QUANTITATIVE $\mathbf{L}^2$ ERROR ANALYSIS FOR INTERPOLATION METHODS AND WAVELET EXPANSIONS

Thierry Blu
France Télécom—CNET DSE/SGV
38–40 rue du Général Leclerc
92794 Issy Moulineaux Cedex 9, France
e-mail:blu@issy.cnet.fr

Michael Unser
Swiss Federal Institute of Technology
EPFL - DMT/IOA
CH-1015 Lausanne, Switzerland
e-mail:michael.unser@epfl.ch

Abstract— **Our goal in this paper is to set a theoretical basis for the comparison of re-sampling and interpolation methods. We consider the general problem of the approximation of an arbitrary continuously-defined function $f(x)$ —not necessarily bandlimited— when we vary the sampling step T. We present an accurate $\mathbf{L}^2$ computation of the induced approximation error as a function of T for a general class of linear approximation operators including interpolation and other kinds of projectors. This new quantitative result provides exact expressions for the asymptotic development of the error as $T \to 0$, and also sharp (asymptotically exact) upper bounds.**

I. Introduction

Re-sampling and interpolation play a central role in image processing. These operations are required to rescale, rotate images or to correct for spatial distortions. Shannon's theory [1] provides an exact sampling/interpolation system for bandlimited signals. However, this method is rarely used in practice because of the slow decay of $\mathrm{sinc}(x)$. Instead, practitioneers rely on more localized methods such as bilinear interpolation, short kernel convolution and polynomial spline interpolation, which are much more efficient to implement, especially in higher dimensions. These methods can all be studied from the general perspective of approximation theory [2], [3], [4], [5], [6]. Here, we are aiming for results that are more quantitative and directly applicable to the comparison of practical interpolation and approximation methods. We will consider N-dimensional signals that belong to the space $\mathbf{L}^2(\mathbb{R}^N)$. We will use the associated $\mathbf{L}^2(\mathbb{R}^N)$ norm to characterize the behavior of the error as a function of the sampling step T. Our computation is based on theoretical results in [7]; the case of refinable functions (wavelets) is treated in more details in [8]. These formulæ may be used to identify the interpolation/approximation methods that give the smallest approximation error for a given computational complexity and/or class of signals. The results should also be useful for comparing the performance of wavelet transforms.

We define the N-Dimensional Fourier transform $\hat{f}(\omega)$ of a function f by the integral $\int f(\xi)e^{-i\omega^{\mathrm{t}}\xi}d^N\xi$.

II. Approximation of functions by wavelets

Classically, approximation techniques are linear operators: this means that the image of the approximation operator is a vector space. We assume that this space is wavelet-like, i.e., that there exists a generating function $\varphi(x)$ in $\mathbf{L}^2(\mathbb{R}^N)$ and a sampling step (or scale parameter [5]) $T > 0$ such that the approximation function belongs to $V_T = \mathrm{span}\{\varphi(\frac{\cdot}{T} - n)\}_{n\in\mathbb{Z}^N}$. The generating function is further assumed to satisfy the Riesz conditions for reasons which are discussed in [7], i.e. there exist two positive constants $B > A > 0$ such that for all $c \in \ell^2(\mathbb{Z}^N)$

$$A\|c\|_{\ell^2(\mathbb{Z}^N)} \leq \Big\| \sum_n c_n\varphi(\cdot - n)\Big\|_{\mathbf{L}^2(\mathbb{R}^N)} \leq B\|c\|_{\ell^2(\mathbb{Z}^N)} \quad (1)$$

The more general case, where the approximating function belongs to a multi-wavelet (or finite elements) space is dealt with in [7].

A. Sampling and interpolation

The approximating operator $\mathcal{Q}_T$ makes use of a sampling operator $\mathcal{S}_T$ which maps the functions f of $\mathbf{L}^2(\mathbb{R}^N)$ to $\ell^2(\mathbb{Z}^N)$ sequences: its form is

$$\mathcal{S}_T : f \longmapsto \left\{ \int f(T\xi)\tilde{\varphi}(\xi - n)\, d^N\xi \right\}_{n\in\mathbb{Z}^N} \quad (2)$$

where $\tilde{\varphi}$ is a distribution and T the sampling step (the sampling frequency is obviously $F = T^{-1}$). Thus, the approximation operator becomes

$$\mathcal{Q}_T : f \longmapsto \sum_n \{\mathcal{S}_T f\}_n \varphi(\frac{\cdot}{T} - n) \quad (3)$$

