \theta_k' \omega^k$, then $(2^k - 1)\theta_k' = \theta_k$ for $k = 0 \ldots 2(L + \tilde{L}) - 1$, according to (24). We must restrict the development of $|\hat{d}_\varphi|^2$ to the powers of $\omega$ that are less than $4L$ since this is the highest power of the development of $E_{\min}$ given in [18]. Thus, two cases should be considered: Either, $\tilde{L} < \frac{L}{3}$, which ensures that the accuracy of the development of $|\hat{d}_\varphi|^2$ (namely, $3(L + \tilde{L})$) is less than $4L$, or $\tilde{L} \geq \frac{L}{3}$, and then we should consider only the first $4L$ coefficients of the development of $|\hat{d}_\varphi|^2$. In both cases, we have to develop the expression $\frac{1}{2\pi} \int |\hat{s}(\omega)|^2 |\hat{d}_\varphi(\omega T)|^2 \, d\omega$ up to the corresponding power, which ultimately leads to (22). □

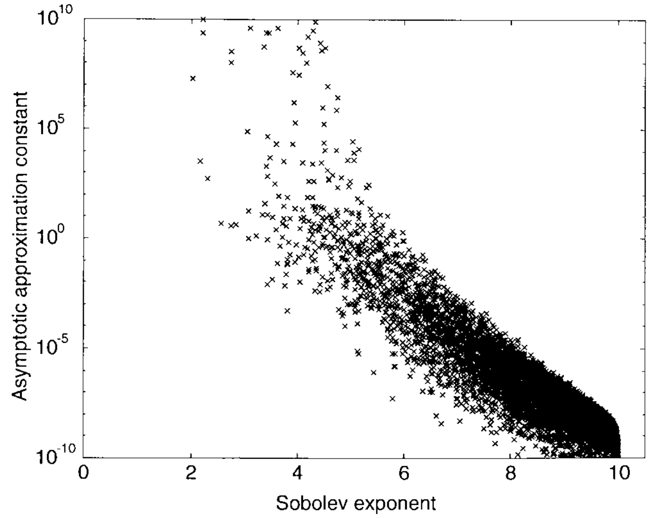


Fig. 1. Scatterplot of $C_\varphi^-$ (log-scale) as a functionof the Sobolev regularity exponent: A total of 10 000 randomfilters $H$ of degree 19, having approximation order is $L = 10$, have been computed.

Given a filter $H(z)$, it is obviously possible to compute the coefficients $(\theta_k, \lambda_k)$ exactly. This is also true for $\gamma_k$ since $\hat{a}_\varphi(\omega)$ can be known as well (see Section III-C).

For the least-squares error case where $\tilde{\varphi} = \varphi_d$, we clearly have $\tilde{L} = L$. Thus, the asymptotic development reduces to the first summation in (22), which was the result initially reported in [18]. Equation (22) also confirms the result of Sweldens *et al.* [14], who proved that the expansions of the biorthonormal and the orthonormal error should match up to (not including) the power $T^{2(L+\tilde{L})}$; however, these authors did not give the explicit form of the developments.

An essential contribution of (22) is that we now have an explicit form for the the first order equivalent $C_\varphi^-$ defined in [15]

$$C_\varphi^- = \frac{|Q(-1)| \sqrt{\hat{a}_\varphi(\pi)}}{2^{L+1} \sqrt{4^L - 1}} \tag{26}$$

where $H(z) = [\frac{z^{-1}+1}{2}]^L Q(z)$. Using this simplified formula, we illustrate the link between the Sobolev regularity "$r$" and $C_\varphi^-$ that was surmised in [15]. For this purpose, we have estimated both the constant and the regularity exponent (using the exact formula given in [3] and [19]) for a large—10 000—set of randomly designed filters that have ten zeros at $z = -1$. As can be seen in Fig. 1, the first-order asymptotic constant decreases roughly exponentially as the Sobolev regularity increases.

*First-Order Interpolation Error:* The technique used in the proof of Theorem 1 can also be exploited in the interpolation case. We show below how to obtain the first order of the corresponding error. First, we rewrite the error kernel from (11) and put it under the form

$$E(\omega) = \left| 1 - \frac{\hat{\varphi}(\omega)}{\hat{b}_\varphi(\omega)} \right|^2 + \frac{\hat{a}_\varphi(\omega)}{|\hat{b}_\varphi(\omega)|^2} E_{\min}(\omega)$$

where $\hat{b}_\varphi$ is the interpolation filter (8). We know from above that $E_{\min}(\omega) = (C_\varphi^-)^2\omega^{2L}$, where $C_\varphi^-$ is given by (26). Thus, we can concentrate on the first term of $E$, which is denoted $|u(\omega)|^2$. Using both the Fourier two-scale relation (14) and (18), we have

$$u(2\omega) = 1 - \frac{\hat{\varphi}(2\omega)}{\hat{b}_\varphi(2\omega)} \tag{27}$$

$$= u(\omega) + \frac{\hat{\varphi}(\omega)}{\hat{b}_\varphi(\omega)} \frac{H(-e^{i\omega})\hat{b}_\varphi(\omega+\pi)}{2\hat{b}_\varphi(2\omega)}. \tag{28}$$

