# Efficient Wavelet Prefilters
# with Optimal Time-Shifts

Stefan Ericsson, and Niklas Grip *Member, IEEE*

*Abstract*— A *wavelet prefilter* maps sample values of an analyzed signal to the scaling function coefficient input of standard discrete wavelet transform (DWT) algorithms. The prefilter is the inverse of a certain *postfilter* convolution matrix consisting of integer sample values of a noninteger-shifted wavelet scaling function. For the prefilter and the DWT algorithms to have similar computational complexity, it is often necessary to use a "short enough" approximation of the prefilter. In addition to well-known quadrature formula and identity matrix prefilter approximations, we propose a Neumann series approximation, a band matrix truncation of the optimal prefilter and derive simple formulas for the operator norm approximation error. This error shows a dramatic dependence on how the postfilter noninteger shift is chosen.

We explain the meaning of this shift in practical applications, describe how to choose it and plot optimally shifted prefilter approximation errors for 95 different Daubechies, Symlet and B-spline wavelets.

Whereas the truncated inverse is overall superior, the Neumann filters are by far the easiest ones to compute and for some short support wavelets, they also give the smallest approximation error. For example, for Daubechies 1–5 wavelets the simplest Neumann prefilter provide an approximation error reduction corresponding to 100–10 000 times oversampling in a non-prefiltered system.

*Index Terms*— Biorthogonal wavelet, prefilter, initialization, Neumann series, quadrature mirror filter, quadrature formula, Lagrange interpolant, Sard optimal, pyramid algorithm, FWT algorithm, Daubechies wavelet, Symlet, B-spline wavelet, sampling.

## I. INTRODUCTION

**M**ODERN wavelet theory is based on a division of $L_2(\mathbb{R})$ into subspaces $V_j$ spanned by integer-translated and dilated copies $\varphi_{j,k} \stackrel{\text{def}}{=} 2^{j/2}\varphi(2^j \cdot -k)$ of a *scaling function* $\varphi$, that will be assumed to have *compact support* in this paper. For any $J \in \mathbb{Z}_+$ and $f_J(x) = \sum_{k\in\mathbb{Z}} a_{J,k}\varphi_{J,k}(x) \in V_J$ with $x$ restricted to some finite interval $I$, the index set can be reduced to a set $I_J$ consisting of those integers for which the support of $\varphi_{J,k}$ overlap $I$:

$$f_J(x) = \sum_{k \in I_J} a_{J,k}\varphi_{J,k}(x), \qquad \forall f \in V_J, \quad \forall x \in I \quad (1)$$

with index set size $|I_J| = 2^J |I| + M = O(2^J)$ and $M \in \mathbb{Z}$ depending on the support of $\varphi$. For unique definition of the coefficients also if only the restriction of $f_J$ to $I$ is known, there are standard approaches like zeropadding, periodic extension or a more complicated modification of all basis functions whose support overlap the endpoints of $I$ [1, Section 7.5].

If $\varphi$ and the subspaces $V_j$ satisfy certain *multiresolution analysis* properties (see, e.g., [1], [2], [3]), then there is a corresponding *mother wavelet* $\psi$ and an orthonormal basis

$$\varphi_n \stackrel{\text{def}}{=} \varphi(\cdot - n) \quad \text{and} \quad \psi_{j,k} \stackrel{\text{def}}{=} 2^{j/2}\psi(2^j \cdot -k) \quad (2)$$

for $V_J$ with $j = 0, 1, \ldots, J - 1$ and $k, n \in \mathbb{Z}$. The *discrete wavelet transform* computes the coefficients $a_{0,n}$ and $d_{j,k}$ necessary for decomposing $f_J$ into a wavelet series expansion

$$f_J(x) = \sum_{n\in I_0} a_{0,n}\varphi(x-n) + \sum_{j=0}^{J-1}\sum_{k\in I_j} d_{j,k}\psi_{j,k}(x), \qquad \forall x \in I, \quad (3)$$

where $|I_j| = O(2^j)$ with exact length depending on the length of the supports of $\varphi$ and $\psi$. These coefficients are usually computed with the very fast pyramid algorithm (also known as the fast wavelet transform or Mallat's algorithm) or the even faster lifting scheme approach (see, e.g, [1], [3], [4]). Both algorithms use the coefficients $a_{J,k}$ of (1) as input (see also Figure 1). For a more complete treatment of both the pyramid algorithm and wavelet theory in general, see, e.g., [1], [2], [3].

