Quantitative Fourier Analysis of Approximation Techniques: Part I—Interpolators and Projectors

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

***Abstract*— We present a general Fourier-based method that provides an accurate prediction of the approximation error as a function of the sampling step T. Our formalism applies to an extended class of convolution-based signal approximation techniques, which includes interpolation, generalized sampling with prefiltering, and the projectors encountered in wavelet theory. We claim that we can predict the $\mathbf{L}^2$-approximation error by integrating the spectrum of the function to approximate—not necessarily bandlimited—against a frequency kernel $E(\omega)$ that characterizes the approximation operator. This prediction is easier yet more precise than was previously available. Our approach has the remarkable property of providing a global error estimate that is the average of the true approximation error over all possible shifts of the input function. Our error prediction is exact for stationary processes, as well as for bandlimited signals. We apply this method to the comparison of standard interpolation and approximation techniques.**

Our method has interesting implications for approximation theory. In particular, we use our results to obtain some new asymptotic expansions of the error as $T \to 0$, as well as to derive improved upper bounds of the kind found in the Strang–Fix theory. We finally show how we can design quasi-interpolators that are near optimal in the least-squares sense.

I. Introduction

RESAMPLING and interpolation play a central role in image processing [1]–[3]. These operations are required to rescale or rotate images or to correct for spatial distortions. Those are also standard tools in signal processing for performing sampling rate conversions or implementing time delays [4], [5]. Shannon's theory [6] provides an exact sampling/interpolation system for bandlimited signals. However, his method is rarely used in practice—especially for images—because of the slow decay of $\mathrm{sinc}(x)$. Instead, practitioners rely on more localized methods such as bilinear interpolation, short kernel convolution [7], and polynomial spline interpolation [8], [9], which are much more efficient to implement, especially in higher dimensions.

Although interpolation techniques are widely used in practice, it should be realized that they often constitute a rather crude approach to the problem of approximating a function or a signal $s(x)$ that is continuously defined. For instance, it is well understood that interpolation is not appropriate for sampling rate reduction because it usually gives rise to aliasing artifacts. The standard remedy is to prefilter the data prior to sampling. This is done implicitly in the theory of the wavelet transform, when a signal is projected onto some multiresolution subspace [10]–[13]. Indeed, the combination of prefiltering and sampling is equivalent to forming a sequence of inner products between the input $s(x)$ and some translated versions of an analysis function $\tilde{\varphi}$. This type of Hilbert space formulation has been the basis for various extensions of Shannon's sampling theory for splines [14], [15] and other wavelet-like expansions [16], [17]. The principle behind these sampling theories is that the prefilter has to be tuned to the approximation space defined by the interpolation function φ; in particular, this means that applying an ideal lowpass filter is not necessarily the best solution.

These methods can all be studied from the general perspective of approximation theory [18]–[22]. The most relevant aspect is the behavior of the approximation error as the sampling step T gets sufficiently small. Although there are many error bounds available, they tend to be rather qualitative and not sharp enough to be of direct use to signal processors. What is desirable in this context is a simple and accurate way of predicting the error so that we can compare algorithms and select the interpolation function and sampling step accordingly. Thus far, this goal has been achieved only partially with an exact computation of the error in the asymptotic regime, that is, when T is small or when the signal to approximate is sufficiently smooth. Specific results have been published for wavelets [21], [23], [24] and other types of convolution-based approximation operators [22].

The main purpose of this paper is to introduce a Fourier-based method that will simplify the analysis of the $\mathbf{L}^2$-error, in addition to producing more accurate estimates with a much wider range of applicability. This technique is based on a powerful approximation theorem that we have presented in [25] in the more general case of multiple generators. What makes it very attractive for signal processing is its very natural frequency domain formulation: It involves nothing but a weighting kernel $E(\omega)$ that characterizes the error behavior of an algorithm entirely. This error kernel can be readily computed for any given algorithm; it can also be used to optimize the approximation technique. The simplicity of the concept should appeal to practitioners. The method is quite general and directly applicable to the evaluation of a large variety of approximation techniques, including interpolators and projectors. We will provide numerous examples to illustrate this point. Another important aspect is that despite its apparent simplicity, the present technique has some important

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)07565-0.

approximation theoretic ramifications. In particular, we will show that it can be the basis for obtaining a whole variety of error bounds that are sharper than what has been available before.

The paper is organized as follows. In Section II, we start by defining a general class of linear approximation operators that falls within the classical signal processing acquisition paradigm: prefiltering, sampling, and postfiltering. This relatively broad family includes most commonly used interpolation algorithms, as well as many kinds of projectors that have been used recently to obtain spline and wavelet approximations of signals. In Section III, we will introduce the frequency kernel $E(\omega)$ that provides a simple, convenient characterization of the performance of a given approximation operator. Specifically, we will show that we can make a very accurate prediction of the approximation error by integrating the spectrum of the function to approximate against this kernel; this is a property that will be justified theoretically. We will then use those results to compare the performance of various interpolation and approximation algorithms. In Section IV, we will turn to more theoretical issues and determine the asymptotic form of the error as the sampling step tends to zero, adding higher order terms to what has been published before. We will also derive improved error bounds, including one that is asymptotically sharp; these should be of interest to approximation theorists and signal processors alike because of the way in which the bound constants are directly related to $E(\omega)$. Finally, in Section V, we will use our results to design quasi-interpolation algorithms that are near optimal in the least-squares sense. These should provide essentially the same performance as the more sophisticated projection operators but at a lower cost since no analog prefiltering is necessary.

Most of our present results are directly applicable to the characterization of the approximation error of wavelet expansions [21], [23], [24]. This is an aspect that will be further investigated in a companion paper. In particular, we will show how the use of the two-scale relation can simplify the determination of many of the bounds constants that are defined in Section IV.

II. Preliminaries: Signal Representation and Approximation

In this paper, we consider the general problem of the reconstruction of a function $s(x)$ of the continuous variable x from a discrete set of measurements (e.g., sample values) collected on a uniform grid with step size T. In general, the reconstruction $\mathcal{Q}_T s(x)$ will only be an approximation of $s(x)$ in some signal subspace V_T; $\mathcal{Q}_T$ is a linear operator that depends explicitly on the sampling step T. Our main interest here will be to quantify the difference between $s(x)$ and its approximated version $\mathcal{Q}_T s(x)$. In principle, we should expect the approximation to improve as the sampling step gets smaller. In the limit, as T approaches zero, we want it to be exact for any "reasonable" input function $s(x)$. In this section, we define all the relevant signal spaces and specify our class of approximation operators $\mathcal{Q}_T$. We also review some basic concepts from approximation theory.

A. Notations

The conventional inner product $\int s_1(x)s_2(x)\,dx$ between two $\mathbf{L}^2$ functions s_1, s_2 is denoted $(s_1 s_2)$, 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 $\int(1+\omega^2)^r|\hat{s}(\omega)|^2\,d\omega<\infty$. By analogy to this definition of regularity, we extend $\|s^{(r)}\|_{\mathbf{L}^2}$ to noninteger values of r by equating it to $\sqrt{\frac{1}{2\pi}\int|\omega|^{2r}|\hat{s}(\omega)|^2\,d\omega}$. 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 *point-wise* 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 [26].

The infinite norm $\sup_{x\in\mathbb{R}}|f(x)|$ of a function f will be denoted by $\|f\|_\infty$.

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 or by their z-transform $H(z)=\sum_n h_n z^n$, for which we use upper-case symbols.

Most of the asymptotic expansions are presented with "$o(\cdot)$" and "$O(\cdot)$" terms, which allows us to give 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 writing $\limsup_{x\to 0}|f(x)/x^n|<\infty$ (i.e., not necessarily 0).

B. Signal Subspaces

We want our interpolation/approximation algorithms to have a simple "shift-invariant" structure that is well adapted to signal processing. For this purpose, we consider the generic reconstruction formula

$$s_c(x)=\sum_n c_n\varphi\Big(\frac{x}{T}-n\Big) \tag{1}$$

where the c_n's are some signal coefficients, and where $\varphi(x)$ is a user-specified "interpolation" function. The coarseness of this representation is controlled through the scale parameter T (sampling step); this gives the expansion formula (1) a wavelet-like flavor. If we allow the c_n's to take arbitrary values, then (1) defines a vector space $V_T=\text{span}_{n\in\mathbb{Z}}\{\varphi(\frac{\cdot}{T}-n)\}$. Although we want to have as much freedom as possible for selecting the function $\varphi(x)$, it is important that V_T be a well-defined subspace of $\mathbf{L}^2$ and that each of its functions $s_c(x)\in V_T$ have a unique and stable representation in terms of the coefficients c_n. In other words, we want (1) to be unambiguous. This turns out to be the case if the functions $\{\varphi(.-n)\}_{n\in\mathbb{Z}}$ constitute a Riesz basis of V_1. Mathematically, this means that there exist two positive constants $B\geq A>0$ such that for all $c\in\ell^2$

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

This constraint is equivalent to $A\leq\hat{a}_\varphi(\omega)\leq B$ almost everywhere (cf. [16]), where the 2π-periodic function $\hat{a}_\varphi(\omega)$

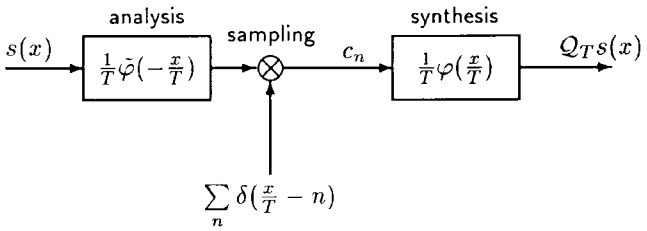

Fig. 1. Block diagram representation of the approximation operator $\mathcal{Q}_T$. The boxes represent continuous-time filters characterized by their impulse response. Sampling is modeled by a multiplication with a train of Dirac functions. When $\tilde{\varphi}(x) = \delta(x)$ (identity operator), the system is an interpolator. The operator is a projector if and only if φ and $\tilde{\varphi}$ are biorthonormal.

is defined by

$$\hat{a}_\varphi(\omega) = \sum_n |\hat{\varphi}(\omega + 2n\pi)|^2. \tag{3}$$

From now on, we will use the term *generating function* to indicate that φ satisfies this hypothesis. Note that the Riesz condition is not restrictive at all: It is satisfied by virtually any interpolation function used in practice. In particular, there is no requirement that φ be compactly supported. This admissibility property is also necessarily satisfied by all scaling functions encountered in the multiresolution theory of the wavelet transform [11], [27], [28], [29]. These latter functions are much more constrained because they must also be solution of a two-scale relation: an assumption that is not made in this paper.