We know that $u(\omega) = O(\omega^L)$, and we want to find $c$ such that $u(\omega) = c\omega^L + O(\omega^{L+1})$. Thus, $2^L c\omega^L = c\omega^L + (i\frac{\omega}{2})^L \frac{Q(-1)}{2}\hat{b}_\varphi(\pi)$ at the first order in $\omega$. After some rearrangements, we finally get

$$\|s - \mathcal{I}_T s\|_{\mathbf{L}^2} = \underbrace{C_\varphi^- \sqrt{1 + \frac{2^L+1}{2^L-1}\frac{|\hat{b}_\varphi(\pi)|^2}{\hat{a}_\varphi(\pi)}}}_{C_\varphi^{\text{int}}} T^L \|s^L\|_{\mathbf{L}^2} + O(T^{L+1}). \tag{29}$$

Therefore, $C_\varphi^{\text{int}} = C_\varphi^-$ if and only if $\hat{b}_\varphi(\pi) = 0$; we insist that this is true only when $\varphi$ satisfies a two-scale equation. For the centered B-splines of even order $L$ (i.e., of odd degree), we check that $|\hat{b}_\varphi(\pi)| = (\frac{2}{\pi})^L 2\zeta(L)(1-2^{-L})$, and thus, using (29), we get $C_\varphi^{\text{int}} = C_\varphi^- \sqrt{1 + \frac{2\zeta(L)^2}{\zeta(2L)}}$, which provides an explicit relation for the constant introduced in [34]. This formula is valid for even order splines only. Note that, asymptotically, when $L \to \infty$, we have $C_\varphi^{\text{int}} = \sqrt{3}C_\varphi^-$.

### B. Upper Bound

A general expression for upper bounds that are asymptotically optimal (up to a given order) was given in [17]. When $\varphi$ satisfies a two-scale equation, we show that we can give an explicit expression of the constants involved in the general formula. We first apply the technique to the lower order upper bound calculation; then, we show how to obtain higher orders. Asymptotically optimal upper bounds follow.

*First-Order Upper Bound:* The first-order bound given in [17] takes the following form:

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} \le \underbrace{\left[ \overbrace{\sup_{|\omega|<\pi} \frac{E(\omega)}{\omega^{2L}}}^{(C_\varphi^0)^2} + \frac{\zeta(2L)}{\pi^{2L}} \right]^{\frac{1}{2}}}_{C_\varphi} \left\|s^{(L)}\right\|_{\mathbf{L}^2} T^L. \tag{30}$$

Moreover, the smallest constant $C_\varphi$ that satisfies (30) is necessarily greater than $C_\varphi^0$, which is the minimum constant over the subclass of Nyquist-bandlimited functions (cf. [17, Sec. IV–B]). This implies that

$$C_\varphi^0 T^L \le \sup_{s\in\mathbf{L}^2} \frac{\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}}{\left\|s^{(L)}\right\|_{\mathbf{L}^2}}.$$

TABLE I
WAVELET UPPER BOUNDS AND SHIFT-INVARIANCE ERROR

| order | Daubechies | | | splines | | |
|---|---|---|---|---|---|---|
| $L$ | $C_\varphi$ | $\frac{C_\varphi}{C_\varphi^0}$ | $\frac{\sigma_\varphi}{\|\varphi\|_{\mathbf{L}^2}}$ | $C_\varphi$ | $\frac{C_\varphi}{C_\varphi^0}$ | $\frac{\sigma_\varphi}{\|\varphi\|_{\mathbf{L}^2}}$ |
| 1 | 0.5 | 57% | 0.577 | 0.5 | 57% | 0.577 |
| 2 | 0.154 | 72% | 0.446 | 0.129 | 56% | 0.273 |
| 3 | 0.0595 | 83% | 0.392 | 0.0399 | 57% | 0.159 |
| 4 | 0.0254 | 91% | 0.360 | 0.0126 | 57% | 0.0959 |
| 5 | 0.0115 | 95% | 0.339 | 0.00401 | 57% | 0.0585 |
| 6 | 0.00536 | 98% | 0.323 | 0.00128 | 57% | 0.036 |
| 7 | 0.00256 | 99% | 0.31 | 0.000406 | 57% | 0.0222 |
| 8 | 0.00123 | 99% | 0.3 | 0.00013 | 57% | 0.0137 |

As a result, the upper bound in (30) is all the more sharp as the second term in the argument of the square root is smaller with respect to the first one.

In order to estimate the upper value of $\omega^{-2L}E(\omega)$ for $\omega \in [-\pi, \pi]$, we use (13) and the two-scale induction equation $4\hat{a}_\varphi(2\omega) = |H(e^{i\omega}|^2\hat{a}_\varphi(\omega) + |H(-e^{i\omega}|^2\hat{a}_\varphi(\omega+\pi)$. After some algebra, we obtain $E(2\omega) - E(\omega) = \rho(\omega)(1 - E(\omega))$, where $\rho(\omega)$ is given by (23); this implies that $E(2\omega) \le E(\omega) + \rho(\omega)$ since $E(\omega)$ is positive. Then, we define $M = \sup_{|\omega|\le\frac{\pi}{2}} \frac{\sqrt{\rho(\omega)}}{|\omega|^L}$ and get by induction $E(\omega) \le \frac{M^2}{4^L-1}\omega^{2L}$. Finally, (30) yields

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \le \left[\frac{M^2}{4^L-1} + \frac{\zeta(2L)}{\pi^{2L}}\right]^{\frac{1}{2}} \left\|s^{(L)}\right\|_{\mathbf{L}^2} T^L. \tag{31}$$

This new bound, and the "sharpness percentage" $\frac{C_\varphi^0}{C_\varphi}$, have been computed for Daubechies and B-spline wavelets (see Table I). A sharpness percentage of 100% indicates that there are signals for which the inequality (31) is an equality; lower values of this percentage indicate that (31) is strict as well as indicating how tight it is. We observe in particular that the bounds for Daubechies wavelets are very sharp and that their sharpness tends to 100% when the approximation order increases. By contrast, the bounds for spline wavelets are less sharp, and their sharpness depends only loosely on the order.