As it is defined, $\mathcal{S}_T f$ is not ensured to be in $\ell^2(\mathbb{Z}^N)$. We therefore make the additional assumption that $\|\omega\|^r\hat{\varphi}(\omega)$ is square integrable for some $r > \frac{N}{2}$ (the Sobolev regularity exponent of f) and that $\tilde{\varphi}$ is a distribution whose Fourier transform is bounded (such as the Dirac mass), which ensures that the samples are in $\ell^2(\mathbb{Z}^N)$ (cf. [7] for the 1-D case; the proof in N-D is similar). Alltogether, the conditions on φ, $\tilde{\varphi}$ and f ensure that $\mathcal{Q}_T f$ belongs to $\mathbf{L}^2(\mathbb{R}^N)$.

III. An approximation theorem

We shall use the $\mathbf{L}^2$ norm to evaluate the approximation error, and define $\varepsilon_T(f) = \|f - \mathcal{Q}_T f\|_{\mathbf{L}^2(\mathbb{R}^N)}$. One of the fundamental results of [7] (1-D case) is that this error can be computed exactly for bandlimited functions: this extends (for multi-wavelets, and for arbitrary sampling distributions) an already known result for the minimal approximation arror [9]. Our results are even more accurate, providing sharp bounds in the case of arbitrary $\mathbf{L}^2(\mathbb{R}^N)$ functions f (with a slight Sobolev constraint).

Theorem 1: With the abovementioned hypotheses for φ, $\tilde{\varphi}$ and f we have

$$\varepsilon_T(f) = \left\{ \frac{1}{(2\pi)^N} \int |\hat{f}(\omega)|^2 E(T\omega)\, d\omega^N \right\}^{\frac{1}{2}} + e(f,T) \quad (4)$$

where $e(f,T)$ vanishes if f is not aliased when sampled at $F = \frac{1}{T}$. Else, we have

$$|e(f,T)| \leq KT^r \left\| \|\omega\|^r \hat{\varphi}(\omega) \right\|_{\mathbf{L}^2(\mathbb{R}^N)}$$

where r is the Sobolev regularity exponent of f and K a constant which does not depend on f.
In this theorem, the positive error kernel $E(\omega)$ is

$$E(\omega) = 1 - 2\Re\{\overline{\hat{\tilde{\varphi}}(\omega)}\hat{\varphi}(\omega)\} + |\hat{\tilde{\varphi}}(\omega)|^2 A(\omega). \quad (5)$$

where $A(\omega) = \sum_{n \in \mathbb{Z}^N} |\hat{\varphi}(\omega + 2n\pi)|^2$. Note that $E(\omega)$ can be also written as the sum of two terms

$$E(\omega) = \underbrace{1 - \frac{|\hat{\varphi}(\omega)|^2}{A(\omega)}}_{E_0(\omega)} + \underbrace{A(\omega)\left|\hat{\tilde{\varphi}}(\omega) - \frac{\hat{\varphi}(\omega)}{A(\omega)}\right|^2}_{E_1(\omega)} \quad (6)$$

where it clearly appears that $E_0(\omega)$ corresponds to the minimal approximation error (as studied in [9]), when $\tilde{\varphi}$ is chosen to be the dual function of φ. In that case, $\mathcal{Q}_T$ reduces to an orthogonal projector $\mathcal{P}_T$. In the practical cases where $\tilde{\varphi}$ has to be a sum of equally distributed Dirac masses $\delta(x-n)$ (quasi-interpolation), the coefficients can be determined by minimizing the contribution of $E_1(\omega)$ in (4).

Considering asymptotic expansions, (4) provides an exact development up to the regularity order of the function to approximate: this is a consequence of $e(f,T) \propto T^r$; which is more, theorem 1 tells us how to find a sampling distribution $\tilde{\varphi}$ such that the approximation error exhibit asymptotically the same development, up to a given order, as the *minimal* approximation error.