C. Interpolation and Approximation Operators

The next question is how to get the coefficients c_n in (1) so that $\mathcal{Q}_T s(x) = s_e(x)$ is a reasonable representation of some desired signal $s(x)$. The natural signal processing answer is through sampling. However, we want our scheme to be more general than simple interpolation, and we want it to account for wavelet-like approximation methods that have been proposed recently [16], [17], [27], [28]. Therefore, we include an additional prefiltering step prior to sampling, which is also consistent with the standard discretization procedure dictated by Shannon's sampling theorem. This leads to the signal processing system that is schematically represented in Fig. 1. Mathematically, the combination of prefiltering and sampling is conveniently described by the inner product integral

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

where $\tilde{\varphi}$ is the so-called analysis (or sampling) function—it is simply the time-reversed version of the normalized impulse response of the prefilter in Fig. 1. Combining (1) and (4), we end up with the following definition of the approximation operator Q_T:

$$\mathcal{Q}_T s(x) = \sum_n \left[\int s(\xi)\tilde{\varphi}\left(\frac{\xi}{T} - n\right) d\frac{\xi}{T}\right]\varphi\left(\frac{x}{T} - n\right). \tag{5}$$

Our last constraint is that its Fourier transform $\hat{\tilde{\varphi}}$ be upper bounded [25]. This gives us the flexibility of considering non$\mathbf{L}^2$ analysis functions such as the Dirac point distribution $\delta(x)$.

A classification and summary of the different linear, integer shift-invariant, approximation methods is given in Table I. The most critical choice in the design of approximation operators is the selection of φ because it determines the approximation space V_T. A key concept is the order of approximation that requires that φ has some very specific properties (cf. Section II-D). Once this choice is made, we have some freedom in selecting $\tilde{\varphi}$ such that the approximation scheme performs appropriately. A design constraint found in many recent approximation schemes is that $\tilde{\varphi}$ must be biorthonormal to φ, i.e., $(\tilde{\varphi}(x-n), \varphi(x)) = \delta_n$. The direct implication of this property is that the operator $\mathcal{Q}_T$ is a projector for it can be verified that $\mathcal{Q}_T\mathcal{Q}_T = \mathcal{Q}_T$. This means that the operator $\mathcal{Q}_T$ has the desirable property of reproducing exactly any function $f \in V_T$. The biorthonormality property is central to the construction of wavelet bases and all the associated multiresolution approximation operators [11], [28], [29]. It has also been used to formulate generalized sampling theories [16], [17].

Interestingly, the standard interpolation schemes (which use no prefiltering at all) can also be interpreted from this newer perspective. In this case, we have that $\tilde{\varphi}(x) = \delta(x)$. The biorthonormality property is then equivalent to the standard interpolation condition $\varphi(k) = \delta_k$, which ensures that the expansion coefficients in (1) are precisely the values of the function at the grid points.

The best approximation scheme in V_T is the one that minimizes the error (least-squares solution) [16]. In this case, the optimal prefilter is uniquely specified through the biorthonormality condition, along with the additional constraint $\tilde{\varphi} \in V(\varphi)$. The corresponding function φ_d is called the dual of φ and is defined as

$$\hat{\varphi}_d(\omega) = \frac{\hat{\varphi}(\omega)}{\hat{a}_\varphi(\omega)} \tag{6}$$

where $\hat{a}_\varphi(\omega)$ is given by (3). This optimal scheme will provide an *orthonormal projection*, which we denote by $\mathcal{P}_T$ instead of $\mathcal{Q}_T$. A special case that falls into this category is $\varphi(x) = \varphi_d(x) = \mathrm{sinc}(x)$, which describes Shannon's classical reconstruction formula for bandlimited signals [6].

D. Lth-Order Approximation Operators

A crucial notion in approximation theory is the order of approximation, which describes the rate of decay of the error as the sampling step goes to zero. Since this is primarily a property of the approximation space, mathematicians have been especially interested in characterizing what happens in the best possible case (least-squares approximation and other $\mathbf{L}^p$ norms) [18], [30], [31], [32], [33]. The basic result in this area, due to Strang and Fix [18], is that the minimum error has an Lth-order decay i.e., $\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \propto T^L$, if and only if

$$\hat{\varphi}(0) \neq 0, \quad \text{and} \quad \hat{\varphi}^{(k)}(2n\pi) = 0 \quad \text{for} \begin{cases} n \neq 0 \\ k = 0 \ldots L-1 \end{cases}. \tag{7}$$

This last equation imposes some strong constraints on the Fourier transform of φ. It is often referred to as the Strang–Fix conditions of order L. In [18], Strang and Fix had originally assumed that φ has compact support; their powerful equivalence

TABLE I
LINEAR APPROXIMATION METHODS

approximation method	analysis	synthesis	references
interpolation	$\tilde{\varphi}(x) = \delta(x)$	$\varphi(n) = \delta_n$	[6], [7], [38], [37]
quasi-interpolation of order L	$\tilde{\varphi}(x) = \delta(x)$	$\hat{\varphi}(\omega + 2n\pi) = \delta_n + O(\omega^L)$	[20], [22], [30], [19], [34]
$\mathbf{L}^2$ optimal	$\tilde{\varphi} = \varphi_d$	not specified	[31], [18]
$\mathbf{L}^2$ projection	$(\tilde{\varphi}(x-n), \varphi(x)) = \delta_n$	not specified	[21], [42]
quasi-biorthonormality of order L	$\hat{\tilde{\varphi}}(\omega)^* \hat{\varphi}(\omega + 2n\pi) = \delta_n + O(\omega^L)$		[36], [25], [23]
general case	not specified		this paper

has since then been extended for noncompactly supported φ with suitable polynomial decay [34] or even less restrictions [25], [31].

Another equivalent formulation of (7) is that all polynomials of degree $n = L - 1$ must be expressible as a linear combination of the translates of φ. This connection can be established with the help of Poisson's summation formula; for instance, the first-order Strang–Fix condition with $\hat{\varphi}(0) = 1$ is equivalent to a partition of unity, namely, $1 = \sum_n \varphi(x - n)$, which plays a crucial role in wavelet theory. In this latter context, the Strang–Fix conditions usually appear in a more disguised form for the order L relates to the number the vanishing moments of the analysis wavelet [11].

Although the bounds of the classical Strang–Fix theory apply to the case of an orthonormal projection, it is not too difficult to conceive of suboptimal schemes that achieve the same rate of decay of the error. The better-known examples in approximation theory are interpolators and quasi-interpolators [19], [20], [30], [35]. By extension, we can also characterize the most general class of linear approximation operators $\mathcal{Q}_T$ that are of order L. The fundamental constraint is that $\mathcal{Q}_T$ must reproduce the polynomials of degree $n = L-1$ perfectly [30], which is obviously possible only if V_1 is an Lth-order subspace. For the class of "integer shift-invariant" operators considered here, this will be the case if and only if φ and $\tilde{\varphi}$ are *quasibiorthonormal* of order L [25], which is equivalent to the requirement that, in addition to the Strang–Fix conditions of order L on φ, the moments of $\tilde{\varphi}$ and of φ_d be equal up to the order L (i.e., $\int x^k \tilde{\varphi}(x)\,dx = \int x^k \varphi_d(x)\,dx$ for $k = 0 \ldots L-1$). This condition can also be expressed in the condensed form of Table I. Specifically, when φ and $\tilde{\varphi}$ satisfy some additional mild conditions on their decrease (which is automatically the case when φ, $\tilde{\varphi}$ are compactly supported [25]), we have the equivalence

$$\forall f \in \mathbf{W}_2^L, \quad \|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} = O(T^L) \Leftrightarrow \varphi \text{ and } \tilde{\varphi} \text{ are quasibiorthonormal of order } L \tag{8}$$

which comes as a special case of [25, Th. 3].

For a given φ, the implication of this result is that we still have a large freedom in designing $\tilde{\varphi}$ so that the operator $\mathcal{Q}_T$ is well behaved. There are only L linear constraints (quasibiorthonormality conditions) that need to be satisfied for $\mathcal{Q}_T$ to have an Lth order of approximation. This is much less restricting than the usual biorthonormality condition since the latter also implies the Lth-order property [21]. Quasi-biorthonormality may therefore be a very relevant condition for the design of simplified approximation algorithms such as the quasiprojectors in [36] that provide essentially the same type of performance as the least-squares solution (see also [23]). Note that $L = 1$ is the minimum requirement for the error to vanish as T tends to zero.