The pyramid algorithm can very well be used in a purely discrete setting, for example when designing a filter bank to produce zero distortion effects, alias cancellation or cancellations of discrete polynomials in the highpass channel. Quite many applications, however, include analysis of sample values of some non-discrete signal $f$ or of its orthogonal projection $f_J$ on $V_J$. Figure 1 shows all computational steps involved both in the analysis and the synthesis.

Sections II–III are devoted to the problem of how to efficiently compute the coefficients $a_{J,k}$ (the vector $\mathbf{a}_J$ in Figure 1) from a vector $\mathbf{f}_J$ of sample values of $f_J$. We derive a solution in the form of a convolution matrix multiplication

$$\mathbf{a}_J = \Phi^{-1}\mathbf{f}_J. \quad (4)$$

Practical restrictions such as time-, memory- or chip area-constraints may necessitate a sparse matrix approximation $\Phi^{-1}_{\text{approx}} \approx \Phi^{-1}$ (or corresponding short enough filter) to assure roughly the same computational complexity as for the pyramid algorithm. We describe some traditional such approximations and compare them with a band matrix truncation of $\Phi^{-1}$ and with a Neumann series family of approximations.

For this comparison, we derive simple formulas for the operator norm approximation errors in Section IV and plot these errors for a large number of Daubechies, Symlet and B-spline wavelets in Section V.

*Example 1:* In a wavelet-based multicarrier signal transmission system, the transmitted signal is by construction a

function $s \in V_J$. Hence the received signal is $f = s + n$, where $n$ is noise. Thus the projection $\mathrm{P}_J$ (computed, e.g., as described in Remark 1) will simply remove some of the noise. For this particular application, basis functions of Gabor type are commonly used today, for example in OFDM (for mobile radio channels and digital audio broadcast) and DMT (for wireline systems), but wavelets can also be used. For more details and further references, see, e.g., [5, Section 2.3].

*Remark 1:* Under certain conditions on $x_0$ and $\varphi$, the initial projection $\mathrm{P}_J$ in Figure 1, assures that $f_J$ can be reproduced from the sample values $\left( f_J(2^{-J}(x_0 + k)) \right)_{k \in \mathbb{Z}}$ (see, e.g., [6], [7], [8]). For the Shannon wavelet, this is nothing but well-known classical sampling theory and $\mathrm{P}_J$ is simply a lowpass filtering at the Nyquist frequency. For other wavelets, there are two main options for how to handle this projection:

1) *To apply* $\mathrm{P}_J$ *as in Figure 1* has clear advantages even if $f$ is supposed to be in $V_J$, as in Example 1. The sample values of $f_J$ must be computed from the *available* input, which typically is sample values of a signal $f$ that is *lowpass filtered* at the Nyquist frequency. The hard way to do this would be to first reconstruct $f$ from its sample values. An efficient way to do it is described in [9].

2) *Without* $\mathrm{P}_J$, $\mathbf{f}_J$ is replaced by a vector $\mathbf{f}$ of sample values of the original input $f$, which, together with a sparse matrix approximation $\Phi_{\mathrm{approx}}^{-1}$, introduces *two* different errors in (4):

$$\left\| \Phi_{\mathrm{approx}}^{-1}\mathbf{f} - \Phi^{-1}\mathbf{f}_J \right\| \leq \left\| \Phi_{\mathrm{approx}}^{-1} \right\| \left\| \mathbf{f}_J - \mathbf{f} \right\|$$
$$+ \left\| \Phi_{\mathrm{approx}}^{-1} - \Phi^{-1} \right\| \left\| \mathbf{f}_J \right\|,$$

e.g., with $l_2$-norm for vectors and the corresponding operator norm for matrices (defined in (10)). For deriving error estimates of the first term, we suggest to use known properties of $f$ (e.g., $f$ bandpass filtered) and some estimate of the error $\|f_J - f\|$, such as (for several different norms) the sharp bounds in [10], [11] or (for $L_2$-norm, $\widetilde{\psi}$ with $m$ vanishing moments and with $D$

for differentiation) the well-known bound $\|f - f_J\| \leq C2^{-jn} \|D^n f\|$ for $n = 0, 1, \ldots, m-1$ [3, Section 4.7]. This paper is primarily devoted to approach 1). We will derive simple estimates for the error $\left\| \Phi_{\mathrm{approx}}^{-1}\mathbf{f}_J - \Phi^{-1}\mathbf{f}_J \right\| \leq \left\| \Phi_{\mathrm{approx}}^{-1} - \Phi^{-1} \right\| \|\mathbf{f}_J\|$ and derive band matrices $\Phi_{\mathrm{approx}}^{-1}$ with small operator norm approximation error $\left\| \Phi_{\mathrm{approx}}^{-1} - \Phi^{-1} \right\|$.

