Quantitative Error Analysis for the Reconstruction of Derivatives

Laurent Condat, *Member, IEEE,* and Torsten Möller, *Member, IEEE*

***Abstract*—We present a general Fourier-based method which provides an accurate prediction of the approximation error, when the derivative of a signal $s(t)$ is continuously reconstructed from uniform point samples or generalized measurements on s. This formalism applies to a wide class of convolution-based techniques. It provides a key tool, the frequency error kernel, for designing computationally efficient reconstruction schemes which are near optimal in the least-squares sense.**

***Index Terms*—derivatives, reconstruction, sampling, interpolation, approximation, error analysis, frequency error kernel.**

I. Introduction

Reconstruction of a continuous function and its derivatives from a set of samples is one of the fundamental operations in signal processing, numerical analysis, and many other fields. In visualization, for instance, the gradient is employed in volume classification and shading [1]. It has to be evaluated at arbitrary locations and not only at the discrete points where the underlying signal has been sampled. Edge detection, segmentation, motion estimation and super-resolution are other applications where partial derivatives may be required at subpixel resolution.

A. Motivation

We denote by $s(t) \in L^2(\mathbb{R})$ a continuously defined function (the signal) which is prefiltered and sampled at uniform locations to yield the discrete measurements

$$u[k] = \int_{\mathbb{R}} s(t)\tilde{\varphi}\big(k - \frac{t}{T}\big)\, d\frac{t}{T} \quad \forall k \in \mathbb{Z}, \tag{1}$$

where T is the sampling step and the analysis function $\tilde{\varphi}(t)$ is, for instance, the impulse response of the acquisition device. This generalized sampling scenario encompasses the case where ideal point samples $u[k] = s(Tk)$ are available, simply by letting $\tilde{\varphi}(t)$ be the Dirac distribution $\delta(t)$.

The signal $s(t)$ is unknown and the sequence $u = (u[k])_{k\in\mathbb{Z}}$ represents the only available data. We are interested in constructing from u an estimate of the N-th derivative $s^{(N)}(t)$ of $s(t)$, for some integer $N \geq 1$. We look for a reconstruction in a *linear shift-invariant space* $\mathcal{V}_T(\varphi) = \mathrm{Span}(\{\varphi(\frac{t}{T} - k)\}_{k\in\mathbb{Z}})$ generated by the translates of a template function $\varphi(t) \in L^2(\mathbb{R})$:

$$f(t) = \frac{1}{T^N} \sum_{k\in\mathbb{Z}} c[k]\, \varphi(\frac{t}{T} - k) \qquad \forall t \in \mathbb{R}, \tag{2}$$

where the coefficients $c[k]$ are obtained by discrete filtering with the stable prefilter $p \in \ell^1$:

$$c[k] = (u * p)[k] \quad \forall k \in \mathbb{Z}. \tag{3}$$

L. Condat is with the Image Team of the GREYC laboratory, a joint CNRS-UCBN-ENSICAEN research unit in Caen, France. Contact: see http://www.greyc.ensicaen.fr/∼lcondat/. T. Möller is with the GrUVi Lab of the Simon Fraser University (SFU), Vancouver, CA. Contact: see http://www.cs.sfu.ca/∼torsten/.

This work was initiated during the visit of L. Condat at SFU, supported by a fellowship of the Computer Science Department of SFU and the NSERC.

Manuscript received XX 2010.

Using this general recipe for reconstruction, we denote by f_{app} an estimate of s (with $N = 0$ in (2)), while an estimate of the derivative $s^{(N)}$ is denoted by f_{der}. We remark that the reconstruction method involves a discrete prefiltering step followed by the fit of the continuous model itself. In practical applications, the prefiltering step is performed once. Its computation time is negligible in comparison with the many calls to (2) to evaluate f at the desired locations.

Estimating $s(t)$ itself is the classical problem of interpolation, for which there is a vast amount of literature; see e.g. the survey papers [2]–[5] and some recent developments [6]–[8]. Of course, once an estimate $f_{\mathrm{app}}(t)$ of $s(t)$ has been reconstructed, one can consider its derivative $f_{\mathrm{app}}^{(N)}(t)$ as a valid estimate of $s^{(N)}(t)$. But there is no *a priori* guarantee that whenever f_{app} is close to s in the least-squares sense, then $f_{\mathrm{app}}^{(N)}$ is close to $s^{(N)}$. Moreover, since efficiency considerations generally steer the design of the method, one may be interested in deriving direct estimation schemes of $s^{(N)}$, without the conceptual intermediary step of evaluating s, which unnecessarily constrains the conditions on accuracy and smoothness. The aim of this work is to provide a way to quantify the error between $s^{(N)}$ and its estimate f_{der}, so that the design of reconstruction schemes minimizing this error is made easy.