III. FOURIER-DOMAIN CHARACTERIZATION OF THE APPROXIMATION ERROR

The theoretical notion of order described in Section II is still rather qualitative. There are many applications (e.g., image processing) for which it would be very useful to have a more quantitative way of estimating $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$. This would allow for an objective comparison of the various approximation techniques available. It would also be convenient to have error formulæ that are simple enough so that the determination of the optimal $\tilde{\varphi}$ can be carried out using standard optimization techniques.

We will now present a new method that has the desired features. First, it has a very simple Fourier-domain formulation that should be appealing to signal processors. Second, it allows for a more accurate prediction of the $\mathbf{L}^2$ approximation error than any of the approximation theoretic estimates that have been available so far. This claim of increased accuracy will be further justified in Section IV. Third, the proposed form of the error criterion is directly amenable to standard filter design techniques: a possibility that is illustrated in Section V.

The key quantity used to characterize an approximation operator $\mathcal{Q}_T$ is the frequency kernel $E(\omega)$, which is defined as follows:

$$E(\omega) = |1 - \hat{\tilde{\varphi}}(\omega)^* \hat{\varphi}(\omega)|^2 + |\hat{\tilde{\varphi}}(\omega)|^2 \sum_{n \neq 0} |\hat{\varphi}(\omega + 2n\pi)|^2 \tag{9}$$

$$= \underbrace{1 - \frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}}_{E_{\min}(\omega)} + \underbrace{\hat{a}_\varphi(\omega) |\hat{\tilde{\varphi}}(\omega) - \hat{\varphi}_d(\omega)|^2}_{E_{\mathrm{res}}(\omega)}. \tag{10}$$

Note that when φ and $\tilde{\varphi}$ are exact duals (least-squares approximation), this kernel reduces to $E_{\min}(\omega)$.

The approximation error is simply predicted from the integral formula

$$\eta_s(T) = \left[\frac{1}{2\pi}\int |\hat{s}(\omega)|^2 E(T\omega)\, d\omega\right]^{\frac{1}{2}} \tag{11}$$

which involves the spectrum of the function to approximate and a rescaled version of the kernel. Relation (10) implies that the criterion $\eta_s(T)$ with a given choice of φ is minimized when $\hat{\tilde{\varphi}}(\omega) = \hat{\varphi}_d(\omega)$, that is, when, $\tilde{\varphi} = \varphi_d$. This solution is precisely the orthonormal projection $\mathcal{P}_T$, which is known to minimize the true error approximation in the $\mathbf{L}^2$ sense. We will now present three arguments that justify the use of the criterion $\eta_s(T)$ to characterize the behavior of any approximation algorithm. The first one is approximation theoretic [e.g., $\eta_s(T)$ is an excellent approximation of $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$ for smooth or bandlimited functions], whereas the two others are more pragmatic because they show that $\eta_s(T)$ defined by (11) represents an average measure of the error, which is often better suited for signal processing than the exact $\mathbf{L}^2$ measure $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$. We then conclude Section III by applying our methodology to the comparison of interpolation and approximation algorithms.

A. Approximation Theoretic Argument

The primary justification for the use of (11) to represent the error is given by the main approximation theorem in [25], which we particularize here for the case of a single generator. To describe this result, it is useful to define $e_g(T)$: the relative out-of-band energy of an $\mathbf{L}^2$ function g

$$e_g(T)^2 = \frac{\int_{|\nu|\geq\frac{\pi}{T}} |\hat{g}(\omega)|^2\, d\omega}{\int |\hat{g}(\omega)|^2\, d\omega} \tag{12}$$

This relative energy is always smaller than 1 and tends to zero as the sampling step $T \to 0$. Thus, we have $e_g(T) = o(1)$.

Theorem 1: Let s be in $\mathbf{W}_2^r$ with $r > \frac{1}{2}$; then, the approximation error is given by the equality

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} = \eta_s(T) + \gamma e_{s^{(r)}}(T) T^r \left\|s^{(r)}\right\|_{\mathbf{L}^2} \tag{13}$$

where

$$|\gamma| \leq K = \frac{2}{\pi^r}\sqrt{\zeta(2r)\|E\|_\infty}. \tag{14}$$

In addition, the second term of (13) exhibits a double aliasing character (in s and in φ). In other words, it vanishes whenever one of the following conditions is met.

- Either s is bandlimited in $\left[-\frac{(k+1)\pi}{T}, -\frac{k\pi}{T}\right] \cup \left[\frac{k\pi}{T}, \frac{(k+1)\pi}{T}\right]$, where k is some positive integer, or
- φ and $\tilde{\varphi}$ are bandlimited in $[-\pi, \pi]$.

The first term in (13), namely, $\eta_s(T)$, is thus the dominant error contribution. The second term is an $o(T^r)$ error correction that may take positive or negative values, depending on the sign of γ, the absolute value of which is bounded by (14). Thus, there are essentially two cases in which (11) provides an exact measure of the error: i) when the input signal $s(x)$ is bandlimited and ii) when $s(x)$ is sufficiently smooth, i.e., when its intrinsic scale T_0 is large with respect to the sampling step T.

The regularity constraint $r > 1/2$ in Theorem 1 is not very restrictive. In particular, it is satisfied whenever $|\hat{s}(\omega)| \leq C|\omega|^{-1-\alpha}$ for any $\alpha > 0$, which implies that $s(x)$ should be not much more than continuous. Note, however, that this does not mean that *any* continuous function satisfies the constraint.

A special case of this result can also be found in [31]. These authors examined the more restrictive least-squares case $\tilde{\varphi} = \varphi_d$ and identified a quantity that is the same as the first error term in (13). Considering only the interpolating case $\tilde{\varphi}(x) = \delta(x)$, another version of the second part of Theorem 1 (restricted to Shannon bandlimited functions) also appeared in [37] and [38].

As we have shown in [25], Theorem 1 is a quite powerful result that has many interesting implications for approximation theory. We will further investigate this aspect in Section IV. In particular, we will exploit Theorem 1 to derive a variety of very accurate error estimates that can be directly applicable in practice. Our new results will include an exact asymptotic expansion of the error as $T \to 0$, as well as some upper bounds that are asymptotically sharp. The interest for signal processors is that all the underlying bounds constants are directly tied to the kernel $E(\omega)$.

Although we have just seen that the second error term in (13) vanishes provided that s is sufficiently well-behaved (smooth or bandlimited), we will now show that there is still another much less restrictive way to let it vanish, namely, by averaging the error over all possible sampling phases or, similarly, over all realizations of a stationary process.

B. Average Approximation Error

In signal processing, where we often characterize signals by their Fourier spectrum, the precise origin (i.e., starting point) of the signal is usually irrelevant. It is, thus, of interest to find a shift-invariant version of Theorem 1 where the error would be averaged over all possible shifts of the input signal.

Assume that we want to approximate s_ξ defined as $s(.-\xi)$, where the sampling phase ξ is any real number. The amplitude of the approximation error varies with ξ: Obviously, it is T periodic. Thus, we can obtain a delay-independent version of the approximation error by averaging $\|s_\xi - \mathcal{Q}s_\xi\|_{\mathbf{L}^2}^2$ over the period interval $[0, T]$.

The following result, which is remarkable for its simplicity, was proven in [25].

Theorem 2: Under the conditions of Theorem 1 on the input signal $s(x)$, the average approximation error is *exactly* the same as the quantity defined by (11)

$$\frac{1}{T}\int_0^T \|s_\xi - \mathcal{Q}s_\xi\|_{\mathbf{L}^2}^2\, d\xi = \eta_s^2(T). \tag{15}$$

This result thus provides the expected (in a probabilistic sense) value of the approximation error when the time of presentation of the signal—the delay or shift ξ—is assumed to be random and uniformly distributed.

As a nontrivial application, we may use this result to quantify the lack of shift-invariance of the basic representation

space $V_1 = \text{span}_{n\in\mathbb{Z}}\{\varphi(x-n)\}$. For this purpose, we simply consider the orthonormal projection of φ_ξ onto $V(\varphi)$ and average the square approximation error $\|\varphi_\xi - \mathcal{P}_1\varphi_\xi\|^2_{\mathbf{L}^2}$ over $\xi \in [0,1]$ (here, we have $T=1$). Denoting this average by σ^2_φ, (11) yields

$$\sigma^2_\varphi = \frac{1}{2\pi}\int |\hat{\varphi}(\omega)|^2\left(1 - \frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}\right)d\omega. \tag{16}$$

Note that $\sigma^2_\varphi = 0$ when φ is bandlimited; for instance, this is the case when $\varphi(x) = \text{sinc}(x)$.

C. Approximation Error for Random Processes

Let us now momentarily assume that s is a realization of a zero-mean, wide-sense stationary process. Such a random process is characterized by its autocorrelation function $\rho_s(t)$ whose Fourier transform is the power spectrum density (PSD) $P_s(\omega)$. Under ergodicity hypothesis, the Wiener–Khinchin theorem states that the PSD can also be obtained through the following limit process:

$$P_s(\omega) = \lim_{M\to\infty}\frac{1}{2M}\left|\int_{-M}^{M} s(\xi)e^{-i\omega\xi}\,d\xi\right|^2.$$

This expression leads us to foresee that an estimate of the approximation error for a random process will be obtained merely by replacing $|\hat{s}(\omega)|^2$ by $P_s(\omega)$ in (11). We give below a formal proof of this fact.