## II. OPTIMAL PREFILTERING

As described after (1), the exact values of $a_{J,l}$ for $l$ close to the endpoints of $I_J$ depend on how interval boundary issues are treated. This does not affect the following derivations, except for some modification of matrix elements near the matrix borders in the approach with modified basis functions, so that $\Phi$ no longer is a convolution matrix. However, for "large enough" matrices, which also is the primary case when prefilter approximations are useful, we expect these modifications to have negligible effects on the approximation errors that we analyze in the following sections.

The prefilter computes the most fine-scale coefficients

$$\mathbf{a}_J = \begin{pmatrix} a_{J,0} & a_{J,1} & \cdots & a_{J,|I_J|-1} \end{pmatrix}^T$$

(with $I_J = \{0, 1, \ldots, |I_J| - 1\}$ for simplicity of notation) from a vector $\mathbf{f}_J$ of sample values $f_J(2^{-J}x_0 + kT_s)$, where $k = 0, 1, \ldots, |I_J| - 1$. We will show that prefilter approximations can have a very dramatic dependence on $x_0$, but without loss of generality, we can set $T_s = 2^{-J}$, so that

$$\mathbf{f}_J = \begin{pmatrix} f_J(2^{-J}x_0) & \cdots & f_J(2^{-J}(x_0 + |I_J| - 1)) \end{pmatrix}^T. \quad (5)$$

From (1) we have

$$f_J(2^{-J}(x_0 + k)) = \sum_{l \in I_J} a_{J,l}\varphi_{J,l}(2^{-J}(x_0 + k)), \quad (6)$$

or, with linear algebra notation,

$$\mathbf{f}_J = \Phi_J \mathbf{a}_J \overset{\mathrm{def}}{=} 2^{J/2}\Phi\mathbf{a}_J, \qquad (\Phi)_{k,l} \overset{\mathrm{def}}{=} \varphi(x_0 + k - l). \quad (7)$$

In the following sections, we will consider the asymptotic case with doubly infinite matrices, for which we conclude in
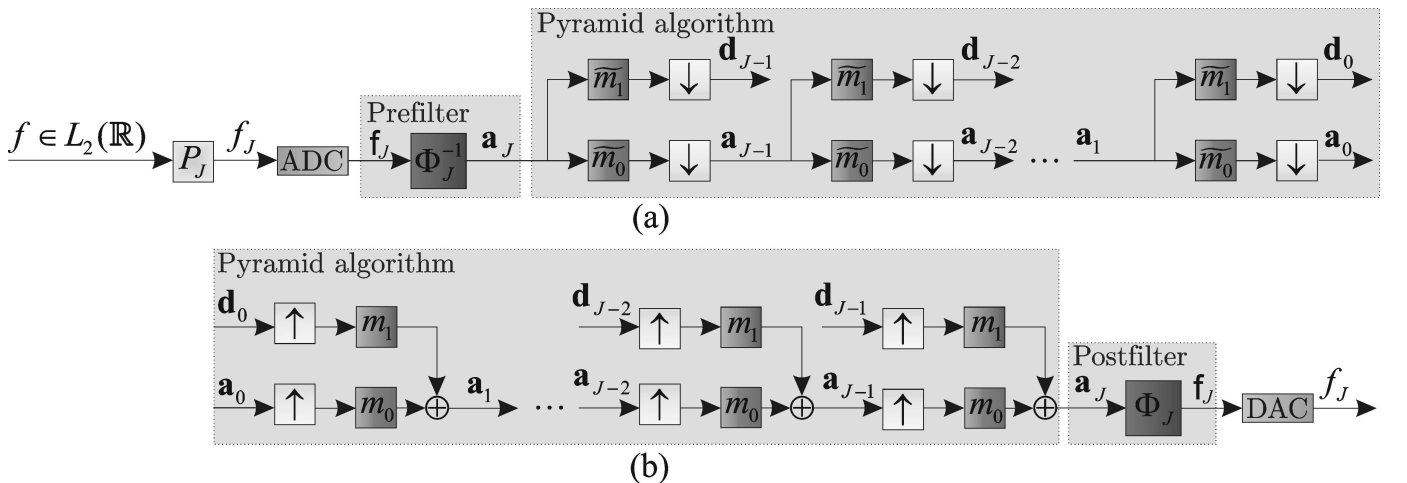


Fig. 1.   Discrete wavelet transform analysis (a) and synthesis (b) of a function $f_J \in V_J$ for biorthogonal (or orthonormal) wavelets. Digital-to-analog (DAC) and analog-to-digital (ADC) converters handle the conversion between $f_J$ and the vector $\mathbf{f}_J$ of sample values. Both the well-known pyramid algorithm and the slightly faster lifting scheme approach require pre- and postfilters for conversion between $\mathbf{f}_J$ and the most fine-scale scaling function coefficients $\mathbf{a}_J$.