B. Related Work

There is a vast literature on designing so-called *digital differentiators*, which are digital filters estimating the derivative at the grid points Tk only, see e.g. [9] and references therein. In [10], pointwise estimates of the derivative are derived, which are optimal in the minimax sense. By contrast, we consider the context in which the derivative is reconstructed continuously in a shift-invariant space, so that it can be evaluated at every arbitrary location.

Shannon's theory provides an exact way to recover a bandlimited signal from its samples, using the sinc interpolator. Similarly, the "ideal" derivative reconstruction filter was shown in [11] to be the derivative of the sinc. However, its slow decay and the ringing artifacts it may introduce, prevent its practical use. Moreover, for non-bandlimited signals, the sinc-based theory is not valid any more [3], [12], [13]. That is why practitioners rely on convolutions with more localized kernels having compact support, like windowed sinc [14], [15], splines and other piecewise polynomial functions [2], [11]. In none of these works, there is an analytic comparison of different filters nor a quantitative analysis of the estimation error.

In [16], the second author and its collaborators developed tools and derived absolute error bounds for the spatial analysis of both interpolation and derivative filters of arbitrary order. Approximation theory also provides a general framework, which focuses on the asymptotic error behavior of the reconstruction method as $T \to 0$ [17], [18]. These qualitative error bounds are generally not sharp enough to be of direct use to practitioners and accurate ways of predicting the approximation error are desirable, so that algorithms can be compared and the parameters φ, p, $\tilde{\varphi}$, T can be chosen appropriately. For this, Blu *et al.* proposed a remarkable Fourier-based method which provides an accurate estimate of the approximation error, with a wide range of applicability [18]. Their approach makes the design of reconstruction algorithms simple and accurate, and it is at the heart of recent developments in interpolation theory [6]–[8]. In this paper, we extend this theory to the setting of derivative reconstruction.

C. Notations and Paper Organization

The Fourier transform of a function $f(t)$ is denoted by $\hat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-j\omega t}dt$. We require $\tilde{\varphi}$ in (1) to have a well-defined bounded Fourier transform.

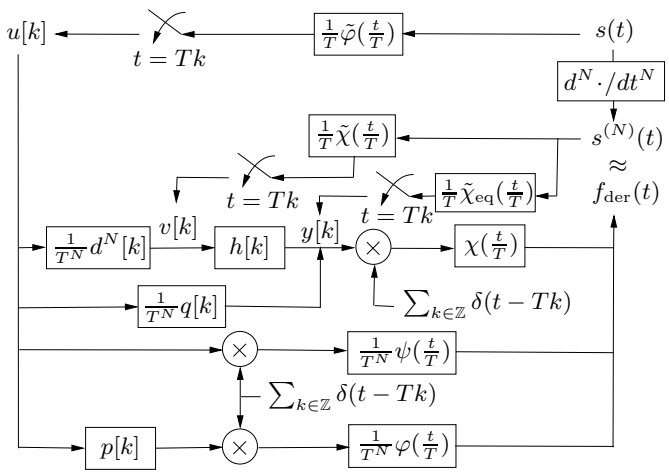

Fig. 1. Flowgraph of the different equivalent procedures for the sampling process $s \mapsto u$ and the derivative reconstruction process $u \mapsto f_{\mathrm{der}}$.

The $\mathcal{Z}$-transform of a discrete signal $v = (v[k])_{k\in\mathbb{Z}}$ is $V(z) = \sum_{\mathbb{Z}} v[k]z^{-k}$ and its Fourier transform is $\hat{v}(\omega) = V(e^{j\omega})$.

For any real $r > 0$, the Sobolev space W_2^r is the set of functions f such that $\int_{\mathbb{R}}(1+\omega^2)^r|\hat{f}(\omega)|^2 d\omega < \infty$. Therefore, the Sobolev regularity of f is the maximum value of r such that $f \in W_2^r$.

$\mathcal{P}^n$ is the space of polynomial functions of degree at most $n \in \mathbb{N}$. We define the causal B-spline $\beta_+^n(t)$ of degree $n \in \mathbb{N}$ by $\beta_+^0 = \mathbb{1}_{[0,1)}$ and $\beta_+^n = \beta_+^{n-1} * \beta_+^0$. The centered B-spline of degree n is $\beta^n(t) = \beta_+^n(t + \frac{n+1}{2})$.