Although we have to make sure that (4) and (1) are well defined for the approximation scheme to preserve a meaning, we shall not examine in details the conditions under which such a property holds. Note, however, that this is true when φ and $\tilde{\varphi}$ are compactly supported.

Since s is not in $\mathbf{L}^2$, we use a time averaged form ε_s of the approximation error that is also averaged over all realizations of the process. We thus define

$$\begin{aligned}\varepsilon^2_s &= \left\langle \lim_{M\to\infty}\frac{1}{2M}\int_{-M}^{M}|s(t) - \mathcal{Q}_T s(t)|^2\,dt\right\rangle \\ &= \frac{1}{T}\int_0^T \langle |s(t) - \mathcal{Q}_T s(t)|^2\rangle\,dt \end{aligned}\tag{17}$$

where the second equality follows from the observation that $\langle |s(t) - \mathcal{Q}_T s(t)|^2\rangle$ is T periodic.

With this definition, the following theorem, which is proven in Appendix A, is the exact equivalent of the deterministic result (11).

Theorem 3: For a stationary random process $s(t)$ with PSD $P_s(\omega)$, the approximation error ε_s is given by

$$\varepsilon^2_s = \frac{1}{2\pi}\int P_s(\omega)E(T\omega)\,d\omega. \tag{18}$$

This result sheds a new light on the error kernel $E(\omega)$. Indeed, if the process $s(t)$ is convolved with an analog filter $h(t)$, then the output process $y(t)$ has an average energy $\langle y^2\rangle = \frac{1}{2\pi}\int P_s(\omega)|\hat{h}(\omega)|^2\,d\omega$; this value coincides with ε^2_s (19) when h is such that $|\hat{h}(\omega)|^2 = E(\omega T)$. In particular, if the process is such that its PSD is concentrated at a given frequency ω_0, i.e., $P_s(\omega) = P_0\delta(\omega - \omega_0)$, then we have $\varepsilon^2_s = \frac{P_0}{2\pi}E(\omega_0 T)$. Thus, we can predict the approximation error by computing the energy of the result of prefiltering s with an analog filter whose response is $\sqrt{E(\omega T)}$.

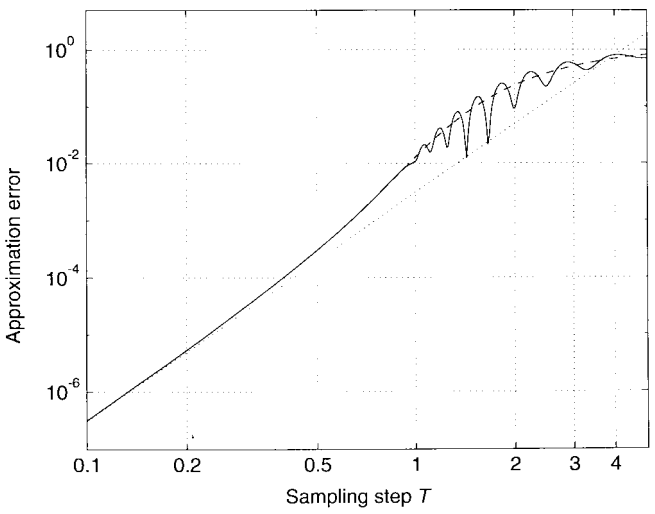

Fig. 2. Least-squares approximation error of the function $s(x) = e^{-\frac{x^2}{2}}$ by cubic splines as a function of the sampling step T. The error estimate η_s (dashed line) is an unbiased smoothed version of the true error (solid line), as stated in Theorem 2; the dotted line is the first order asymptote $C^-_\varphi\|s\|_{\mathbf{L}^2}T^4$, computed according to [21].

This interpretation is very useful since it shows how we can adapt the sampling/generating functions to the frequency content of the input process $s(t)$ or, even better, how we choose them to obtain a desired accuracy within a given frequency band, regardless of the other parts of the spectrum.

D. Quantitative Assessment of Interpolation and Approximation Algorithms

We now illustrate the accuracy and efficiency of our theoretical results by some examples. We also compare the merits of various approximation algorithms.

1) Comparison Between the True $\mathbf{L}^2$ Error and η_s: First, we want to demonstrate experimentally that η_s is a very good indicator of the true approximation error. For this purpose, we approximate the test function $s(x) = e^{-\frac{x^2}{2}}$ by cubic splines with step size T. The exact least-squares approximation error as a function of T is plotted in Fig. 2; it can be compared with our calculated error estimate η_s. The two curves are very close to each other and almost indistinguishable for $T < 1$. Moreover, for $T \geq 1$, the plot illustrates the average property satisfied by η_s (see Theorem 2).

2) Exact Expression of the Kernel $E(\omega)$: For B-splines of any order, the minimal kernel $E(\omega)$ (i.e., when $\tilde{\varphi} = \varphi_d$) and the interpolation one can be computed exactly. This is due to the fact that $\varphi(x)$ has a simple expression in the Fourier domain given by $\hat{\varphi}(\omega) = \hat{\varphi}_L(\omega) = \left[\frac{\sin(\omega/2)}{\omega/2}\right]^L$, where L is the order of the B-spline (note that the spline order is one more than the spline degree).

If we consider the least-squares error, we have that $E_{\min}(\omega) = 1 - \frac{\hat{\varphi}_L(\omega)^2}{\hat{a}_L(\omega)}$, where $\hat{a}_L(\omega)$ is the autocorrelation filter defined by (3). According to (10), the general

TABLE II
APPROXIMATION KERNELS FOR THE FIRST SIX SPLINES

order L	$\hat{a}_L(2\omega)$	$\hat{b}_L(2\omega)$	$E_{\min}(2\omega)$	$E_{\text{res}}(2\omega)$
1	1	1	$1-\frac{\sin^{2L}(\omega)}{\omega^{2L}\hat{a}_L(2\omega)}$	$\left\lvert\frac{\sqrt{\hat{a}_L(2\omega)}}{\hat{b}_L(2\omega)}-\frac{\sin^L(\omega)}{\omega^L\sqrt{\hat{a}_L(2\omega)}}\right\rvert^2$
2	$1-\frac{2}{3}\sin^2(\omega)$	1		
3	$1-\sin^2(\omega)+\frac{2}{15}\sin^4(\omega)$	$1-\frac{1}{2}\sin^2(\omega)$		
4	$1-\frac{4}{3}\sin^2(\omega)+\frac{2}{5}\sin^4(\omega)$ $-\frac{4}{315}\sin^6(\omega)$	$1-\frac{2}{3}\sin^2(\omega)$		
5	$1-\frac{5}{3}\sin^2(\omega)+\frac{7}{9}\sin^4(\omega)$ $-\frac{17}{189}\sin^6(\omega)+\frac{2}{2835}\sin^8(\omega)$	$1-\frac{5}{6}\sin^2(\omega)$ $+\frac{1}{24}\sin^4(\omega)$		
6	$1-2\sin^2(\omega)+\frac{19}{15}\sin^4(\omega)$ $-\frac{256}{945}\sin^6(\omega)+\frac{62}{4725}\sin^8(\omega)$ $-\frac{4}{155925}\sin^{10}(\omega)$	$1-\sin^2(\omega)$ $+\frac{2}{15}\sin^4(\omega)$		

approximation kernel is obtained by adding a residual term $E_{\text{res}}(\omega)$ to the least-squares error. In the interpolation case, this extra term reads $E_{\text{res}}(\omega)=|\frac{\sqrt{\hat{a}_L(\omega)}}{\hat{b}_L(\omega)}-\frac{\hat{\varphi}_L(\omega)}{\sqrt{\hat{a}_L(\omega)}}|^2$, where the interpolation filter $\hat{b}_L$ is defined by

$$\hat{b}_L(\omega)=\sum_n \hat{\varphi}_L(\omega+2n\pi). \tag{19}$$

By definition, $\hat{a}_L=\hat{b}_{2L}$. One way of computing $\hat{b}_L$ is to use the identities $\cot(x)=\sum_n(x+n\pi)^{-1}$ when L is even and $\sin^{-1}(x)=\sum_n(-1)^n(x+n\pi)^{-1}$ when L is odd (for a different computing method, see [39] and [40]. We thus find

$$\hat{b}_L(2\omega)=-\frac{\sin^L(\omega)}{(L-1)!}\frac{d^{L-1}}{d\omega^{L-1}}(\cot(\omega)),\quad \text{if } L \text{ is even}$$

$$\hat{b}_L(2\omega)=\frac{\sin^L(\omega)}{(L-1)!}\frac{d^{L-1}}{d\omega^{L-1}}(\sin(\omega)^{-1}),\quad \text{if } L \text{ is odd.}$$

The analytical values of the kernels for the first six orders are given in Table II. The plots of these kernels are shown in Fig. 3. We can verify that the interpolation is always worse than the least-squares scheme, although the difference between both tends to become negligible as the order increases.