A. Averaging the $\mathbf{L}^2$ error

In general, we want to be able to approximate not only $f(.)$, but also any shifted version $f_u(.) = f(.-u)$. Since this may produce different values of the approximation error, we may be interested in a more "global" measure obtained by averaging the error over all the possible values of the shift parameter u. Since $\varepsilon_T(f_u)$ is 1-periodic in any canonical direction of $\mathbb{R}^N$, we define this average by

$$\eta_T(f)^2 \stackrel{\text{def}}{=} \frac{1}{T^N} \int_{[0,T]^N} \varepsilon_T(f_u)\, d^N u. \quad (7)$$

A remarkable result is that $\eta_T(f)$ can be computed exactly: it reduces to the first term of (4).

Theorem 2: Under the same conditions on φ, $\tilde{\varphi}$ and f as theorem 1, we have

$$\eta_T(f)^2 = \frac{1}{(2\pi)^N} \int |\hat{f}(\omega)|^2 E(T\omega)\, d\omega^N \quad (8)$$

This interpretation of the first term of (4) shows that the second is the variability that can be expected when considering shifted versions of one same function; moreover, the correction term to $\varepsilon_T(f)^2$ cancels on the average.

IV. Refinable wavelets

From now on, we shall restrict ourselves to 1-D functions, and even more, to refinable generating functions (wavelets), which have gained such a wide success in Signal Processing. We thus require that the approximating function $\varphi(x)$ satisfies a two-scale difference equation [10], [11]

$$\varphi(x) = \sum_n g_n \varphi(2x - n) \quad (9)$$

In this case, the space V_T generates a multiresolution analysis [12], [13].

In fact, since the approximation kernel can be separated into two parts E_0 and E_1, we shall only concentrate on E_0 assuming that the sampling distribution has been chosen so as to minimize the approximation error (i.e. $E_1(\omega) = 0$).

The filter $G(z) = \sum_n g_n z^{-n}$ in (9) is supposed here to be FIR and to generate a regular function, which satisfies the Riesz condition (1). In particular, we assume that $(1 + z^{-1})^L$ divides $G(z)$ and define the quotient to be $2^{-L}Q(z)$.

A. Asymptotic development

For f regular enough, we can find [7], [8] the asymptotic development of $\varepsilon_T(f)^2$ as $T \to \infty$ by knowing the development of $E(\omega)$ near 0. In particular, if we concentrate on the first $4L$ coefficients (needing $f^{(2L)} \in \mathbf{L}^2$), the development can easily be computed from the generating filter $G(z)$.

Before giving the expression of the development, we recall that the 2π-periodic function $A(\omega) = \sum_n |\hat{\varphi}(\omega + 2n\pi)|^2$ can be *computed exactly* through the resolution of a linear system of equations [14], [8] due to the existence of the refinement equation (9): this technique will not be detailed here. It will thus be possible to find the exact coefficients of the development of $R(\omega)$ defined by

$$R(\omega) = \frac{A(\omega+\pi)}{4A(2\omega)}\left|G(-e^{i\omega})\right|^2 \tag{10}$$

in the neighborhood of $\omega = 0$; we denote them by a_k, i.e. $R(\omega) = \sum_{k\geq 0} a_k\omega^k$. In fact, the first $2L$ coefficients cancel due to the divisibility condition on $G(z)$; note also that the odd coefficients cancel as well. The error $\varepsilon_T(f)^2$ satisfies the following development

$$\varepsilon_T(f)^2 \approx \sum_{k=L}^{2L-2} \frac{a_k}{4^k-1}\|f^{(k)}\|^2_{\mathbf{L}^2}T^{2k} \tag{11}$$

up to the order $4L-1$ in the neighbourhood of $T=0$. In particular, we can provide a closed expression for the infinite summation [5] yielding the first non vanishing coefficient C_φ^-; specifically $C_\varphi^- = \frac{|Q(-1)|\sqrt{A(\pi)}}{2^{L+1}\sqrt{4^L-1}}$. Note, however, that (11) provides much more, since it gives access to the next asymptotic orders as well.

A.1 B-Splines

In the case of B-splines, generated by $G(z) = 2^{-L+1}(1+z^{-1})^L$, the expression of C_φ^- is already known [5], but not the next asymptotic equivalents. A direct computation [8] shows that

$$\varepsilon_T(f)^2 \approx \sum_{k=L}^{2L-1} \frac{2\zeta(2k)}{(2\pi)^{2k}}\binom{2k-1}{2L-1}\|f^{(k)}\|^2_{\mathbf{L}^2}T^{2k} \tag{12}$$

where $\zeta(s)$ is Riemann's zeta function $\sum_{n\geq 1} n^{-s}$, which can here (s is even) be computed exactly with the help of Bernoulli numbers.