We introduce the dual ϕ_d of a function $\phi \in L^2(\mathbb{R})$ by $\widehat{\phi}_d(\omega) = \hat{\phi}(\omega)^*/\hat{a}_\phi(\omega)$, where the discrete autocorrelation filter a_ϕ is defined by $a_\phi[k] = \int_{\mathbb{R}} \phi(t)\phi(t-k)dt$ and the star is for complex conjugation.

The different functions and filters used throughout the paper are illustrated by the flowgraph in Fig. 1.

The outline of this paper is as follows. In Sect. II, we introduce the *frequency error kernel*, which is the cornerstone for quantifying the error between a function and its estimate from discrete measurements. We present our main results in Sect. III, based on a new error kernel dedicated to the reconstruction of derivatives. Finally, in Sect. IV, we discuss the consequences of the formalism for the design of efficient reconstruction methods.

II. The Frequency Error Kernel

An important result of Blu *et al.* is that the error $\|s - f_{\mathrm{app}}\|_{L^2}$ between s and its estimate reconstructed using (2) can be predicted very accurately by the estimate [18]:

$$\eta_s(T) = \sqrt{\frac{1}{2\pi}\int_{\mathbb{R}} |\hat{s}(\omega)|^2 E(T\omega)\, d\omega}, \tag{4}$$

using the *frequency error kernel* defined by

$$E(\omega) = \underbrace{1 - \frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}}_{E_{\min}(\omega)} + \underbrace{\hat{a}_\varphi(\omega)\left|\widehat{\tilde{\varphi}}(\omega)\hat{p}(\omega) - \widehat{\varphi}_d(\omega)\right|^2}_{E_{\mathrm{res}}(\omega)}. \tag{5}$$

Remarkable properties of the global error estimate $\eta_s(T)$ include its exactness for bandlimited signals and for the average of the true approximation error over all possible shifts of the input function s. In the general case, we have the approximation $\|s - f_{\mathrm{app}}\|_{L^2} = \eta_s(T) + o(T^r)$, assuming that s has Sobolev regularity $r > \frac{1}{2}$. Moreover, in a stochastic framework where s is a realization of a random stationary process instead of a deterministic function of L^2, $\eta_s(T)$ is the exact expression of the time averaged expectation of the quadratic pointwise error, by replacing the energy $|\hat{s}(\omega)|^2$ by the power spectrum density $\hat{c}_s(\omega)$ in (4). These properties and several others are detailed in [18] and [19].

In practical situations, $|\hat{s}(\omega)|^2$ or $\hat{c}_s(\omega)$ is unknown, but the multiplicative form in the integral (4) ensures that the error is small if $E(\omega)$ is close to zero. Hence, the frequency error kernel is a tool of choice for characterizing a reconstruction scheme. More precisely, the value $E(\omega)$ at a given frequency ω can be interpreted as the average power of the approximation error, in case $s(t)$ is the pure unit sinusoid $e^{j\omega t/T}$ [18, Thm. 3]. Therefore, the study of $E(\omega)$ allows to characterize the behavior of a reconstruction method at different frequency components. For instance, $E(\omega)$ for ω close to π indicates to which extent the salient features, fine details and textures are preserved and aliasing is enhanced. By contrast, an asymptotic study for ω around 0 characterizes the reconstruction quality for the low frequency content of the signal s.

Thus, φ and p can be tuned to minimize the error kernel, so that the reconstruction quality is improved for virtually every function s [6]–[8], [20]. Given the reconstruction space $\mathcal{V}_T(\varphi)$, the error kernel $E(\omega)$ attains its minimum value $E_{\min}(\omega)$, for every $\omega \in \mathbb{R}$, when f_{app} is the minimum error reconstruction of s in $\mathcal{V}_T(\varphi)$; that is, its orthogonal projection onto $\mathcal{V}_T(\varphi)$. Thus, the prefilter p should be designed so that $E(\omega)$ is close to $E_{\min}(\omega)$, in order for the method to behave like this optimal, but generally unattainable, least-squares approximation [8], [18].

III. Computation of the Error Estimates

We first define the function $\psi(t)$ by

$$\psi(t) = \sum_{k\in\mathbb{Z}} p[k]\,\varphi(t-k) \;\Leftrightarrow\; \hat{\psi}(\omega) = \hat{p}(\omega)\hat{\varphi}(\omega). \tag{6}$$

Then, $f_{\mathrm{der}}(t) = \frac{1}{T^N}\sum_{k\in\mathbb{Z}} u[k]\,\psi(\frac{t}{T} - k)$. Although in practice φ is chosen with compact support, so that (2) is computationally attractive, the function ψ, which is the impulse response of the reconstruction operation, can have infinite support.