3) Comparison with Keys' Interpolating Kernel: A family of short piecewise polynomial kernels was proposed by Keys [7] in 1981. Unlike B-splines, these functions satisfy the interpolation property, i.e., $\varphi(n)=\delta_n$. The family that is made of piecewise cubic polynomials, depends on one parameter a and can be defined in the Fourier domain by

$$\begin{aligned}\omega^4\hat{\varphi}(\omega)=&12(a+2-2\cos\omega-a\cos 2\omega)\\&-4\omega((4a+3)\sin\omega+a\sin 2\omega).\end{aligned} \tag{20}$$

As shown in Fig. 4, for $a=-0.7$, -0.5 and -0.3, Keys' functions perform significantly worse than the interpolating splines of order 3 and 4 (i.e., quadratic and cubic splines). This is also true asymptotically as $\omega\to 0$, although it is not visible in Fig. 4. This is an important observation because Keys' short cubic kernel is still considered by many to be the state-of-the art interpolation method in image processing. Even though

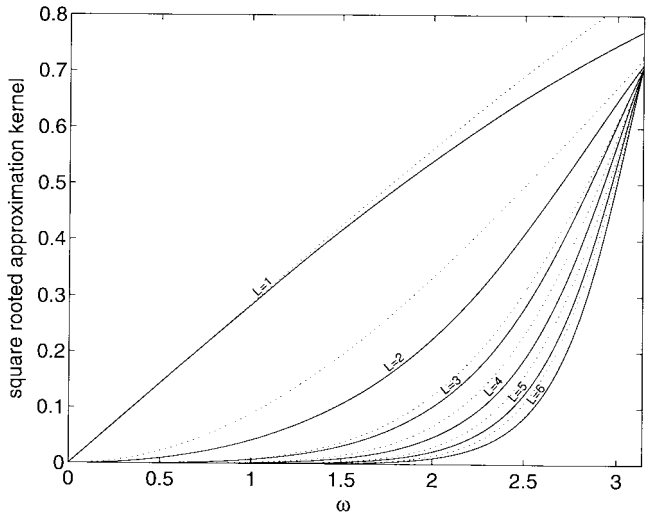

Fig. 3. Least-squares (solid lines) and interpolating (dotted lines) rooted approximation kernels for the B-splines of order 1 to 6. The kernel is all the closer to the ω-axis as the approximation order is higher.

Keys' functions have the interpolation property, they do not result in a faster algorithm in two dimensions if we take into account the whole interpolation process (a timed comparison can be found in [41]). Quadratic or cubic spline interpolators are quite competitive computationally because the cost of the additional prefiltering that is required is negligible; most computational resources are spent in the kernel evaluation. Thus, there does not appear to be any reason for not preferring splines.

4) Oblique Projections: The best way to approximate a function $f(x)$ in some subspace $V_T(\varphi)$ is to compute its least-squares approximation. Unfortunately, this requires the computation of inner products with the dual function φ_d; those can be difficult to obtain in practice. Vrhel *et al.* [42] have proposed to simplify the process by using the shortest first-order analysis function, namely, the Haar function (i.e., a spline of order 1), while still performing a projection in the space of $V_T(\varphi)$ (e.g., cubic splines). The key point is that the inner products can be computed by straightforward

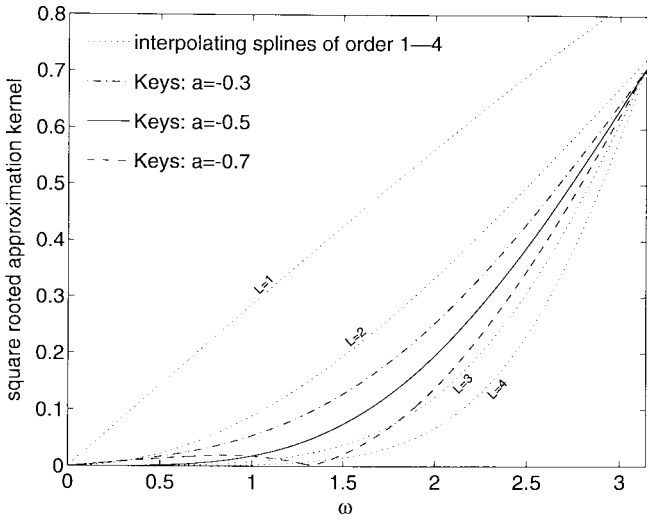

Fig. 4. Comparison between the approximation kernels of Keys' function for different values of the parameter a and interpolating splines of order 1 to 4.

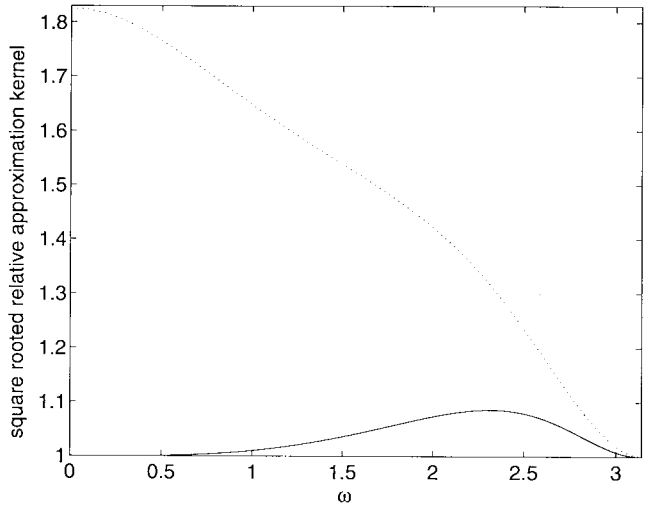

Fig. 5. Comparison between an oblique projection (solid line) and an interpolation (dotted line) in the space of cubic splines. The approximation kernels E for both methods are plotted relative to the minimal kernel, i.e., $\omega \mapsto \sqrt{\frac{E(\omega)}{E_{\min}(\omega)}}$.

integration of $f(x)$ between two bounds. The approximation is then computed using modified synthesis functions; these are chosen biorthonormal to $\varphi_1 = \tilde{\varphi}$ (oblique projection). These authors found experimentally that this biorthonormal scheme results in no more than a very slight loss of performance when compared with the least-squares approximation. This is confirmed by our Fourier analysis: In Fig. 5, the corresponding rooted approximation kernel appears to be off by less than 10% from the minimal approximation kernel. On the other hand, the performance is by far superior to that of the corresponding interpolator.

IV. Approximation Theoretical Results

In this section, we will use Theorem 1 to characterize the behavior of the error as the sampling step gets sufficiently small. We will also derive new error bounds and provide sharp estimates for the leading constants that appear in the Strang–Fix theory of approximation. In addition to those new results, a worthwhile contribution is the simplicity of the argument, which does not involve much more than taking the Taylor series expansion of the symmetrical kernel $E(\omega)$ around the origin $\omega = 0$.

A. Asymptotic Error

From Theorem 1 we see that as $T \to 0$, we have $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} = \eta_s + o(T^r)$. Since η_s has the closed form (11), it is possible to expand the approximation error in power series of T, which gives us the exact behavior of the error as the sampling step becomes sufficiently small. Let us assume that r_0 is an integer lesser than the Sobolev regularity exponent of s. Then, under the assumption that the Fourier kernel can be developed up to the degree r_0 in Taylor series, we have $E(\omega) = \sum_{k=0}^{r_0} \frac{E^{(2k)}(0)}{(2k)!}\omega^{2k} + o(\omega^{2r_0})$ (only even powers of ω are involved since $E(\omega)$ is even). Recalling that $\omega^k \hat{s}(\omega)$ is the Fourier transform of $s^{(k)}(x)$, which is the kth derivative of $s(x)$, we can thus derive an asymptotic error formula

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}^2 = \sum_{k=0}^{r_0} \frac{E^{(2k)}(0)}{(2k)!} \left\|s^{(k)}\right\|_{\mathbf{L}^2}^2 T^{2k} + o(T^{2r_0}). \quad (21)$$

The first nonzero coefficient of the Taylor series expansion will give the asymptotic rate of decay of the error $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}$ as a function of T. Specifically, if φ satisfies the Strang–Fix conditions (7) of order L, and if $\tilde{\varphi}$ is chosen appropriately (i.e., quasibiorthonormal), then the error will have the characteristic form

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} = C_L \left\|s^{(L)}\right\|_{\mathbf{L}^2} T^L \quad \text{as } T \to 0 \quad (22)$$

where the asymptotic error constant is $C_L = \sqrt{\frac{E^{(2L)}(0)}{(2L)!}}$. This is precisely what is meant when we speak about an Lth-order approximation scheme.

The result given by (21) constitutes quite an improvement over the various specialized forms of (22) that can be found in the literature [21], [22], [36]. First, it is a more general formula with higher order terms up to r_0. Second, the Fourier domain derivation made possible by Theorem 1 is easier and more direct than the time domain approaches that had been proposed before. Based on (21), we can retrieve all previously published asymptotic formulæ of the form of (22). For instance, if φ satisfies the Strang–Fix conditions of order L, and if $\tilde{\varphi}$ is biorthonormal to φ and satisfies $\sum_n \tilde{\varphi}(x-n) = 1$ as in [21], then we find that $C_L = \frac{1}{L!}\sqrt{\sum_{n\neq 0} |\hat{\varphi}^{(L)}(2n\pi)|^2}$, as stated in [21]. This result is, in fact, the general first-order asymptotic equivalent for the minimum approximation error. Likewise, if we are interested in the quasi-interpolation error as defined by [22], [36] (i.e., $\tilde{\varphi}(x) = \delta(x)$), the definition (9) of $E(\omega)$ easily provides the result [22, Prop. 5.1], i.e., (22) with $C_L = \frac{1}{L!}\sqrt{\sum_n |\hat{\varphi}^{(L)}(2n\pi)|^2}$.

In particular, note that if s is bandlimited over $[-\frac{\omega_{\max}}{2}, \frac{\omega_{\max}}{2}]$, then $\|s^{(k)}\|_{\mathbf{L}^2}$ does not increase faster than $C(\frac{\omega_{\max}}{2})^k$, where C is bounded. This means that there exists a convergence radius $T_0 = \frac{2\pi}{\omega_{\max}}$ such that

(21) converges (uniformly, exponentially) for all $T < T_0$, whenever E has a convergent MacLaurin series for $|\omega| < \pi$.