Remark 3, page 7, that $\Phi_J$ is *invertible* for several commonly used wavelets. Hence, for compact support $\varphi$, we can truncate $\Phi_J$ to a (not necessarily square) finite matrix with linearly independent columns. Hence, $\|\mathbf{f}_J - \Phi_J \mathbf{a}_J\|_2$ is minimal for the least square error solution $\mathbf{a}_J = (\Phi_J^* \Phi_J)^{-1} \Phi_J^* \mathbf{f}_J$, where $\Phi_J^*$ is the conjugate transpose of $\Phi_J$ (with corresponding truncation of the righthand side approximations in the finite matrix version of (8)) or $\mathbf{a}_J = \Phi_J^{-1} \mathbf{f}_J$ for the infinite $\Phi_J$ of the following sections, as well as for any truncated and square $\Phi_J$ with a finite (for invertibility) condition number that is small enough for the numerical stability constraints of the application at hand. In Figure 2 the numerical stability of high degree Daubechies prefilters show a clear sensitivity to the type of truncation, whereas high degree B-spline prefilters have large condition numbers also without truncation of the remaining columns.

For all commonly used wavelets, $\varphi$ is very fast decaying or even has compact support, so $\Phi_J$ is usually a band matrix or at least diagonally dominant, so that a good banded approximation is likely to exist. Such sparseness gives the computation of $\Phi \mathbf{a}_J$ in (7) the same computational complexity $O(2^J)$ as the pyramid algorithm. The prefilter may be much less sparse but since $\Phi_J$ is a convolution matrix, it can be diagonalized with the fast Fourier transform and the prefilter computation $\mathbf{a}_J = \Phi_J^{-1} \mathbf{f}_J$ can be thus performed in $O(J2^J)$ arithmetic operations. We will now derive prefilter approximations that decrease also the prefiltering complexity to $O(2^J)$ arithmetic operations.

## III. LOW-COMPLEXITY PREFILTER APPROXIMATIONS

We will derive prefilter approximations for biorthogonal wavelets (described in more detail, e.g., in [3]). To each such basis corresponds a *dual* (or biorthogonal) wavelet basis $\left\{ \widetilde{\varphi}_{j,k}, \widetilde{\psi}_{j,k} \in L_2(\mathbb{R}) \,\middle|\, k \in \mathbb{Z}, j \in \mathbb{N} \right\}$, such that the scaling function coefficients are

$$a_{J,k} = \int_{\mathbb{R}} f_J(x) \overline{\widetilde{\varphi}_{J,k}(x)} \, dx.$$

Orthonormal wavelets are included as the special case that occurs if $\widetilde{\varphi} = \varphi$ and $\widetilde{\psi} = \psi$.

The most commonly used prefilter approximation is based on the fact that $\widetilde{\varphi}$ is well localized around some $x_{\max}$ such that $|\widetilde{\varphi}(x_{\max})| = \max_x |\widetilde{\varphi}(x)|$ and on the observation

$$\int_{\mathbb{R}} \widetilde{\varphi}_{J,k}(x) \, dx = 2^{J/2} \int_{\mathbb{R}} \widetilde{\varphi}(2^J x - k) \, dx = c 2^{-J/2},$$

where

$$c \overset{\text{def}}{=} \int_{\mathbb{R}} \widetilde{\varphi}(x) \, dx \quad (= 1 \text{ for orthonormal wavelets}).$$

Consequently, for "large enough" $J$ and continuous $f_J$, it follows from (7) that

$$a_{J,k} = \int_{\mathbb{R}} f_J(x) \overline{\widetilde{\varphi}_{J,k}(x)} \, dx \approx \overline{c} 2^{-J/2} f_J(2^{-J}(x_{\max} + k))$$

$$\Rightarrow \boxed{\Phi_J^{-1} \approx \overline{c} 2^{-J/2} \mathrm{I}}. \quad (8a)$$