We assume that the following equivalent conditions on ψ are satisfied, so that f_{der} does not blow up as T tends to zero (because of the $1/T^N$ factor in (2)):

$$f_{\mathrm{der}} = 0 \text{ if } s \in \mathcal{P}^{N-1}, \tag{7}$$

$$\Leftrightarrow \sum_{k\in\mathbb{Z}} P(k)\,\psi(t-k) = 0, \quad \forall t \in \mathbb{R},\ \forall P \in \mathcal{P}^{N-1}, \tag{8}$$

$$\Leftrightarrow \widehat{\psi}^{(n)}(2k\pi) = 0, \quad \forall k \in \mathbb{Z},\ \forall n = 0\ldots N-1. \tag{9}$$

Although these equivalences may be classical in harmonic analysis, we rederive them in the Appendix A for sake of completeness.

We can remark that, with the naive approach which consists in reconstructing an estimate f_{app} of s and taking its derivative $f_{\mathrm{der}} = f_{\mathrm{app}}^{(N)}$, it is necessary that the approximation process $s \mapsto f_{\mathrm{app}}$ is a quasi-projection of order N [8], for the conditions (7)-(9) to be satisfied.

Our second requirement is that ψ can be decomposed as

$$\psi(t) = \sum_{k\in\mathbb{Z}} q[k]\,\chi(t-k) \;\Leftrightarrow\; \hat{\psi}(\omega) = \hat{q}(\omega)\hat{\chi}(\omega), \tag{10}$$

where $q \in \ell^1$ is a discrete filter and the integer translates of the function $\chi \in L^2$ form a Riesz basis; that is, there exist two constants $B \geq A > 0$ such that $A \leq \hat{a}_\chi(\omega) \leq B$ almost everywhere. Thus, each function of the reconstruction space $\mathcal{V}_T(\chi)$ has a unique and stable expansion in terms of the shifts of χ. Note that this condition is not restrictive; in particular, there is no requirement that χ be compactly supported, even if φ is.

Our last requirement is that $s \in L^2$ has Sobolev regularity $r > N + \frac{1}{2}$, so that s has at least N continuous derivatives in L^2.

The Riesz basis condition together with the requirement (9) imply that 0 is a root of q of multiplicity N. In other words, there exists a stable filter $h \in \ell^1$ such that $q = h * d^N$, where the iterated causal finite difference filter d^N is defined by $D^N(z) = (1 - z^{-1})^N$. Thus, we essentially decomposed $\psi = p * \varphi$ into a moving difference filter q and a well posed reconstruction kernel χ. This allows us to recast derivative reconstruction into the framework summarized in Sect. II, by reasoning on χ instead of φ. For this, we need the following property:

Lemma. We define $v = \frac{1}{T^N}\, u * d^N$. Then, $v[k] = \int_{\mathbb{R}} s^{(N)}(t)\tilde{\chi}(k - \frac{t}{T}) d\frac{t}{T}$, where $\tilde{\chi} = \tilde{\varphi} * \beta_+^{N-1}$.

This important lemma, proved in Appendix B, allows us to interpret the filtered samples $u[k]$ as generalized samples of the derivative $s^{(N)}$. Hence, we now have all the ingredients to formulate the following results, proved in Appendix B.

Theorem 1. $\|s^{(N)} - f_{\mathrm{der}}\|_{L^2} = \eta_{s^{(N)}}(T) + o(T^{r-N})$, where

$$\eta_{s^{(N)}}(T) = \left(\frac{1}{2\pi} \int_{\mathbb{R}} \underbrace{|\hat{s}(\omega)|^2 \omega^{2N}}_{\left|\widehat{s^{(N)}}(\omega)\right|^2} E(T\omega)\, d\omega \right)^{1/2} \tag{11}$$

and the new frequency error kernel characterizing derivative reconstruction is

$$E(\omega) = \underbrace{1 - \frac{|\widehat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}}_{E_{\min}(\omega)} + \underbrace{\hat{a}_\varphi(\omega) \left| \widehat{\tilde{\varphi}}(\omega)\hat{p}(\omega)\frac{1}{(j\omega)^N} - \widehat{\varphi}_d(\omega) \right|^2}_{E_{\mathrm{res}}(\omega)}. \tag{12}$$

Moreover, the correction term $o(T^{r-N})$ vanishes if s is bandlimited in $[-\frac{\pi}{T}, \frac{\pi}{T}]$, or if $\tilde{\varphi}$ and φ are both bandlimited in $[-\pi, \pi]$.