B. Improved Upper Bound for the Error

The natural formulation of Theorem 1 is an upper bound for the error. However, if we want to have sharper bounds, it is better to rewrite (13) using the aliasing property shown in the second part of the theorem. This is actually the same technique as in [25] for the proof of Theorem 1, which amounts to decomposing s into a set of bandlimited functions s_k satisfying the condition of Theorem 1, i.e.,

$$\hat{s}_k(\omega) = \begin{cases} \hat{s}(\omega), & \text{if } \frac{k\pi}{T} \leq |\omega| \leq \frac{(k+1)\pi}{T} \\ 0, & \text{elsewhere} \end{cases}, \quad k \in \mathbb{N}. \tag{23}$$

In that case, due to Minkowski's inequality, we have that

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} \leq \sum_{k \geq 0} \|s_k - \mathcal{Q}_T s_k\|_{\mathbf{L}^2}.$$

Thanks to the second part of Theorem 1, this decomposition gives a closed-form expression for each term of the right-hand side. Still, following the steps of [25], it is not difficult to find the following result, which is more suitable for deriving upper bounds

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} \leq \left[\frac{1}{2\pi} \int_{\frac{-\pi}{T}}^{\frac{\pi}{T}} |\hat{s}(\omega)|^2 E(T\omega)\, d\omega \right]^{\frac{1}{2}} + K_r e_{s^{(r)}}(T) T^r \left\| s^{(r)} \right\|_{\mathbf{L}^2} \tag{24}$$

where $K_r = \pi^{-r} \sqrt{\|E\|_\infty \zeta(2r)}$. The main difference with (13) is that the first integral term is now bandlimited and that we have reduced the upper bound on the second term by a factor of two.

To see how this inequality can be exploited to improve classical results, assume that φ satisfies the Strang–Fix conditions of order L and that $\tilde{\varphi}$ is chosen so that its first L moments are identical to those of φ_d. Then, $\frac{E(\omega)}{\omega^{2L}}$ is bounded over $[-\pi, \pi]$. We choose $r = L$ in (24) so that we can write

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} \leq \left[\sup_{|\omega| < \pi} \frac{E(\omega)}{\omega^{2L}} + K_L^2 \right]^{\frac{1}{2}} \left\| s^{(L)} \right\|_{\mathbf{L}^2} T^L. \tag{25}$$

We have derived this expression by using the Cauchy–Schwartz inequality $(ax+by)^2 \leq (a^2+b^2)(x^2+y^2)$, where x, y are the energy of $\hat{s}$ over $|\omega|T \leq \pi$ and $|\omega|T \geq \pi$, respectively.

The qualitative form of this bound is known in approximation theory, except that no satisfactory estimates for the leading constant were available. To illustrate the quality of the present formula, we recall that if $\hat{s}$ is compactly supported within $[\frac{-\pi}{T}, \frac{\pi}{T}]$, then the second term in (24) cancels, and we end up with an equality as stated in Theorem 1. In this case, we have the same inequality as (25) with the leading constant $C_{\min} = \sup_{|\omega| < \pi} \sqrt{\frac{E(\omega)}{\omega^{2L}}}$. For this particular setup, we can build a sequence of $\mathbf{L}^2$ signals such that the integral formula (24) comes as close as we wish to the right-hand side of the reduced form of (25); this sequence is concentrated around the frequency ω_0 at which $\frac{E(\omega)}{\omega^{2L}}$ achieves its supremum. Thus, the leading constant in the reduced form of inequality (25) is the best that can be achieved within the subclass of bandlimited signals. This proves that the leading constant of a general upper bound of the error cannot be smaller than $C_{\min}$. Thus, the only possible source of degradation in (25) is the K_L term, which makes the inequality less sharp. In other words, (25) will be very good as long as K_L is small with respect to $C_{\min}$.

There are other approaches for estimating the Strang–Fix bound. Depending on the hypotheses on the interpolation and sampling functions, we can derive other accurate results that are compatible with the ones provided in [22] and [36] by including higher order terms. Examples of such bounds, all derived from Theorem 1, can be found in [25].

Here, we propose a new upper bound that is asymptotically exact. To this end, consider the MacLaurin series expansion of $E(\omega)$ up to the order $r_0 - 1$, yielding a remainder $E_{r_0}(\omega)$ defined by

$$E(\omega) = \sum_{k=0}^{r_0 - 1} \frac{E^{(2k)}(0)}{(2k)!} \omega^{2k} + E_{r_0}(\omega) \tag{26}$$

and such that $\omega^{-2r_0} E_{r_0}(\omega)$ is bounded over $[-\pi, \pi]$. An accurate upper bound for the expansion error is

$$\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2}^2 \leq \left| \sum_{k=0}^{r_0 - 1} \frac{E^{(2k)}(0)}{(2k)!} \left\| s^{(k)} \right\|_{\mathbf{L}^2}^2 T^{2k} \right|^{\frac{1}{2}} + \left[\left| \sup_{|\omega| < \pi} \frac{E_{r_0}(\omega)}{\omega^{2r_0}} \right| + K_{r_0}^2 \right]^{\frac{1}{2}} \left\| s^{(r_0)} \right\|_{\mathbf{L}^2} T^{r_0} \tag{27}$$

where r_0 is any positive integer strictly smaller than r, which is the Sobolev regularity of s. The accuracy we claim is justified by the fact that the first expression on the right-hand side is exactly the asymptotic equivalent (21) of the approximation error as $T \to 0$, whereas the second term is of higher order in T, implying that the accuracy of (27) improves as T tends to 0. If, for instance, we select $r_0 = L + 1$, where L is the order of φ and $\tilde{\varphi}$ is chosen so that its first L moments are identical to those of φ_d, then the first term is precisely the asymptotic error in (22). Thus, we end up with an improved bound of the form $\|s - \mathcal{Q}_T s\|_{\mathbf{L}^2} \leq C_L \|s^{(L)}\|_{\mathbf{L}^2} T^L + C' \|s^{(L+1)}\|_{\mathbf{L}^2} T^{L+1}$.

C. Cubic Spline Approximation

We consider the example of a cubic spline approximation to illustrate these various calculations. For simplicity, we concentrate on the case of the least-squares approximation [43]. The corresponding error kernel is given in Table II. Its Taylor series expansion around the origin can be computed exactly. From (21), we get

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2}^2 = \frac{\left\| s^{(4)} \right\|_{\mathbf{L}^2}^2 T^8}{1\,209\,600} + \frac{\left\| s^{(5)} \right\|_{\mathbf{L}^2}^2 T^{10}}{1\,330\,560} + \frac{691 \left\| s^{(6)} \right\|_{\mathbf{L}^2}^2 T^{12}}{3\,962\,649\,600} + \frac{\left\| s^{(7)} \right\|_{\mathbf{L}^2}^2 T^{14}}{43\,545\,600} + O(T^{16}) \tag{28}$$

which confirms the well-known fact that cubic splines have a fourth order of approximation ($L = 4$). In particular, this yields the value of the constant $C_4 = \frac{1}{240\sqrt{21}}$, which is precisely the figure reported in [21]. Likewise, we can also determine the supremum of $\frac{E(\omega)}{\omega^{2L}}$ for $|\omega| < \pi$. A direct application of (25) yields the standard error bound

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \leq 0.0126 \times \left\|s^{(4)}\right\|_{\mathbf{L}^2} T^4. \tag{29}$$

Note that the present value of the upper bound constant is much better than the estimate reported in [22], namely 0.182. The present bound is reasonably sharp since a minimal value of the constant, which is obtained by setting $K_4 = 0$ in (25), is $C_{\min} = 0.007\,26$. In other words, our new bound is only 73% larger than the minimal possible value over the subclass of Shannon bandlimited functions. This percentage may overestimate the difference with the minimal value for *all* functions.

Using (27), we obtain the asymptotically optimal bound

$$\|s - \mathcal{P}_T s\|_{\mathbf{L}^2} \leq C_4 \left\|s^{(4)}\right\|_{\mathbf{L}^2} T^4 + 0.004\,62 \times \left\|s^{(5)}\right\|_{\mathbf{L}^2} T^5 \tag{30}$$

which, to the best of our knowledge, has no previous equivalent in the literature. We shall give more general bounds for splines of any order in the companion paper [44].

V. Quasi-Interpolation (QI)

In its most general form, the approximation operator $\mathcal{Q}_T$ requires an analog prefiltering of the data prior to sampling: an operation that can be difficult to implement in practice. Most often in digital signal processing, we make the assumption that the sampling is ideal, in the sense that our discrete signal values represent the true samples $s_n = s(nT)$ of some signal $s(t)$ that is typically assumed to be bandlimited. If we intend to stay entirely discrete, the only sampling functions that can be considered are a linear combination of Dirac masses

$$\tilde{\varphi}(x) = \sum_k p_k \delta(x + k) \tag{31}$$

which is the continuous-time representation of the digital filter p_k. In the Fourier domain, we have that $\hat{\tilde{\varphi}}(\omega) = P(e^{i\omega})$, where $P(z)$ is the z-transform of p_k. With this particular setting, we obtain a subclass of approximation algorithms that are commonly referred to as *quasi-interpolants* [36], [22].