A.2 Daubechies wavelets

Daubechies wavelets [15] are generated by the shortest orthonormal filter $G(z)$ which has L "regularity" factors $1+z^{-1}$. In particular we have $A(\omega)=1$ which simplifies $R(\omega)$. Here, we rely on the development of $\sin^{2L-1}(x)$ in power series, whose coefficients we denote by $d_{L,k}$

$$\sin^{2L-1}(x) = \sum_{k\geq 0} d_{L,k}x^k$$

Additionally we define the constant $C_L = 4^{-L+1}L\binom{2L}{L}$. Then the first $4L$ coefficients of the asymptotic development are given by

$$\varepsilon_T(f)^2 \approx \sum_{k=L}^{2L-2} \frac{C_L d_{L,k}}{8k(4^k-1)}\|f^{(k)}\|^2_{\mathbf{L}^2}T^{2k} \tag{13}$$

The first non-vanishing order is $C_\varphi^- = \left\{\frac{C_L}{8L(4^L-1)}\right\}^{1/2}$ and can be compared to the B-spline first equivalent: an asymptotic study as $L\to\infty$ shows that, in order for the Daubechies wavelets to achieve the same approximation error as the B-splines, it is necessary that the sampling step be π times smaller, which of course increases by the same amount the sampled data. This demonstrates the superiority of B-splines over Daubechies' wavelets for the approximation of smooth functions.

B. Upper bound

Theorem 1 provides a powerful tool for deriving a whole variety of bounds for $\varepsilon_T(f)$[7]. Due to its asymptotic properties, we can go even further and exhibit a new sequence of bounds that are asymptotically optimal up to a given order, as shown in [8]. We shall not give all the sequences, but only some of them.

Recalling the definition (10) of $R(\omega)$, one of our results for $\varepsilon_T(f)$ is

$$\varepsilon_T(f) \leq \sup_{|\omega|\leq\frac{\pi}{2}}\sqrt{\frac{\omega^{-2L}R(\omega)}{4^L-1}+\frac{\zeta(2L)}{\pi^{2L}}}\|f^{(L)}\|_{\mathbf{L}^2}T^L, \tag{14}$$

which gives an explicit value to the constant C_φ that appears in the Strang-Fix error estimate [2].

B.1 Splines

In the case of splines, a more direct computation provides

$$\varepsilon_T(f) \leq \frac{\sqrt{2\zeta(2L)-1/2}}{\pi^L}\|f^{(L)}\|_{\mathbf{L}^2}T^L. \tag{15}$$

We can also be more accurate by adding another term to the right hand side. With the following bound, our result is asymptotically sharp

$$\varepsilon_T(f) \leq \underbrace{\frac{\sqrt{2\zeta(2L)}}{(2\pi)^L}}_{C_\varphi^-}\|f^{(L)}\|_{\mathbf{L}^2}T^L + \frac{\sqrt{2}}{\pi^{L+1}}\|f^{(L+1)}\|_{\mathbf{L}^2}T^{L+1}. \tag{16}$$

This is because the first term on the right hand side is exactly the B-spline asymptotic first order equivalent C_φ^-.

B.2 Daubechies wavelets

Using (14), it can be checked that a first order upper bound can be written as

$$\varepsilon_T(f) \leq \sqrt{\left(C_\varphi^-\right)^2+\frac{\zeta(2L)}{\pi^{2L}}}\|f^{(L)}\|_{\mathbf{L}^2}T^L \tag{17}$$

where $C_\varphi^- = \left\{\frac{C_L}{8L(4^L-1)}\right\}^{1/2}$ is here the first order asymptotic equivalent for Daubechies wavelets. In general, and in particular as L increases, the $\left(C_\varphi^-\right)^2$ term is larger than the second one, so that this first order bound is close to the asymptotic equivalent: this is thus a sharp bound.

V. Conclusion

The consistency of our results can be checked in figures 1 and 2: in both log-log plots, the exact approximation error $\varepsilon_T(f)$ and the first member of the rhs of (4) coincide as soon as $T \lessapprox 1$; above this value, $\varepsilon_T(f)$ oscillates in the expected way, that is to say, on the (squared) average it is identical to the first member of (4) (cf theorem 2). We also check the sharpness of our new upper bounds: in particular, we note how close the Daubechies upper bound (17) is to the first order asymptotic equivalent, and how close these two (easily computable) estimates are to the *exact* approximation error when $T \lessapprox 0.5$.