*Higher Order Bounds:* Higher orders can be obtained by bounding the remainder of the MacLaurin series of $E(\omega)$ (see [17, (28)]). The use of the two-scale relation also yields an explicit form of this sharpened bound.

*Theorem 2:* Assume that $0 \le N \le L$; then, we have the upper bound for the least-squares approximation error by dyadic wavelets

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \le \left|\sum_{k=L}^{L+N-1} \frac{\gamma_k}{4^k-1}\left\|s^{(k)}\right\|_{\mathbf{L}^2}^2 T^{2k}\right|^{\frac{1}{2}} + \left[\frac{M_N^2}{4^{L+N}-1} + \frac{\zeta(2L+2N)}{\pi^{2L+2N}}\right]^{\frac{1}{2}} \left\|s^{(L+N)}\right\|_{\mathbf{L}^2} T^{L+N} \tag{32}$$

TABLE II
WAVELET QUASI-INTERPOLANT PREFILTERS

| order $L$ | splines |
|---|---|
| 1 | $\frac{1}{2}z+\frac{1}{2}$ |
| 2 | $-\frac{1}{12}z^2+\frac{7}{6}z-\frac{1}{12}$ |
| 3 | $-\frac{1}{8}z^3+\frac{5}{8}z^2+\frac{5}{8}z-\frac{1}{8}$ |
| 4 | $7/240z^4-17/60z^3+181/120z^2-17/60z+7/240$ |

| order $L$ | Daubechies |
|---|---|
| 1 | $\frac{1}{2}z+\frac{1}{2}$ |
| 2 | $-0.116z^2+0.866z+0.25$ |
| 3 | $-0.06199z^2+0.928743z+0.145885-0.012638z^{-1}$ |
| 4 | $0.0094259z^4-0.0567535z^3+0.1164163z^2+0.9051175z+0.0257938$ |

where $M_N$ is a constant obtained by

$$M_N^2 = \left| \sup_{|\omega|\le\frac{\pi}{2}} \frac{\rho(\omega)-\sum_{k=L}^{L+N-1}\gamma_k\omega^{2k}}{\omega^{2L+2N}} \right| \tag{33}$$

and where $(\rho(\omega), \gamma_k)$ have the same definition as in (23).

The proof follows the same steps as for the first order and is a direct consequence of the two-scale relation (13). Notice that (32) contains (31) if we let $N=0$.

The interest of this result is that (32) and the true approximation error have the *same asymptotic development* in the neighborhood of $T=0$ up to the power $T^{L+N}$. We can thus say that (32) is asymptotically optimal up to the order $N$; this provides a new series of unreported bounds with optimal characteristics.

### C. Initialization Filters for the Wavelet Transform

Unless the scaling function already satisfies a so-called quasi-interpolation property [35], it is advisable to use a prefilter to compute the fine scale expansion coefficient in (3), given the discrete values of the function $s(nT)$. Some examples of wavelet prefiltering algorithms are discussed in [34] and [36].

The problem amounts to finding a prefilter $P(e^{i\omega})=\hat{\tilde{\varphi}}(\omega)$ such that $\|s-\mathcal{Q}_T s\|_{\mathbf{L}^2}^2$ has the same asymptotic development as $\|s-\mathcal{P}_T s\|_{\mathbf{L}^2}^2$ when $T$ tends to zero, up to the order $T^{2N}$. This problem has been explicitly solved in [17, (36)], where it was shown that it is equivalent to choose $P$ such that $P(e^{i\omega})=\hat{\varphi}_d(\omega)+O(\omega^N)$. Since we can compute exactly $\hat{a}_\varphi$ (see Section III-C) and since we have the exact value of the moments of $\varphi$ (see Section III-D), we can express the exact MacLaurin development of $P(e^{i\omega})$ up to the order $N$. As noticed in [17], this condition is satisfied by an infinite collection of filters having various length and delay.

In Table II, we give quasi-interpolant prefilters for (non-centered) B-splines and Daubechies wavelets of various approximation orders $L$: Using these filters as sampling distributions ensures that $\|s-\mathcal{Q}_T s\|_{\mathbf{L}^2}=\|s-\mathcal{P}_T s\|_{\mathbf{L}^2}+O(T^{L+1})$. Moreover, we have chosen the filters so that the resulting approximation kernels are well-behaved away from $\omega=0$.