Note that there are two equivalent ways of computing such a quasiinterpolation operator $\mathcal{Q}_T$:

- either by first prefiltering the discretized data $s_n = s(nT)$ by $P(z)$, which yields a sequence c_n, and then by applying (1) (this is the approach that we recommend in practice);
- simply by evaluating the interpolation formula

$$Q_T s(x) = \sum_k s(nT) \varphi_p(x - k) \tag{32}$$

which uses the "quasi-interpolant" function

$$\varphi_p(x) = \sum_k p_k \varphi(x - k). \tag{33}$$

We mention this second possibility because it corresponds to the more standard description of a quasi-interpolant. The term "quasi-" is used to signify that the interpolation formula (32) only needs to be exact for the polynomials of degree $n = L-1$, i.e., $x^d = \sum_n n^d \varphi_p(x - n)$ for $d = 0 \ldots L - 1$. Note that the quasi-interpolant function φ_p generates the same subspace as φ, provided that $|P(e^{i\omega})|^2$ is essentially bounded from above and below (nonvanishing), cf. [16, Prop. 6].

A. Optimal Filter

We will now consider the problem of finding the best filter $P(e^{i\omega})$ so that the quasi-interpolator (32) is as close as possible to the original signal $s(x)$. The criterion that we will minimize is the mean square error given by (11). In general, once the generating function has been chosen, the optimal filter depends both on the signal s and the value of T.

Since $E(\omega)$ is a *quadratic* function of P, the optimization of the criterion (11) yields a linear system of equations in terms of the coefficients of P. Such a system is not difficult to solve, especially when P is FIR.

The expression (10) of $E(\omega)$ suggests that the optimal P minimizes the contribution of the second term of the right-hand side in (13), that is, $\int |\hat{s}(\omega)|^2 \hat{a}_\varphi(\omega T) |P(e^{i\omega T}) - \hat{\varphi}_d(\omega T)|^2 \, d\omega$. For example, if we assume that $|\hat{s}| = 1$ over its bandlimited support, then the optimal filter P is given by

$$p_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ni\omega} \frac{\hat{\varphi}(\omega)}{\hat{a}_\varphi(\omega)} \, d\omega. \tag{34}$$

In other words, $P(e^{i\omega})$ is chosen so that it matches $\hat{\varphi}_d(\omega)$ in the primary frequency band $|\omega| \leq \pi$. Thus, the system attempts to replicate the least-squares solution (orthonormal projection).

B. Asymptotically Optimal Filter

For applications where most of the spectral energy of the signal is concentrated in the neighborhood of $\omega = 0$ (e.g., images), it may be sufficient that the quasiinterpolation approximation be optimal only in the limit as $T \to 0$. This can be achieved by requiring that the square asymptotic error given by (21) matches the least-squares error up to some order $2N$. Note that the prospect of finding such an asymptotically optimal solution is new. Even though there has been a strong interest for quasi-interpolation in the approximation theory literature, the possibility that the asymptotic expansion error can be arbitrarily close to the minimum error has never been examined before, perhaps due to the absence of a formula such as (21).

This problem is mathematically translated into the condition

$$P(e^{i\omega}) - \hat{\varphi}_d(\omega) = O(\omega^N). \tag{35}$$

We can obviously impose the constraint (35) if P is represented by a trigonometric polynomial of length $N_0 \geq N$. Note that apart from N_0, we have another free integer parameter, namely, the highest power of z in the z-transform representation of P. This parameter can be adjusted in such a way that the approximation kernel behaves well not only near $\omega = 0$ but also in the whole Shannon bandwidth. Intuitively,

this value should be chosen in practice so that $P(e^{i\omega})$ has approximately the same phase behavior as $\hat{\varphi}_d(\omega)$. Moreover, if $N_0 > N$, we can take advantage of the $N_0 - N$ additional free parameters to improve the filter response when one moves away from the origin.

If we are interested in IIR (realizable) filters of the type described in [39] and [40], then it is possible to have an explicit expression for an admissible filter $P(e^{i\omega})$, provided φ has compact support. The idea is to choose a compactly supported function u and to require that the developments of $P(e^{i\omega})$ and $\frac{\hat{\varphi}(\omega)\hat{u}(\omega)^*}{\hat{a}_\varphi(\omega)\hat{u}(\omega)^*}$ around 0 agree up to ω^N. Since $P(e^{i\omega})$ is a 2π-periodic function, we then periodize both the numerator and the denominator of $\frac{\hat{\varphi}(\omega)\hat{u}(\omega)^*}{\hat{a}_\varphi(\omega)\hat{u}(\omega)^*}$

$$P(e^{i\omega}) = \frac{\sum_n \hat{\varphi}(\omega+2n\pi)\hat{u}(\omega+2n\pi)^*}{\hat{a}_\varphi(\omega)\sum_n \hat{u}(\omega+2n\pi)^*}. \tag{36}$$

Of course, since u and φ are compactly supported, the trigonometric polynomials that appear in the numerator and in the denominator have a finite length (using Poisson's summation formula). Now, we readily check that $P(e^{i\omega}) = \hat{\varphi}_d(\omega) + O(\omega^N)$, if $\hat{u}$ is chosen such that $\hat{u}(\omega+2n\pi) = O(\omega^N)$ for $n \neq 0$. Moreover, the right-hand side of (36) is closer to $\hat{\varphi}_d(\omega)$ as $\hat{u}(\omega)$ is more selective in the Shannon bandwidth. This means that P is not only an optimal quasi-interpolator up to the order N—which is an asymptotic property—but that it also remains close to the *ideal* analog sampling function within the sampling frequency interval. This suggests that the use of this filter should offer a smaller approximation error than an exact interpolation scheme that uses $P(e^{i\omega}) = [\sum_n \hat{\varphi}(\omega+2n\pi)^*]^{-1}$, provided u is chosen selective enough.

In order to minimize the degree of the filter in (36), it is advantageous to use a very short function u of order N. This suggests the use of the B-spline of order N; in particular, if φ is symmetrical around 0—which implies that $\hat{\varphi}$ is real—then the centered B-spline

$$\hat{u}(\omega) = \left[\frac{\sin\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}}\right]^N \tag{37}$$

seems to be a good candidate. Note that if we choose $u = \varphi$, then we obtain a prefilter that corresponds to a standard interpolator [40].

In Fig. 6, we plotted the (square root) approximation kernels obtained using a prefilter built with (36) and (37) when φ is a cubic spline and $N = 5, 6, 7$ and 8, which means that $E(\omega) - E_{\min}(\omega) = O(\omega^{10,12,14 \text{and } 16})$, respectively. As can be observed, even for $N = 5$, the kernel is very close to the minimal kernel with an error less than 10%. As an example, the filter for $N = 5$ is given in (38), shown at the bottom of the page. To obtain it, we used the induction formula and the B-spline filters that appeared in [39] and [40].

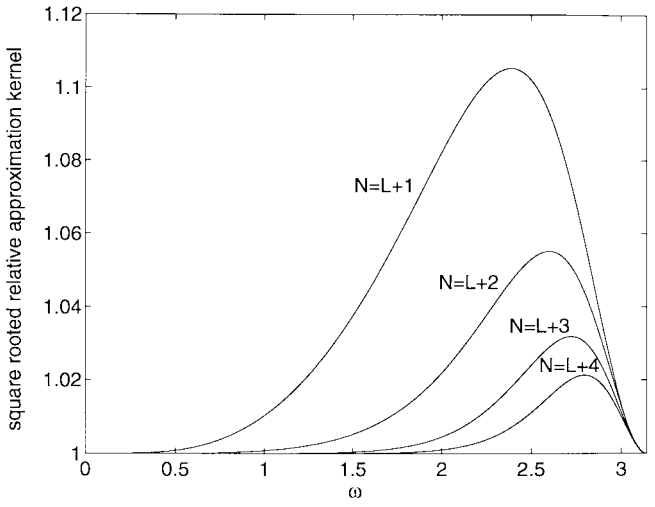

Fig. 6. Optimal cubic spline ($L = 4$) quasi-interpolants: reduction of the error as the order N of the windowing B-spline increases (the same convention as in Fig. 5 has been used for the plots of the kernels).

VI. Conclusion

In this first paper of a series, we have introduced a Fourier method for the analysis of shift-invariant approximation schemes. This method provides simple and direct tools for evaluating the quality of a whole variety of approximation algorithms. It shows in many details how to choose both the approximating and the sampling functions. We have applied our theoretical results to the comparison of some commonly used algorithms for the interpolation and approximation of images (see Sections III-D and IV-C). In particular, we have demonstrated that spline interpolators of a degree greater than one are always superior to short kernel cubic convolution: a popular high-quality method used for image interpolation. This is a result of practical relevance because spline interpolation for degrees less than four is not more expensive computationally. We have also used our error formulæ to design a new type of quasi-interpolator that is asymptotically optimal.

In a second paper [44], we will restrict somewhat the range of admissible approximating functions by requiring that they satisfy a two-scale relation. This further investigation will be more oriented toward the wavelet community; in particular, we will show how advantageous it can be to use spline wavelets instead of Daubechies' in applications in which approximation power is the key factor.