Theorem 2. In a stochastic framework where s is a realization of a random stationary process with power spectrum density $\hat{c}_s(\omega)$, instead of a deterministic function of L^2, we have

$$\eta_{s^{(N)}}(T) = \left(\frac{1}{T} \int_0^T \mathcal{E}\{|s^{(N)}(t) - f_{\mathrm{der}}(t)|^2\} dt \right)^{1/2}, \tag{13}$$

by replacing $|\hat{s}(\omega)|^2$ by $\hat{c}_s(\omega)$ in (11).

In Fig. 2, we give an example of the error estimate $\eta_{s'}(T)$ for a practical experiment in which we reconstruct the derivative of a Gaussian from point samples. This shows that $\eta_{s^{(N)}}(T)$ is an accurate, shift invariant approximation of the true error. We note that if $s \in W_2^r$ for every $r \geq 0$, which is the case in our example, then the difference between $\eta_{s^{(N)}}(T)$ and the true error decays faster than every polynomial in T as T tends to zero. This means that $\eta_{s^{(N)}}(T)$ can be considered as the exact value of the error in some non-infinitesimal interval; e.g. for $T \in [0, 0.9]$ in Fig. 2. In the general case, $\eta_{s^{(N)}}(T)$ is a reliable estimate of the error for practical values of T. By contrast, error analysis approaches based on Taylor series only apply to the asymptotic regime where s is highly oversampled.

IV. Asymptotic Approximation Performance

In this section, we focus on the reconstruction of lowpass signals; that is, we assume that $s^{(N)}$ has most of its energy around $\omega = 0$. This is the case for natural images, at least for $N = 1$. For other types of signals, the same study could be performed around another frequency than 0, with exponential splines taking the role of the B-splines [21].

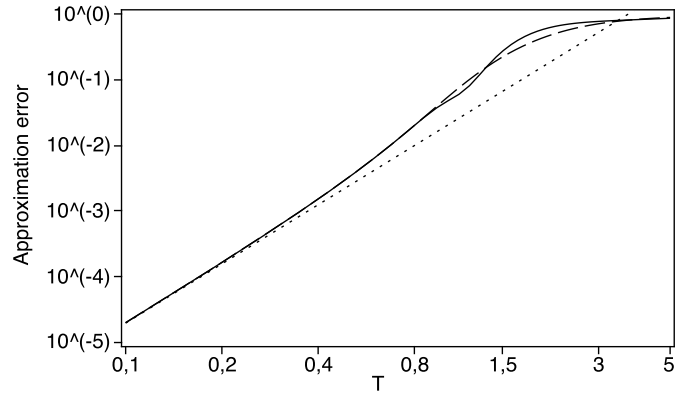

Fig. 2. Approximation error as a function of the sampling step T. We estimate the first derivative $s'(t)$ of $s(t) = e^{-\frac{(t-1)^2}{2}}$ by the first derivative of the cubic spline interpolating the point samples $u[k] = s(Tk)$. Thus, we have $\tilde{\varphi} = \delta$, $\varphi = (\beta^3)'$ et $P(z) = 6/(z + 4 + z^{-1})$. The error estimate $\eta_{s'}(T)$ (dashed line) is close to the true error $\|s' - f_{\mathrm{der}}\|_{L^2}$ (solid line). The dotted line is the asymptote $C\|s^{(4)}\|_{L^2}T^3$, where the asymptotic constant $C = 1/\sqrt{30240}$ is obtained from the Taylor development $E(\omega)^{1/2} \sim C\omega^3$.

A. The Approximation Order

From Theorem 1, due to the closed form of $\eta_{s^{(N)}}(T)$, it is easy to expand this estimate in a power series of T to obtain the exact behavior of the error as $T \to 0$. Specifically, if $s^{(N)}$ has at least Sobolev regularity L, we have the equivalence

$$E(\omega)^{1/2} \sim C\,\omega^L \quad \text{as } \omega \to 0 \tag{14}$$

$$\text{iff} \quad \|s^{(N)} - f_{\mathrm{der}}\|_{L^2} \sim C\|s^{(N+L)}\|_{L^2}T^L \quad \text{as } T \to 0. \tag{15}$$

In that case, we speak about a Lth-order approximation scheme. When most of the spectral energy of the signal is concentrated in the neighborhood of $\omega = 0$, like with natural images, the approximation order L is the most crucial determinant of the reconstruction quality and should be chosen as large as possible. To have an approximation order L, it is necessary that χ satisfies the Strang-Fix conditions of order L [17]:

$$\hat{\chi}(0) \neq 0 \quad \text{and} \quad \hat{\chi}^{(n)}(2k\pi) = 0 \;\text{ for } \begin{cases} k \neq 0 \\ n = 0 \ldots L-1 \end{cases}. \tag{16}$$

It was shown in [6] that a function χ with approximation L has a support size $S \geq L$ with equality iff χ is a MOMS. Therefore, the reconstruction schemes having the optimal tradeoff between the reconstruction quality and the computational complexity are obtained by choosing the reconstruction kernel φ as a MOMS.