### D. Shift-Invariance Error

Except for ideal unrealizable functions such as the Nyquist interpolator, the approximation space $V$ generated by integer shifts of $\varphi$ is not globally shift invariant. Thus, if $x_0\notin\mathbb{Z}$, we have $\varphi(x-x_0)\notin V$ in general. However, the results given in [17] allow us to compute the average value over all the possible real shifts $x_0$ of the approximation error when projecting $\varphi_{x_0}(x)=\varphi(x-x_0)$ orthogonally onto $V$. This is a consequence of [17, Th. 2], which yields

$$\begin{aligned}\sigma_\varphi^2 &= \int_0^1 \left\|\varphi_{x_0}-\mathcal{P}_1\varphi_{x_0}\right\|_{\mathbf{L}^2}^2 dx_0 \\ &= \frac{1}{2\pi}\int |\hat{\varphi}(\omega)|^2\left(1-\frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}\right)d\omega.\end{aligned}$$

Here, we show that this equation can be rewritten in a form that is suitable for exact computation. Specifically, if we let $\hat{c}_\varphi(\omega)=\sum_n|\hat{\varphi}(\omega+2n\pi)|^4$, and then

$$\sigma_\varphi^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\hat{a}_\varphi(\omega)-\frac{\hat{c}_\varphi(\omega)}{\hat{a}_\varphi(\omega)}\right)d\omega. \tag{34}$$

Note that as for $\hat{a}_\varphi(\omega)$, $\hat{c}_\varphi(\omega)$ can be computed exactly; it corresponds to the integer samples of the dyadic wavelet generated by the filter $\frac{1}{8}H(z)^2H(z^{-1})^2$ (see Section III-C). Let us denote by $A(z)$ and $C(z)$ the FIR filters defined by $A(e^{-i\omega})=\hat{a}_\varphi(\omega)$ and $C(e^{-i\omega})=\hat{c}_\varphi(\omega)$, respectively. Then, using the formulation of complex analysis, we have

$$\sigma_\varphi^2 = a_0 - \frac{1}{2i\pi}\oint_{C_{(0,1)}}\frac{C(z)}{zA(z)}\,dz \tag{35}$$

where $C_{(0,1)}$ is the unit circle (in the complex plane) centered on 0. We then know, by Cauchy's residues theorem, that the integral on the right-hand side equals the sum of the residues of $\frac{C(z)}{zA(z)}$ for the poles lying *strictly inside* $C_{(0,1)}$. Of course, a major simplification occurs when $\varphi$ is orthonormal, that is to say, when $\hat{a}_\varphi(\omega)=1$. In that specific case, we simply have $\sigma_\varphi^2=1-c_0$.

In Table I, we give the reduced value $\frac{\sigma_\varphi}{\|\varphi\|_{\mathbf{L}^2}}$ for Daubechies and spline wavelets of order 1 to 8. We observe that the Daubechies ones have a much slower decrease toward 0 than the splines as the order increases. This trend was confirmed by further investigation: We computed $\frac{\sigma_\varphi}{\|\varphi\|_{\mathbf{L}^2}}$ for the first 100 orders and observed that the shift-invariance error seems to decrease as a polynomial in $L$ for Daubechies wavelets (roughly $\frac{\sigma_\varphi}{\|\varphi\|_{\mathbf{L}^2}}\propto L^{-0.45}\log L$) and as an exponential for splines (roughly $\frac{\sigma_\varphi}{\|\varphi\|_{\mathbf{L}^2}}\propto(\frac{2}{\pi})^L$). This indicates that the approximation spaces generated by Daubechies wavelets converges much more slowly than splines to a shift-invariant space, as their order tends to infinity.

### E. Convergence Rate of Dyadic Schemes

The function defined by (13) is, in general, not known exactly everywhere; in particular, irrational values of $x$ can be obtained only through an infinite number of iterations of the interpolation scheme described in Section III-C. If we restrict ourselves to $j$ iterations, the iterated two-scale difference equation reads

$$\varphi(x) = \sum_k g_{j,k}\varphi(2^j x - k)$$

where $g_{j,n}$ is defined as in (16). This form suggests a way to approximate the limit function $\varphi$: We can simply define a sequence of functions $\varphi_j(x)$ as in [37] by

$$\varphi_j(x) = \sum_k g_{j,k}\chi(2^j x - k) \qquad (36)$$

where the compactly supported function $\chi$ is chosen so that the error $\|\varphi - \varphi_j\|_{\mathbf{L}^2}$ tends to zero as $j \to \infty$. Interestingly, the study of this error also falls within the scope of the present paper: It suffices to observe that we can rewrite (36) as $\varphi_j = \mathcal{Q}_{2^{-j}}\varphi$, where $\mathcal{Q}_T$ is defined by the couple of sampling/approximation functions $(\varphi_d, \chi)$: This is a consequence of the equality $g_{j,n} = \int \varphi(x)\varphi_d(x2^j - n)\, dx2^j$.

For our results to apply, we assume that $\varphi \in \mathrm{W}_2^r$ with $r > \frac{1}{2}$. We thus have

$$\|\varphi - \varphi_j\|_{\mathbf{L}^2} = \left[\frac{1}{2\pi}\int |\hat{\varphi}(\omega)|^2 E(2^{-j}\omega)\, d\omega\right]^{1/2} + o(2^{-jr}).$$

Since the second term on the right-hand side does not *a priori* decrease faster than $2^{-jr'}$ if $r' > r$, we cannot hope to have a convergence rate higher than $r$. Moreover, if $\omega^{-2r}E(\omega)$ is bounded, then $\varphi \in \mathrm{W}_2^r$ implies that the first term on the right-hand side is $O(2^{-jr})$. This means that if we choose $\chi$ such that $E(\omega) = O(\omega^{2r})$, then the convergence rate of $\varphi_j$ to $\varphi$ will be at least $2^{-jr}$. When $r > 1$, this is a significant improvement over the usual rate, namely, $2^{-j}$, which would drive the convergence process if we had performed the approximation with a classical step function. From (12), we see that the constraint on $E$ is satisfied iff $\chi$ satisfies the Strang–Fix conditions of order $N = \lceil r \rceil$, and $\hat{\chi}(\omega) = \hat{\varphi}(\omega) + O(\omega^N)$.