Appendix A
Proof of Theorem 3

We want to estimate $\varepsilon_s^2 = \int_0^T \langle |s(t) - \mathcal{Q}s(t)|^2 \rangle \frac{dt}{T}$. For this, we have to compute three terms since $\langle |s(t) - \mathcal{Q}s(t)|^2 \rangle = \langle s(t)^2 \rangle - 2\langle s(t)\mathcal{Q}_T s(t) \rangle + \langle |\mathcal{Q}s(t)|^2 \rangle$. We shall need the

$$P(z) = \frac{3}{16}\frac{z^4 + z^{-4} + 6552(z^3 + z^{-3}) + 331\,612(z^2 + z^{-2}) + 2\,485\,288(z + z^{-1}) + 4\,675\,014}{z^5 + z^{-5} + 196(z^4 + z^{-4}) + 10\,541(z^3 + z^{-3}) + 120\,608(z^2 + z^{-2}) + 467\,858(z + z^{-1}) + 736\,952}. \tag{38}$$

following results, which are easy to obtain:

$$\langle s(t)c_n\rangle = \int \rho(t-nT-\tau)\tilde{\varphi}\Big(\frac{\tau}{T}\Big)\,d\frac{\tau}{T}, \tag{39}$$

$$u_{n-n'} \stackrel{\text{def}}{=} \langle s_n c_{n'}\rangle = \frac{1}{2\pi}\int P_s(\omega)|\hat{\tilde{\varphi}}(\omega T)|^2 e^{(n-n')i\omega T}\,d\omega. \tag{40}$$

- We have that

$$\int_0^T \langle s(t)^2\rangle\,\frac{dt}{T} = \rho(0) = \frac{1}{2\pi}\int P_s(\omega)\,d\omega. \tag{41}$$

- $\langle s(t)\mathcal{Q}_T s(t)\rangle$ is computed as follows:

$$\begin{aligned}\int_0^T \langle s(t)\mathcal{Q}_T s(t)\rangle\,\frac{dt}{T} &= \sum_n \int_0^T \langle s(t)c_n\rangle\varphi\Big(\frac{t}{T}-n\Big)\,\frac{dt}{T} \qquad (42)\\ &= \iint \rho(t-\tau)\varphi\Big(\frac{t}{T}\Big)\tilde{\varphi}\Big(\frac{\tau}{T}\Big)\,\frac{dt\,d\tau}{T^2}\\ &= \frac{1}{2\pi}\int P_s(\omega)\hat{\varphi}(\omega T)^*\hat{\tilde{\varphi}}(\omega T)\,d\omega. \qquad (43)\end{aligned}$$

- $\langle|\mathcal{Q}_T s(t)|^2\rangle$ is computed as follows:

$$\begin{aligned}\langle|\mathcal{Q}_T s(t)|^2\rangle &= \sum_{n,n'} u_{n-n'}\int_0^T \varphi\Big(\frac{t}{T}-n\Big)\varphi\Big(\frac{t}{T}-n'\Big)\,\frac{dt}{T}\\ &= \sum_n u_n \int \varphi(x-n)\varphi(x)\,dx\\ &= \frac{1}{2\pi}\int \sum_n u_n e^{ni\omega}|\hat{\varphi}(\omega)|^2\,d\omega\\ &= \frac{1}{2\pi}\int P_s(\omega)|\hat{\tilde{\varphi}}(\omega T)|^2\hat{a}_\varphi(\omega T)\,d\omega. \qquad (44)\end{aligned}$$

Putting (41), (43), and (44) together, we find the announced result (18).

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] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978.
[2] J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imag.*, vol. MI-2, pp. 31–39, 1983.
[3] R. Bernstein and G. C. Stierhoff, "Precision processing of earth image data," *Amer. Sci.*, vol. 64, no. 5, pp. 500–508, Sept.–Oct. 1976.
[4] R. E. Crochiere and L. R. Rabiner, "Interpolation and decimation of digital signals—A tutorial review," *Proc. IEEE*, vol. 69, pp. 300–331, Mar. 1991.
[5] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay," *IEEE Signal Processing Mag.*, vol. 13, pp. 30–60, Jan. 1996.
[6] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.
[7] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, pp. 1153–1160, 1981.
[8] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 277–285, Mar. 1991.
[9] M. Unser, "On the optimality of ideal filters for pyramid and wavelet signal approximation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3591–3596, Dec. 1993.
[10] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[11] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.
[12] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
[13] S. Mallat, *A Wavelet Tour of Signal Processing*, San Diego, CA: Academic, 1998.
[14] R. Hummel, "Sampling for spline reconstruction," *SIAM J. Math. Anal.*, vol. 43, no. 2, pp. 278–288, 1983.
[15] M. Unser, A. Aldroubi, and M. Eden, "On the asymptotic convergence of B-spline wavelets to Gabor functions," *IEEE Trans. Inform. Theory*, vol. 38, pp. 864–872, Mar. 1992.
[16] 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.
[17] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, pp. 2915–2925, Nov. 1994.
[18] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspect of Functional Analysis*. Rome, Italy: Cremonese, 1971, pp. 796–830.
[19] C. de Boor and G. Fix, "Spline approximation by quasiinterpolants," *J. Approx. Theory*, vol. 8, pp. 19–45, 1973.
[20] C. K. Chui and H. Diamond, "A characterization of multivariate quasiinterpolation formulas and applications," *Numer. Math.*, vol. 57, pp. 105–121, 1990.
[21] M. Unser, "Approximation power of biorthogonal wavelet expansions," *IEEE Trans. Signal Processing*, vol. 44, pp. 519–527, Mar. 1996.
[22] 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.
[23] 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.
[24] ———, "Asymptotic error expansions for wavelet approximations of smooth functions II," *Numer. Math.*, vol. 68, no. 3, pp. 377–401, 1994.
[25] 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.
[26] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544–1576, Nov. 1992.
[27] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. XLI, pp. 909–996, Nov. 1988.
[28] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet decomposition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.
[29] Y. Meyer, *Ondelettes*. Paris, France: Hermann, 1990 (in French).
[30] C. de Boor, "Quasiinterpolation and approximation power of multivariate splines," in *Computation of Curves and Surfaces*, W. Dahmen *et al.*, Eds. Dordrecht, The Netherlands: Kluwer, 1990, pp. 313–345.
[31] 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.
[32] R. Q. Jia, "A counterexample to a result concerning controlled approximation," *Trans. Amer. Math. Soc.*, vol. 97, no. 4, pp. 647–654, Aug. 1986.
[33] R. Q. Jia and J. J. Lei, "Approximation by multiinteger translates of functions having global support," *J. Approx. Theory*, vol. 72, pp. 2–23, 1993.
[34] E. W. Cheney and W. A. Light, "Quasiinterpolation with basis functions having noncompact support," *Constr. Approx.*, vol. 8, pp. 35–48, 1992.
[35] C. de Boor, "The polynomials in the linear span of integer translates of a compactly supported function," *Constr. Approx.*, vol. 3, pp. 199–208, 1987.
[36] M. Unser, "Quasiorthogonality and quasiprojections," *Appl. Comput. Harmon. Anal.*, vol. 3, pp. 201–214, 1996.
[37] S. K. Park and R. A. Schowengerdt, "Image reconstruction by parametric cubic convolution," *Comput. Vision Graph. Image Process.*, vol. 23, pp. 258–272, 1983.
[38] A. Schaum, "Theory and design of local interpolators," *Comput. Vision Graph. Image Process.*, vol. 55, no. 6, pp. 464–481, Nov. 1993.

[39] 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.
[40] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Trans. Signal Processing*, vol. 41, pp. 834–848, Feb. 1993.
[41] M. Unser, P. Thévenaz, and L. Yaroslavsky, "Convolution-based interpolation for fast, high-quality rotation of images," *IEEE Trans. Image Processing*, vol. 4, pp. 1371–1381, Oct. 1995.
[42] M. J. Vrhel, C. Lee, and M. Unser, "Rapid computation of the continuous wavelet transform by oblique projections," *IEEE Trans. Signal Processing*, vol. 45, pp. 891–900, Apr. 1997.
[43] M. Unser, A. Aldroubi, and M. Eden, "Polynomial spline approximation: Filter design and asymptotic equivalence with shannon's sampling theorem," *IEEE Trans. Inform. Theory*, vol. 38, pp. 95–103, Jan. 1992.
[44] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part II—Wavelets," *IEEE Trans. Signal Processing*, this issue, pp. 2796–2806.

Thierry Blu (M'96) was born in Orléans, France, in 1964. He graduated from École Polytechnique, Paris, France, in 1986 and from Télécom Paris (ENST) in 1988. In 1996, he received the Ph.D. in electrical engineering from ENST for a study on iterated rational filter banks applied to wide-band audio coding.

He is currently with the Biomedical Imaging Group at the Swiss Federal Institute of Technology, Lausanne, on leave from France Télécom National Center for Telecommunications Studies, Issy-les-Moulineaux, France. His research interests include (multi-) wavelets, multiresolution analysis, multirate filter banks, approximation and sampling theory, and psychoacoustics.

Michael Unser (M'89–SM'94–F'99) was born in Zug, Switzerland, on April 9, 1958. He received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology (EPFL), Lausanne.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, where he was heading the Image Processing Group. He is now Professor and Head of the Biomedical Imaging Group, EPFL. His main research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing. He is the author of over 70 published journal papers in these areas.

Dr. Unser is a Member of the Image and Multidimensional Signal Processing Committee of the IEEE Signal Processing Society. He is on the editorial boards of *Signal Processing, The Journal of Visual Communication and Image Representation*, and *Pattern Recognition.* He was an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1992 to 1995 and the IEEE SIGNAL PROCESSING LETTERS from 1994 to 1998. He co-organized the 1994 IEEE-EMBS Workshop on Wavelets in Medicine and Biology and serves as conference chair for SPIE's Wavelet Applications in Signal and Image Processing, held annually since 1993. He received the Dommer Prize for excellence from the EPFL in 1981, the Research Prize of the Brown-Boveri Corporation for his thesis in 1984, and the IEEE Signal Processing Society's 1995 Best Paper Award (in the IMDSP technical area) for a TRANSACTIONS paper with A. Aldroubi and M. Eden on B-spline signal processing.