On this point, there is an interesting connection with wavelet theory. The equivalent conditions (7),(8),(9) imply that ψ has N vanishing moments: $\int_{\mathbb{R}} t^n \psi(t) dt = 0,\ \forall n = 0 \ldots N-1$. Therefore, ψ is similar to a wavelet function. It is known that wavelets behave like differentiators [22]. Moreover, in [23], it is shown that there is a B-spline function at the heart of each scaling function and wavelet associated with a multiresolution analysis. In our setting, there is no such requirements like the two-scale relations associated with multiresolution. Also, the building component of ψ carrying the approximation order is χ. Therefore, these are the MOMS functions, a broader class than the B-splines, which naturally appear when designing χ.

B. The choice of the Prefilter

We now assume that φ is fixed. This determines χ and its approximation order L (eqn. (16)). Then, we have to choose p so as to

exploit at best the properties of the reconstruction space $\mathcal{V}_T(\varphi)$; that is, so that the scheme has approximation order L (eqns (14),(15)). In fact, $f_{\rm der} \in \mathcal{V}_T(\varphi) \subset \mathcal{V}_T(\chi)$ and p controls which approximation of $s^{(N)}$ in $\mathcal{V}_T(\varphi)$ is picked by the method. The best possible reconstruction is the orthogonal projection of $s^{(N)}$ in $\mathcal{V}_T(\varphi)$. The error between this optimal approximation and $s^{(N)}$ decays like T^L and is characterized by the error kernel $E_{\min}$ given in (12). Thus, the reconstruction scheme has approximation order L if and only if

$$E(\omega) = E_{\min}(\omega) + O(\omega^{2L}). \tag{17}$$

To characterize p more precisely, we define the function

$$\tilde{\chi}_{\rm eq}(t) = \sum_{k\in\mathbb{Z}} h[k]\tilde{\chi}(t-k), \tag{18}$$

where we recall that $\tilde{\chi} = \tilde{\varphi} * \beta_+^{N-1}$ and $\hat{p}.\hat{\varphi} = \hat{h}.\hat{d}^N.\hat{\chi}$. So, we have $f_{\rm der}(t) = \sum_{k\in\mathbb{Z}} y[k]\chi(\frac{t}{T}-k) \quad \forall t \in \mathbb{R}$, with $y[k] = \int_{\mathbb{R}} s^{(N)}(t)\tilde{\chi}_{\rm eq}(k-\frac{t}{T})d\frac{t}{T} \quad \forall k \in \mathbb{Z}$. Then, the reconstruction scheme has approximation order L if and only if $\tilde{\chi}_{\rm eq}$ and χ are *quasi-biorthonormal* with order L [18]; that is,

$$\widehat{\tilde{\chi}}_{\rm eq}(\omega)\hat{\chi}(\omega+2k\pi) = \delta_k + O(\omega^L), \quad \forall k \in \mathbb{Z}, \tag{19}$$

which is equivalent for $\tilde{\chi}_{\rm eq}$ and χ_d to have the same moments up to order L:

$$\int_{\mathbb{R}} t^n \tilde{\chi}_{\rm eq}(t)dt = \int_{\mathbb{R}} t^n \chi_d(t)dt, \qquad \forall n = 0\ldots L-1. \tag{20}$$

We can express these conditions in terms of $\tilde{\varphi}$, φ and p as follows:

$$\frac{\hat{p}(\omega)}{(j\omega)^N}\widehat{\tilde{\varphi}}(\omega)\hat{\varphi}(\omega+2k\pi) = \delta_k + O(\omega^L), \quad \forall k \in \mathbb{Z}, \tag{21}$$

or, equivalently,

$$\hat{p}(\omega)\widehat{\tilde{\varphi}}(\omega) = \frac{(j\omega)^N}{\hat{\varphi}(\omega)} + O(\omega^{L+N}). \tag{22}$$

We note that (22) can be obtained directly from (17).

Thus, it is preferable to choose p so that these quasi-biorthonormality conditions are satisfied. There is a great freedom in this respect, since only the $L+N$ linear constraints given by (22) have to be satisfied for the scheme to have the maximal approximation order, given φ.