Another approximation (suggested, e.g., in [3], [12], [13] for orthonormal wavelets) is based on the rectangle integral approximation

$$a_{J,k} = \int_{\mathbb{R}} f_J(x) \overline{\widetilde{\varphi}_{J,k}(x)} \, dx$$
$$\approx 2^{-J} \sum_{l \in \mathbb{Z}} f_J(2^{-J}(x_0 + l)) \overline{\widetilde{\varphi}_{J,k}(2^{-J}(x_0 + l))}$$
$$= 2^{-J/2} \sum_{l \in \mathbb{Z}} f_J(2^{-J}(x_0 + l)) \overline{\widetilde{\varphi}(x_0 + l - k)}$$

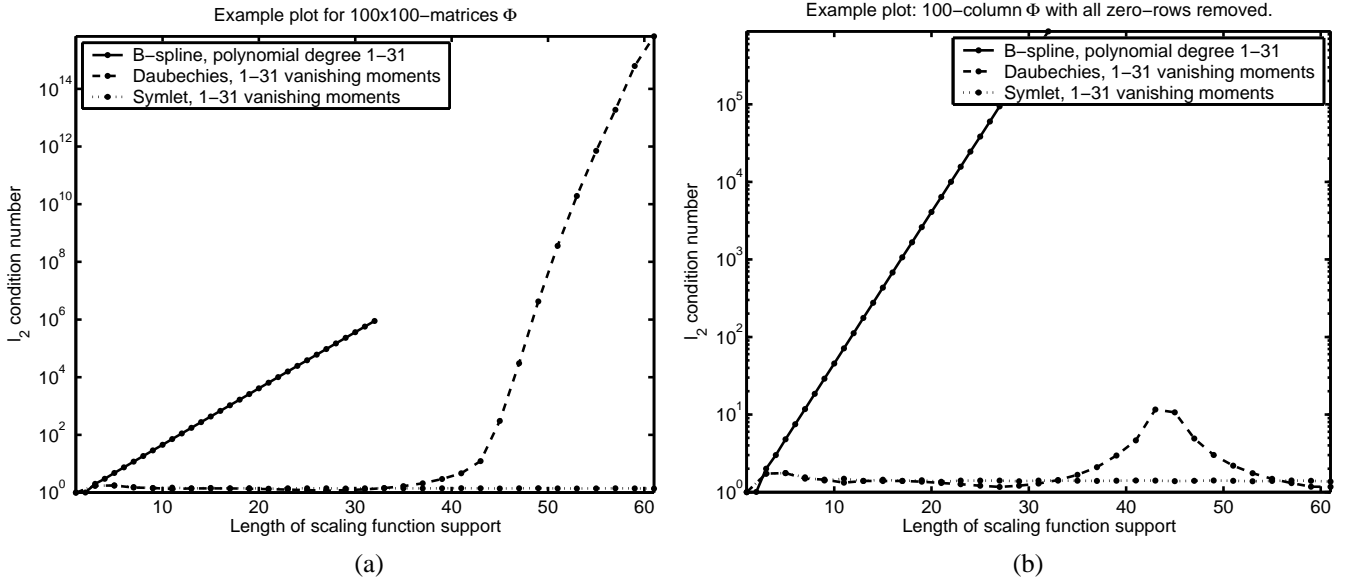$$\Rightarrow \boxed{\Phi_J^{-1} \approx 2^{-J/2} \widetilde{\Phi}^*}, \quad (8b)$$



Fig. 2. Typical example plot of $l_2$-condition numbers when $\Phi_J$ (with $x_0$ given by (16)) is truncated to a $100 \times 100$ matrix (a) and to a minimal 100 column matrix with no nonzero column entries removed (b). While high degree Daubechies prefilters show a clear sensitivity to the truncation of nonzero entries, the increasing B-spline condition numbers seems to rather be caused by increasingly slow-varying $\varphi$ and increasingly more similar columns.

where $\widetilde{\Phi}^*$ is the conjugate transpose of $\widetilde{\Phi}$ and

$$(\widetilde{\Phi})_{k,l} \stackrel{\text{def}}{=} \widetilde{\varphi}(x_0 + k - l).$$

If $\psi$ and $\widetilde{\psi}$ have $n$ vanishing moments, then for polynomial $f$ of degree $\leq n$, (8b) gives perfect reconstruction of the coefficients $a_{J,k}$ and (on finite intervals) $f$ (see, e.g., [3], [1]). In Remark 6, page 10, we describe some wavelet spaces for which (8b) also becomes an equality.