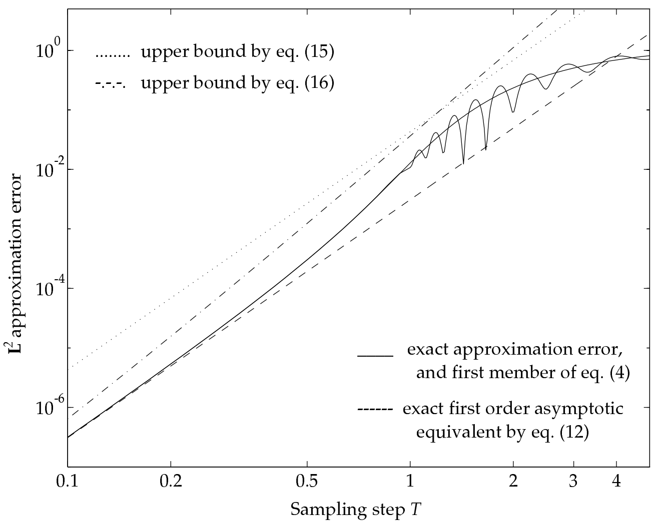

Fig. 1. Results of the paper for the minimal approximation of a gaussian $f(x) = e^{-x^2/2}$ using a B-spline of order 4 (i.e. $L = 4$)

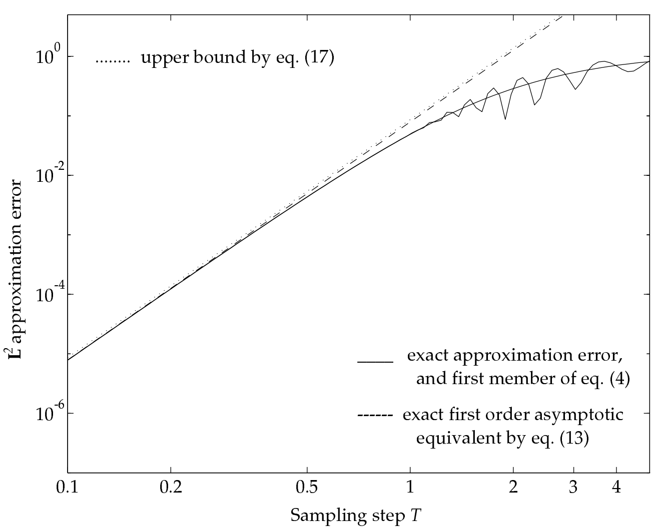

Fig. 2. Results of the paper for the minimal approximation of a gaussian $f(x) = e^{-x^2/2}$ using a Daubechies wavelet of order 4

We expect that the general theory described here, and more generally in [7] may be helpful for the design of optimal interpolation methods under various constraints (restrictions on the size of the support, on the sampling scheme etc...). This has numerous applications in digital signal processing, as well as in the discretization of continuous systems.

References

[1] C.E. Shannon, "Communication in the presence of noise," in *Proc. IRE*, January 1949, vol. 37, pp. 10–21.

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

[3] C. de Boor and G. Fix, "Spline approximation by quasi-interpolants," *J. Approx. Th.*, vol. 8, pp. 19–45, 1973.

[4] C.K. Chui and H. Diamond, "A characterization of multivariate quasi-interpolation formulas and applications," *Numer. Math.*, vol. 57, pp. 105–121, 1990.

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

[6] M. Unser and I. Daubechies, "On the approximation power of convolution-based least squares versus interpolation," *IEEE Trans. Sig. Proc.*, 1996, to appear.

[7] T. Blu and M. Unser, "Approximation error for quasi-interpolators and (multi-) wavelet expansions," *Appl. and Comp. Harm. An.*, 1997, submitted.

[8] T. Blu and M. Unser, "Quantitative $\mathbf{L}^2$ error for wavelet expansions," 1997, in preparation.

[9] 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, February 1994.

[10] 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.

[11] 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.

[12] Y. Meyer, *Ondelettes*, Hermann, Paris, France, 1990, in French.

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

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

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