V. Conclusion

We introduced a generic Fourier methodology to evaluate the quality of shift-invariant methods that continuously reconstruct the derivative of a function from discrete measurements. In our future work, we will focus on the use of this theory to design efficient reconstruction schemes [24]. Some reconstruction methods for the second derivative are compared in [25] to illustrate our approach. We will also investigate the extension of the formalism to noisy measurements [26], [27].

Since the frequency error kernel can be defined for multi-dimensional signals on lattices, like in [20], the extension of this work to the evaluation of partial derivatives of multi-dimensional signals is straightforward [24]. Interesting applications include finite difference methods and the numerical resolution of PDEs.

Appendix A
Proof of the Equivalence of (7),(8),(9)

We first prove that (8) $\Leftrightarrow$ (9). Let n be in $0\ldots N-1$. We define $y_n = (y_n[k])_{k\in\mathbb{Z}}$ by $y_n[k] = \widehat{\psi}^{(n)}(2k\pi)$. Then, $y_n[k] = 0 \ \forall k \Leftrightarrow \widehat{y_n} = 0 \Leftrightarrow \sum_k (t-k)^n\psi(t-k) = 0 \ \forall t$, where the last equivalence is Poisson's sum formula applied to the function $t^n\psi(t)$. So, (9) $\Leftrightarrow \sum_{k\in\mathbb{Z}} Q(t-k)\,\psi(t-k) = 0, \ \forall t \in \mathbb{R}, \ \forall Q \in \mathcal{P}^{N-1}$. This is equivalent to (8) by letting, for any t, $Q(X)$ be $P(t-X)$. □

We then show that (8) $\Rightarrow$ (7). We assume (8). Let $s(t) = \sum_{n=0}^{N-1} a_n t^n$ be in $\mathcal{P}^{N-1}$ and k be in $\mathbb{Z}$. We introduce the function $g(t) = \frac{1}{T}s(t)\tilde{\varphi}(k-\frac{t}{T})$. Then, $u[k] = \hat{g}(0) = \sum_{n=0}^{N-1} a_n j^n \frac{d^n}{d\omega^n}\{e^{-j\omega Tk}\widehat{\tilde{\varphi}}(T\omega)^*\}|_{\omega=0}$ is a polynomial in k of degree $N-1$. Applying (8) to this polynomial P, we get $f_{\rm der} = 0$. □

We don't give the details of (7) $\Rightarrow$ (8). This involves exhibiting, for a given $P \in \mathcal{P}^{N-1}$, a polynomial $s \in \mathcal{P}^{N-1}$ such that $u[k] = P(k)$, which is always possible.

Appendix B
Proofs of the Results in Sect. III

We first prove the lemma. We define the function $s_T(t) = s(Tt)$. Then, for every n in $1\ldots N$, $s_T^{(n)} * \beta_+^0(t) = \int_{t-1}^t s^{(n)}(x)dx = s_T^{(n-1)}(t) - s_T^{(n-1)}(t-1) = s_T^{(n-1)} * \big(\delta - \delta(\cdot-1)\big)(t)$. By recurrence, using the property $\beta_+^n = \beta_+^{n-1} * \beta_+^0$, we get $s_T^{(N)} * \beta_+^{N-1} * \tilde{\varphi} = s_T * \big(\delta - \delta(\cdot-1)\big)^N * \tilde{\varphi}$. Therefore, since $u[k] = s_T * \tilde{\varphi}\,(k)$, we have $s_T^{(N)} * \tilde{\chi}\,(k) = T^N \int_{\mathbb{R}} s^{(N)}(t)\tilde{\chi}(k-\frac{t}{T})d\frac{t}{T} = u * d^N\,[k]$. □

The theorems 1 and 2 are adaptations of [18, Theorem 1] and [18, Theorem 3], respectively, by making the substitutions $s \to s^{(N)}$, $u \to v$, $\varphi \to \chi$, $\tilde{\varphi} \to \tilde{\chi}$, $p \to h$. Thus, the theorems are satisfied with the following expression of $E(\omega)$ in (11):

$$E(\omega) = 1 - \frac{|\hat{\chi}(\omega)|^2}{\hat{a}_\chi(\omega)} + \hat{a}_\chi(\omega)\big|\widehat{\tilde{\chi}}(\omega)\hat{h}(\omega) - \widehat{\chi}_d(\omega)\big|^2. \tag{23}$$

Then, we obtain (12) by substituting in (23) the equalities $\hat{\chi} = \frac{\hat{p}}{\hat{q}}\hat{\varphi}$, $\hat{a}_\chi = |\frac{\hat{p}}{\hat{q}}|^2\hat{a}_\varphi$, $\widehat{\tilde{\chi}} = \widehat{\tilde{\varphi}}.\widehat{\beta_+^{N-1}}$ with $\widehat{\beta_+^{N-1}}(\omega) = \hat{d}^N(\omega)/(j\omega)^N$, $\hat{q} = \hat{h}.\hat{d}^N$. □

References

[1] G. Kindlmann and J. W. Durkin, "Semi-automatic generation of transfer functions for direct volume rendering," in *Proc. of the IEEE Symposium on Volume Visualization (VVS)*, 1998, pp. 79–86.