Next, we exploit the fact (see Remark 3, page 7) that for a large number of wavelet scaling functions $\varphi$, $\|\mathrm{I} - \Phi\| < 1$, so that $\Phi^{-1}$ exists and can be approximated with a truncated Neumann series expansion:

$$\boxed{\Phi_J^{-1} = 2^{-J/2}\Phi^{-1} \approx 2^{-J/2}\sum_{k=0}^{N}(\mathrm{I} - \Phi)^k} \qquad (8c)$$

For orthonormal wavelets and $N = 0$, (8c) coincides with the standard approach (8a). For $N = 1$, (8c) gives $\Phi_J^{-1} \approx 2^{-J/2}(2\mathrm{I} - \Phi)$, which is an equally sparse alternative to (8b).

We will also consider two higher degree quadrature formula generalizations of (8b), denoted *Sard* and *Lagrange* approximation: For wavelets with $p$ vanishing moments and arbitrary integers $m$ and $n \geq m + p$, Ehrich [14] derived an algorithm for computing a quadrature formula approximation $(\Phi_{\mathrm{approx}}^{-1})_{k,l} = h(k - l)$ of the prefilter with nonzero $h(k)$ only if $k = -n, \ldots, -m$. The approximation is derived only for $x_0 = 0$, but is "Sard-optimal", which means that no other prefilter with the same nonzero entries give a smaller maximum error $\sup_{f \in M_{2,p}} \left| (\Phi_J^{-1}\mathbf{f}_J)_k - a_{J,k} \right|$ for $f$ in the Sobolev space $M_{2,p} \stackrel{\text{def}}{=} \left\{ f \,\middle|\, f \in C^{p-1} \text{ and } \|f^{(p)}\|_2 \leq 1 \right\}$. It is not completely clear from [14] whether these Sard-optimal prefilters can be generalized to arbitrary $x_0$. We leave this question open for future research, but investigate the quadrature formula for arbitrary $x_0$ that follows from a Lagrange polynomial interpolation $f_J(x) \approx \sum_{l \in I_k} f_J(\lambda_l) L_{I_k,l}(x)$ (proposed in [15]) with sampling points $\lambda_l = 2^{-J}(x_0 + l)$, index set $I_k = \{k - M, k - M + 1, \ldots, k + N\}$ (symmetric around k) and $L_{I_k,l}(x) \stackrel{\text{def}}{=} \prod_{m \in I_k, m \neq l} \frac{x - \lambda_m}{\lambda_l - \lambda_m}$. Thus

$$a_{J,k} = \langle f_J, \widetilde{\varphi}_{J,k} \rangle \approx \sum_{l \in I_k} f_J(\lambda_l) \langle L_{I_k,l}, \widetilde{\varphi}_{J,k} \rangle,$$

whereas by (7), $a_{J,k} = \sum_{l \in \mathbb{Z}} (\Phi_J^{-1})_{k,l} f_J(\lambda_l)$. Hence $(\Phi_J^{-1})_{k,l} = 0$ for $k - l \notin [-N, M]$ and for integers $k - l \in [-N, M]$ the substitutions $y = 2^J x - k$ and $n = k - m$ gives

$$(\Phi_J^{-1})_{k,l} \approx \int_{\mathbb{R}} \left( \prod_{m \in I_k, m \neq l} \frac{x - \lambda_m}{\lambda_l - \lambda_m} \right) \overline{\widetilde{\varphi}_{J,k}(x)} \, \mathrm{d}x$$

$$= 2^{-\frac{J}{2}} \int_{\mathbb{R}} \overline{\widetilde{\varphi}(y)} \prod_{m \in I_k, m \neq l} \frac{y - (x_0 + m - k)}{l - m} \, \mathrm{d}y$$

$$\boxed{(\Phi_J^{-1})_{k,l} \approx 2^{-\frac{J}{2}} \int_{\mathbb{R}} \overline{\widetilde{\varphi}(y)} \prod_{n \in [-N,M], n \neq k-l} \frac{y - (x_0 - n)}{n - (k - l)} \, \mathrm{d}y}$$
$$(8d)$$

We will use notation like *Sard N* and *Lagrange N* for prefilter approximations that have the same number of nonzero diagonals (or filter coefficients) as the Neumann $N$ prefilter.