A related result appears in [37] and [38] in the (more general) case of $p/q$-adic schemes, i.e., rational up-sampling iterated schemes: It was shown that under the same conditions on $\chi$ as above, the convergence rate *under the* $\mathbf{L}^\infty$ *norm* of the approximating functions $\varphi_j$ is $2^{-jh}$, where $h$ is the Hölder regularity exponent of $\varphi$. Clearly, the $\mathbf{L}^2$ approximation error decreases faster than the $\mathbf{L}^\infty$ error, due to the well-known inequality $h \le r$ [3], [19].

## V. Examples: Splines and Daubechies Wavelets

### A. B-Splines

Non-centered B-splines are piecewise polynomials that satisfy a two-scale relation (13); the generating filter is $H_L(z) = 2^{-L+1}(z^{-1}+1)^L$. We thus have $\hat{\varphi}(\omega) = [\frac{1-e^{-i\omega}}{i\omega}]^L$, which solves the induction equation (14) explicitly. In this case, we can get the exact minimal approximating kernel $E_{\min}(\omega)$. Even though the two-scale technique used in the preceding sections applies to splines as well, we shall, however, consider this exact kernel since it makes it easier to obtain our bounds and asymptotic developments.

*1) Asymptotic Expansions:*

*Theorem 3:* The first $4L$ coefficients of the asymptotic development of the least-squares spline approximation error is given by

$$\|s - \mathcal{P}_T s\|^2_{\mathbf{L}^2} = \sum_{k=0}^{L-1} \frac{2\zeta(2L+2k)}{(2\pi)^{2L+2k}}\binom{2L+2k-1}{2k} \times \left\|s^{(L+k)}\right\|^2_{\mathbf{L}^2} T^{2L+2k} + O(T^{4L}). \qquad (37)$$

*Proof:* Writing down the expression of $E_{\min}(\omega)$, we have

$$E(\omega) = \frac{\omega^{2L}a(\omega)}{1+\omega^{2L}a(\omega)} \qquad (38)$$

where $a(\omega)$ is defined as $\sum_{n\neq 0}(\omega+2n\pi)^{-2L}$. In this sum, we develop each term $(\omega+2k\pi)^{-2L}$ into convergent MacLaurin series and exchange the summations. This yields

$$a(\omega) = \sum_{k\ge 0}\frac{2\zeta(2L+2k)}{(2\pi)^{2L+2k}}\binom{2L+2k-1}{2k}\omega^{2k}. \qquad (39)$$

This development is uniformly convergent for any $\omega \in ]-2\pi, 2\pi[$. The proof now follows easily from (38), which shows that $E(\omega) = \omega^{2L}a(\omega) + O(\omega^{4L})$. □

The first coefficient of this development appeared previously in [15], where it was expressed with a Bernoulli number instead of a zeta function. The next terms of the development are new and provide a finer asymptotic characterization of the expansion error.

*2) Upper Bounds:*

*Theorem 4:* The least-squares polynomial spline approximation error is bounded as follows:

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \le \frac{1}{\pi^L}\sqrt{2\zeta(2L) - \frac{1}{2}}\left\|s^{(L)}\right\|_{\mathbf{L}^2} T^L. \qquad (40)$$

Moreover, we can find asymptotically optimal upper bounds. If $1 \le N \le L$, we have

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \le \left|\sum_{k=0}^{N-1}\frac{2\zeta(2L+2k)}{(2\pi)^{2L+2k}}\binom{2L+2k-1}{2k}\left\|s^{(L+k)}\right\|^2_{\mathbf{L}^2}T^{2L+2k}\right|^{\frac{1}{2}} + \frac{\sqrt{2}}{\pi^{L+N}}\left\|s^{(L+N)}\right\|_{\mathbf{L}^2}T^{L+N}. \qquad (41)$$

*Proof:* We consider again the exact expression (38) of $E(\omega)$. Our first task is to bound $\omega^{-2L}E(\omega)$. Let us define $b(\omega) = a(\omega) - (\omega - 2\pi)^{-2L} = \sum_{n\neq 0,-1}(\omega+2n\pi)^{-2L}$. It is easy to check that $\frac{d}{d\omega}a(\omega)$ is strictly negative for every $\omega \in ]-\pi, \pi[$ and, thus, that $b(\omega)$ is strictly decreasing over this same interval. In particular, if $0 \le \omega \le \pi$, we have $0 < b(\omega) \le b(0) = \frac{2\zeta(2L)-1}{(2\pi)^{2L}}$. Using (38), we find

$$\omega^{-2L}E(\omega) \le \frac{b(\omega) + \frac{1}{(\omega-2\pi)^{2L}}}{1+\frac{\omega^{2L}}{(\omega-2\pi)^{2L}}} \le b(0) + \frac{1}{\omega^{2L}+(\omega-2\pi)^{2L}} \le \frac{1+\frac{2\zeta(2L)-1}{2^{2L-1}}}{2\pi^{2L}}.$$