[2] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 739–758, Jul. 2000.

[3] M. Unser, "Sampling—50 Years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.

[4] T. M. Lehmann, C. Gönner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1049–1075, Nov. 1999.

[5] E. Meijering, W. Niessen, and M. Viergever, "Quantitative evaluation of convolution-based methods for medical image interpolation," *Medical Image Analysis*, vol. 5, no. 2, pp. 111–126, Jun. 2001.

[6] T. Blu, P. Thévenaz, and M. Unser, "MOMS: Maximal-order interpolation of minimal support," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1069–1080, Jul. 2001.

[7] ——, "Linear interpolation revitalized," *IEEE Trans. Image Process.*, vol. 13, no. 5, pp. 710–719, May 2004.

[8] L. Condat, T. Blu, and M. Unser, "Beyond interpolation: Optimal reconstruction by quasi-interpolation," in *Proc. of IEEE ICIP*, vol. 1, Sep. 2005, pp. 33–36.

[9] I. Selesnick, "Maximally flat low-pass digital differentiators," *IEEE Trans. Circuits Syst. II*, vol. 49, no. 3, pp. 219–223, Mar. 2002.

[10] H. Kirshner and M. Porat, "On derivative approximation from sampled data," in *Proc. of SampTA*, Jun. 2007.

[11] M. J. Bentum, T. Malzbender, and B. B. Lichtenbelt, "Frequency analysis of gradient estimators in volume rendering," *IEEE Trans. Vis. Comput. Graphics*, vol. 2, no. 3, pp. 242–254, Sep. 1996.

[12] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 8, pp. 959–969, Aug. 1998.

[13] S. Ramani, D. Van De Ville, T. Blu, and M. Unser, "Nonideal sampling and regularization theory," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1055–1070, Nov. 2008.

[14] M. E. Goss, "An adjustable gradient filter for volume visualization image enhancement," in *Proc. of Graphics Interface*, 1994, pp. 67–74.
[15] T. Theußl, H. Helwig, and E. Gröller, "Mastering windows: improving reconstruction," in *Proc. of the IEEE Symposium on Volume Visualization (VVS)*, 2000, pp. 101–108.
[16] T. Möller, R. Machiraju, K. Mueller, and R. Yagel, "Evaluation and design of filters using a Taylor series expansion," *IEEE Trans. Vis. Comput. Graphics*, vol. 3, no. 2, pp. 184–199, April-June 1997.
[17] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspects of Functional Analysis*, Cremonese, Rome, Italy, 1971, pp. 796–830.
[18] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I-interpolators and projectors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2783–2806, Oct. 1999.
[19] M. Jacob, T. Blu, and M. Unser, "Sampling of periodic signals: A quantitative error analysis," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1153–1159, May 2002.
[20] L. Condat and D. Van De Ville, "Quasi-interpolating spline models for hexagonally-sampled data," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1195–1206, May 2007.
[21] M. Unser and T. Blu, "Cardinal exponential splines: Part I—Theory and filtering algorithms," *IEEE Trans. Image Process.*, vol. 53, no. 4, pp. 1425–1438, Apr. 2005.
[22] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1999.
[23] M. Unser and T. Blu, "Wavelet theory demystified," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 470–483, Feb. 2003.
[24] U. R. Alim, T. Möller, and L. Condat, "Gradient estimation revitalized," *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE Visualization 2010)*, vol. 16, no. 6, Nov.-Dec. 2010, to appear.
[25] L. Condat and T. Möller, "Quantitative error analysis for the reconstruction of derivatives," GREYC lab., Caen, France, Tech. Rep. hal-00462203, Sep. 2009.
[26] D. den Hertog, R. Brekelmans, L. Driessen, and H. Hamers, "Gradient estimation schemes for noisy functions," Tilburg University, Center for Economic Research, Tech. Rep. 12, 2003.
[27] Y. C. Eldar and M. Unser, "Nonideal sampling and interpolation from noisy observations in shift-invariant spaces," *IEEE Trans. Image Process.*, vol. 54, no. 7, pp. 2636–2651, Jul. 2006.