We will also compare with the *truncated inverse N* approximation obtained by computing $\Phi^{-1}$, keeping the largest (constant) diagonals and setting the others to zero. This corresponds to convolution with a prefilter that is a minimum $l_2$-norm error $n$-term approximation of the optimal prefilter. In next section, this is equivalent to a minimum Fourier series $L_2$-error instead of the $L_\infty$-error that we will prefer to minimize. So although it's not truly optimal, it is still likely to be a good and easily computed prefilter approximation. In our implementation, we compute the filter coefficients of $\Phi_{\mathrm{Trunc}} \approx \Phi^{-1}$ with the formula $h_{\mathrm{Trunc}} = \mathrm{IFFT}(1/\mathrm{FFT}(h_{\mathrm{zp}}))$, where $h_{\mathrm{zp}}$ is a "sufficiently zeropadded" copy of the filter $h_n = \varphi(x_0 + n) = (\Phi)_{k,k-n}$ ($k$ arbitrary integer). We describe two other ways to compute $\Phi^{-1}$ in remarks 4 and 6, Section V.

*Remark 2:* For biorthogonal wavelets, the quadrature approaches (conjugate transpose, Sard and Lagrange) depend on how $\widetilde{\varphi}$ is chosen. In the *semiorthogonal* special case, $\widetilde{\varphi} = \sum_{l \in \mathbb{Z}} c_l \varphi_l \in V_J$ is the unique $l_2$-solution $(c_l)_{l \in \mathbb{Z}}$ of

$$\delta_{0,k} = \langle \widetilde{\varphi}, \varphi_k \rangle = \left\langle \sum_{l \in \mathbb{Z}} c_l \varphi_l, \varphi_k \right\rangle = \sum_{l \in \mathbb{Z}} c_l \langle \varphi_l, \varphi_k \rangle,$$

or in vector notation,

$$\mathrm{G}\mathbf{c} = \boldsymbol{\delta}, \qquad (9)$$

where $\boldsymbol{\delta}_k = \delta_{0,k}$ and the doubly infinite matrix G is known as the *Gram matrix*. For B-spline wavelets of degree 2–30, a numerical solution of (9) gives an exponentially decaying $\widetilde{\varphi}$ (typically increasing the essential support due to floating point precision with a factor $> 30$), which would have to be truncated to give a banded conjugate transpose approximation, but then (and for the just described optimality reasons) we find it more reasonable to instead truncate $\Phi^{-1}$. The long dual support also made our implementation of the Lagrange interpolation too slow for the large scale comparison in Section V. Thus we will investigate the conjugate transpose and Lagrange approximation errors only for Daubechies and Symlet scaling functions. (A standard orthonormalization procedure would only give exponential decaying prefilter *and* postfilter [2], but one possible future alternative might be to look for a compact support dual, the existence of which is proved for so-called minimally supported scaling functions in [16].)

## IV. COMPUTING THE APPROXIMATION ERROR

For simple signal-length independent error analysis and comparison of the approximations in Section III we consider the limiting case with doubly infinite matrices and with vectors, such as $\mathbf{a}$, replaced by $l_2$-sequences $a = (a_k)$ with the usual norm

$$\|a\|_{l_2} \stackrel{\text{def}}{=} \left( \sum_{k \in \mathbb{Z}} |a_k|^2 \right)^{\frac{1}{2}}. \qquad (10a)$$

Finite length inputs are reperesented by doubly zeropadded sequences. A reasonable performance measure for approximations $\Phi_{\mathrm{approx}}^{-1}$ is the relative $l_2$-error of the resulting coefficient approximations $\Phi_{\mathrm{approx}}^{-1}f$ given by the *operator norm*

$$\|\Phi^{-1} - \Phi_{\mathrm{approx}}^{-1}\| \stackrel{\text{def}}{=} \sup_{0 \neq f \in l_2} \frac{\|\Phi^{-1}f - \Phi_{\mathrm{approx}}^{-1}f\|_{l_2}}{\|f\|_{l_2}}. \qquad (10b)$$

In Proposition 1, we derive simple formulas for this error by using an isomorphism (derived in Lemma 1)

$$U\colon \mathcal{C} \to \mathcal{B}(L_2([0,1])), \qquad U\,H \stackrel{\text{def}}{=} M_{\widehat{h}^*},$$

in the following commutative diagram,

$$
\begin{array}{ccc}
s \in l_2 & \xrightarrow{\;H \in \mathcal{C}\;} & h * s \in l_2 \\[2pt]
{\scriptstyle F}\Big\downarrow & & \Big\uparrow{\scriptstyle F^{-1}} \\[2pt]
\widehat{s}^* \in L_2([0,1]) & \xrightarrow{\;M_{\widehat{h}^*}\;} & \widehat{h}^*\widehat{s}^* \in L_2([0,1])
\end{array}
$$

where F is the *Fourier series operator*

$$F\,(h_k)_k \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}} h_k \mathrm{e}^{-\mathrm{i}2\pi k(\cdot)} \stackrel{\text{def}}{=} \widehat{h}^*,$$