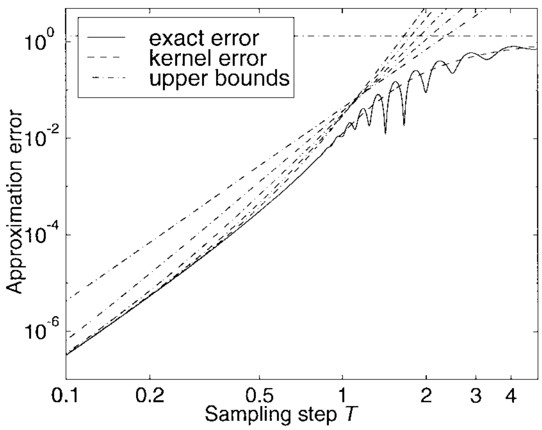


Fig. 2. Least-squares approximation of a Gaussian $s(x) = e^{-\frac{x^2}{2}}$ using cubic splines ($L = 4$), as a function of the sampling step $T$. The solid line is the exact error, whereas the dashed line is the error computed using (10) and involving the kernel (12). The oblique dash-dotted lines are the upper bounds computed with (40), and (41) for $N = 1, 2, 3, 4$. When the order $N$ increases, the bounds get closer to the exact curve for $T \leq 1$. The horizontal dash-dotted line is the trivial upper bound given by $\|s\|_{\mathbf{L}^2}$.

A final simplification occurs if we use the inequality $\frac{2\zeta(2L)-1}{2^{2L}} < \zeta(2L) - 1$. Applying (30), we end with the first-order bound (40).

Finding the asymptotically optimal bounds is hardly different. Let us denote by $a_N(\omega)$ the first $N$ terms of the development (39) of $a(\omega)$. The remainder $r_N(\omega) = \frac{a(\omega)-a_N(\omega)}{\omega^{2N}}$ is strictly positive, as is obvious from (39). Moreover, since this expansion is uniformly convergent over $]-2\pi, 2\pi[$ and since all its coefficients are strictly positive, $r_N$ is a strictly increasing function of $\omega$; thus, $r_N(\omega) \leq r_N(\pi)$. Using (38), these remarks lead to

$$E(\omega) - \omega^{2L} a_N(\omega) \leq \omega^{2L+2N} r_N(\omega) \leq \omega^{2L+2N} r_N(\pi).$$

We have that $\pi^{2N} r_N(\pi) < a(\pi) - a(0)$, where $a(0)$ and $a(\pi)$ can be computed exactly, as in $a(0) = \frac{2\zeta(2L)}{(2\pi)^{2L}}$ (the first-order asymptotic), and $a(\pi) = \frac{1}{\pi^{2L}}[2\zeta(2L)(1-4^{-L})-1]$. Moreover, using the definition of the zeta function, a close inspection shows that $\pi^{2L}(a(\pi) - a(0)) \leq 2 - \zeta(2L + 2N)$. Thus, we finally get (41). □

Note that the simplified form (40) is only slightly less sharp than the bound that would be obtained directly from (31): For example, the relative difference between (40) and (31) is approximately $6 \cdot 10^{-2}, 10^{-2}, 3 \cdot 10^{-3}$ and $7 \cdot 10^{-4}$ for $L = 1, 2, 3, 4$, respectively. As $L$ tends to infinity, this relative difference quickly vanishes.

Fig. 2 shows how sharp our new bounds are in the particular case of the approximation of a Gaussian $e^{-\frac{x^2}{2}}$ by cubic splines (see caption for details). As can be observed, the quality of the bound improves as $N$ increases for sampling steps lesser than 1; in turn, the upper bound worsens for higher values of the sampling step because of the polynomial behavior of (41).

### B. Daubechies Wavelets

The Daubechies filters [6], which we denote by $D_L(z)$, are the shortest orthonormal filters that contain $L$ "regularity" factors $1+z^{-1}$. The product $D_L(z)D_L(z^{-1})$ can be computed exactly by the formula (see Meyer, quoted by [3, p. 168])

$$|D_L(e^{i\omega})|^2 = C_L \int_0^{\pi+\omega} \sin^{2L-1} x \, dx \tag{42}$$

where $C_L = 4^{-L+1} L\binom{2L}{L}$.

*1) Asymptotic Expansions:* We apply here the results of Section IV-A. Since Daubechies wavelets are orthonormal, we simply have $\hat{a}_\varphi(\omega) = 1$ so that $\gamma_k$ are the coefficients of the development of $|D_L(-e^{i\omega})|^2$ up to the order $4L$ as specified by (23). The expression of $|D_L(-e^{i\omega})|^2$ can be obtained by substituting $\omega$ to $\omega + \pi$ in (42).

*Theorem 5:* The first $4L$ coefficients of the asymptotic development of the least-squares approximation error by Daubechies wavelets are given by

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2}^2 = \sum_{k=L}^{2L-1} \frac{C_L d_k^L}{8k(4^k - 1)} \left\|s^{(k)}\right\|_{\mathbf{L}^2}^2 T^{2k} + O(T^{4L}) \tag{43}$$

where the coefficients $d_k^L$ are defined by the MacLaurin development of $\sin^{2L-1}(x)$

$$\sin^{2L-1}(x) = \sum_{k \geq 0} d_k^L x^{2k-1}. \tag{44}$$

*Proof:* Note that $d_k^L = 0$ for all $k < L$. The link between the coefficients $\gamma_k$ of (23) and $d_k$ of (44) is straightforward by using (42). We find that $\gamma_k = \frac{C_L d_k^L}{8k}$. Thus, applying (22), we find (43). □

The coefficients $d_k^L$ are easy to obtain. For example, we have $d_L^L = 1$ and $d_{L+1}^L = -\frac{2L-1}{6}$, which implies that

$$\begin{aligned}\|s - \mathcal{P}_T s\|_{\mathbf{L}^2}^2 = {} & \frac{C_L}{8L(4^L - 1)} \left\|s^{(L)}\right\|_{\mathbf{L}^2}^2 T^{2L} \\ & - \frac{C_L(2L-1)}{48(L+1)(4^{L+1}-1)} \left\|s^{(L+1)}\right\|_{\mathbf{L}^2}^2 T^{2L+2} \\ & + O(T^{2L+4}).\end{aligned}$$

In particular, this expression provides a closed form for the first-order asymptotic constant. This improves the previously reported expression of the first-order asymptotic constant, which was given as the limit of an infinite summation [15]. This exact expression makes it easy to compare theoretically the constants arising in Daubechies and spline wavelet approximations.

*Theorem 6:* The first-order asymptotic constant of Daubechies wavelets is given by

$$C_{\text{Dau}}^{-} = 4^{-L} \sqrt{\frac{\binom{2L}{L}}{2(1 - 4^{-L})}}. \tag{45}$$

Moreover, the asymptotic sampling density required by using Daubechies wavelets tends to be exactly $\pi$ times larger than the splines, as $L \to \infty$, that is, if $T_{\text{Dau}}$ and $T_{\text{spl}}$ are the sampling steps necessary for Daubechies and spline wavelets to provide the same (asymptotical) approximation error, then

$$\lim_{L \to \infty} \frac{T_{\text{Dau}}}{T_{\text{spl}}} = \frac{1}{\pi}. \tag{46}$$

*Proof:* We concentrate on the second part of the theorem since (45) is a direct consequence of Theorem 5.

The relation between $T_{\rm Dau}$ and $T_{\rm spl}$ is obviously $C^-_{\rm Dau}T^L_{\rm Dau} = C^-_{\rm spl}T^L_{\rm spl}$. It follows that $\frac{T_{\rm Dau}}{T_{\rm spl}} = [\frac{C^-_{\rm spl}}{C^-_{\rm Dau}}]^{1/L}$. Using the Stirling formula ($n! \approx \sqrt{2\pi n}n^n e^{-n}$), we find $\binom{2L}{L} \approx \frac{4^L}{\sqrt{\pi L}}$. Thus, $C^-_{\rm Dau} \approx \frac{2^{-L^-}}{\sqrt[4]{4\pi L}}$ as $L \to \infty$. On the other side, $C^-_{\rm spl} \approx \frac{\sqrt{2}}{(2\pi)^L}$ since $\zeta(n) \approx 1$ as $n \to \infty$. The limit result (46) then follows immediately. □

Finding the precise factor $\pi$ in this context is rather unexpected. Our result confirms the previous reports [14], [15] that splines are superior to Daubechies wavelets for the approximation of smooth functions.

*2) Upper Bounds:*

*Theorem 7:* A first-order upper bound of the least-squares approximation error with Daubechies wavelets is

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \le C^-_{\rm Dau}\sqrt{1 + \frac{1}{2}\left[\frac{C^-_{\rm spl}}{C^-_{\rm Dau}}\right]^2}\,\|s^{(L)}\|_{\mathbf{L}^2}T^L. \quad (47)$$

*Proof:* Here we can use the results of Section IV-B. The definition (42) of $D_L$ shows that

$$\rho(\omega) = \frac{C_L}{4}\int_0^\omega \sin^{2L-1} x\, dx.$$

Thus, since $|\sin x| < |x|$, we have $M_0^2 = \frac{C_L}{8L}$, where $M_0$ was defined by (33). The application of (32), and a rearrangement using the value of $C^-_{\rm spl}$, provides (47). □

Since the second term is smaller than the first one, and since it is asymptotically negligible with respect to the first term (as shown by Theorem 6), it appears that the upper bound for Daubechies filters is very close to the asymptotic first order equivalent: This shows how sharp our first bound is and that it is consistent with the indications of sharpness found in Table I.

The higher orders in (32) provide even sharper bounds. However, the constant in (47) is already so close to the optimal asymptotic value that there is not much benefit in refining the bound further.

## VI. Conclusion

In this second paper, we have shown how easy it is to compute various quantities (asymptotics, bounds, shift-invariance error, etc.) linked to the approximation error once we assume the basis functions satisfy a scaling relation. Due to the explicit connection between these quantities and the generating filter, our expressions help to solve optimization problems, such as filter design, whenever optimized approximation characteristics are desired.

The results given here can be extended to the case of arbitrary (integer) scaling factors. The case of fractional scaling factors [39] is currently under investigation by one of us: This study reveals that after some (nontrivial) adaptations, the formulæ take the same qualitative form.

We expect that our investigations in approximation theory will be helpful in signal processing since interpolation (see Shannon's celebrated theorem [40]) is such a common tool here. The advantage of the theorems that have been presented here is that they are valid *irrespective of the amplitude of the sampling step*: This contrasts with most results in approximation theory that concentrate essentially on the limit when the sampling step tends to zero. In signal processing, indeed, the sampling step is most often given in advance and cannot be made smaller.

## Acknowledgment

The authors thank Dr. P. Thévenaz of the Swiss Federal Institute of Technology for his kind and thorough proofreading of this paper.

## References


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

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

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

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

[5] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

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

[7] G. Battle, "A block spin construction of wavelets. Part I: Lemarié functions," *Commun. Math. Phys.*, vol. 110, no. 4, pp. 601–615, 1987.

[8] A. Aldroubi and M. Unser, "Families of multiresolution and wavelet spaces with optimal properties," *Numer. Funct. Anal. Opt.*, vol. 14, nos. 5–6, pp. 417–446, 1993.

[9] M. Unser, A. Aldroubi, and M. Eden, "A family of polynomial spline wavelet transforms," *Signal Process.*, vol. 30, pp. 141–162, 1993.

[10] C. K. Chui and J. Z. Wang, "On compactly supported spline wavelets and a duality principle," *Trans. Amer. Math. Soc.*, vol. 330, no. 2, pp. 903–915, 1992.

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

[12] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Rev.*, vol. 31, pp. 613–627, 1989.

[13] W. Sweldens and R. Piessens, "Quadrature formulæ and asymptotic error expansions for wavelet approximations of smooth functions," *SIAM J. Math. Anal.*, vol. 31, no. 4, pp. 1240–1264, 1994.

[14] ——, "Asymptotic error expansions for wavelet approximations of smooth functions II," *Numer. Math.*, vol. 68, no. 3, pp. 377–401, 1994.

[15] M. Unser, "Approximation power of biorthogonal wavelet expansions," *IEEE Trans. Signal Processing*, vol. 44, pp. 519–527, Mar. 1996.

[16] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspects of Functional Analysis*, Cremonese, Ed. Rome, Italy: 1971, pp. 796–830.

[17] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors," *IEEE Trans. Signal Processing*, this issue, pp. 2783–2795.

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

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

[20] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Funct. Anal. Opt.*, vol. 15, nos. 1–2, pp. 1–21, Feb. 1994.

[21] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 2207–2232, Sept. 1992.

[22] A. Aldroubi, "Oblique projections in atomic spaces," in *Proc. Amer. Math. Soc.*, July 1996, vol. 124, no. 7, pp. 2051–2060.

[23] O. Rioul and P. Duhamel, "A Remez exchange algorithm for orthonormal wavelets," *IEEE Trans. Circuits Syst.*, vol. 41, pp. 550–560, Aug. 1994.

[24] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I—Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821–832, Feb. 1993.

[25] ———, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Trans. Signal Processing*, vol. 41, pp. 834–848, Feb. 1993.
[26] E. W. Cheney and W. A. Light, "Quasiinterpolation with basis functions having noncompact support," *Constr. Approx.*, vol. 8, pp. 35–48, 1992.
[27] C. de Boor, R. A. Devore, and A. Ron, "Approximation from shift invariant subspaces of $\mathbf{L}^2(\mathbb{R}^d)$," *Trans. Amer. Math. Soc.*, vol. 341, no. 2, pp. 787–806, Feb. 1994.
[28] Y. Meyer, *Ondelettes*. Paris, France: Hermann, 1990, in French.
[29] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet decomposition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.
[30] 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, Sept. 1991.
[31] ———, "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.
[32] A. Cohen, I. Daubechies, and G. Plonka, "Regularity of refinable function vectors," *J. Fourier Anal. Appl.*, vol. 3, no. 3, pp. 295–324, 1997.
[33] C. Heil, G. Strang, and V. Strela, "Approximation by translates of refinable functions," *Numer. Math.*, vol. 73, pp. 75–94, 1996.
[34] M. Unser and I. Daubechies, "On the approximation power of convolution-based least-squares versus interpolation," *IEEE Trans. Signal Processing*, vol. 45, pp. 1697–1711, July 1997.
[35] M. Unser, "Quasiorthogonality and quasiprojections," *Appl. Comput. Harmon. Anal.*, vol. 3, pp. 201–214, 1996.
[36] P. Abry and P. Flandrin, "On the initialization of the discrete wavelet transform algorithm," *IEEE Signal Processing Lett.*, vol. 1, pp. 32–34, Feb. 1994.
[37] O. Rioul and T. Blu, "Simple regularity criteria for subdivision schemes II. The rational case," submitted for publication.
[38] T. Blu, "Bancs de filtres itérés en fraction d'octave—Application au Codage de Son," Ph.D. thesis, École Nat. Supérieure Télécommun., Paris, France, 1996, in French.
[39] ———, "Iterated filter banks with rational rate changes—Connection with discrete wavelet transforms," *IEEE Trans. Signal Processing*, vol. 41, pp. 3232–3244, Dec. 1993.
[40] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

**Thierry Blu** (M'96), for a photograph and biography, see this issue, p. 2795.

**Michael Unser** (M'89–SM'94–F'99), for a photograph and biography, see this issue, p. 2795.