$M_u$ is the multiplicative operator $M_u\,g \stackrel{\text{def}}{=} ug$, $\mathcal{B}(L_2([0,1]))$ is the set of linear bounded operators from $L_2([0,1])$ to $L_2([0,1])$ and $\mathcal{C}$ is the set of doubly infinite *convolution matrices* H such that for $H_{k,l} = h_{k-l}$,

$$h = (h_k)_{k \in \mathbb{Z}} \in l_2 \qquad \text{and} \qquad \operatorname*{ess\,sup}_{\xi}\left|\widehat{h}^*(\xi)\right| < \infty. \quad (11)$$

(See, e.g., [17] for definitions of essential infimum/supremum and other functional analysis terminology used in our proofs.)

*Lemma 1:* For $H \in \mathcal{C}$ and $h \in l_2$ as above, it follows that

$$\|H\| = \|U\,H\| = \operatorname*{ess\,sup}_{\xi \in [0,1)}\left|\widehat{h}^*(\xi)\right|, \qquad (12a)$$

$$
\begin{aligned}
H \text{ invertible} \quad &\Leftrightarrow \quad M_{\widehat{h}^*} \text{ invertible} \\
&\Leftrightarrow \quad \operatorname*{ess\,inf}_{\xi \in [0,1)}\left|\widehat{h}^*(\xi)\right| > 0,
\end{aligned}
\qquad (12b)
$$

and that

$$H \text{ invertible} \quad \Rightarrow \quad U(H^{-1}) = M_{\frac{1}{\widehat{h}^*}}. \qquad (12c)$$

Moreover, for $H_1, H_2 \in \mathcal{C}$ with associated $l_2$-sequences $h_1$ and $h_2$, respectively, it follows that $H_1 H_2 \in \mathcal{C}$, $H_1 + H_2 \in \mathcal{C}$,

$$
\begin{aligned}
U(H_1 H_2) &= (U\,H_1)(U\,H_2) \quad \text{and} \\
U(H_1 + H_2) &= U\,H_1 + U\,H_2.
\end{aligned}
\qquad (12d)
$$

*Proof:* For any $g \in L_2([0,1])$ and $\varepsilon > 0$, choose $N$ such that, for $a = (a_k) = F^{-1} g$ and $g_N \stackrel{\text{def}}{=} \sum_{|n| \le N} a_n \mathrm{e}^{-\mathrm{i}2\pi nx}$,

$$\|g - g_N\| < \varepsilon.$$

Then (11) gives that $\sum_{|n| \le N} h_{k-n} a_n \in l_2$ and

$$
\begin{aligned}
F\,H\,F^{-1} g_N &= \sum_{k \in \mathbb{Z}} (H\,F^{-1} g_N)_k \mathrm{e}^{-\mathrm{i}2\pi k(\cdot)} \\
&= \sum_{k \in \mathbb{Z}} \sum_{|n| \le N} h_{k-n} a_n \mathrm{e}^{-\mathrm{i}2\pi k(\cdot)} \\
&= \sum_{|n| \le N} a_n \sum_{l \in \mathbb{Z}} h_l \mathrm{e}^{-\mathrm{i}2\pi (n+l)(\cdot)} \\
&= \sum_{l \in \mathbb{Z}} h_l \mathrm{e}^{-\mathrm{i}2\pi l(\cdot)} \sum_{|n| \le N} a_n \mathrm{e}^{-\mathrm{i}2\pi n(\cdot)} \\
&= \widehat{h}^* g_N = M_{\widehat{h}^*}\, g_N.
\end{aligned}
$$

Hence, $F\,H\,F^{-1} = M_{\widehat{h}^*}$ on a dense subset of $L_2([0,1])$ and thus also on all of $L_2([0,1])$. Thus (12a) follows from the Parseval equation $\|F\,c\| = \|c\|$:

$$
\|H\| = \left\|F\,H\,F^{-1}\right\| = \left\|M_{\widehat{h}^*}\right\| = \sup_{\|g\|_{L_2([0,1])}=1}\|\widehat{h}^* g\|
$$
$$
= \operatorname{ess\,sup}|\widehat{h}^*|.
$$

Since F is bijective, it is also apparent that $U\,H \stackrel{\text{def}}{=} M_{\widehat{h}^*} = F\,H\,F^{-1}$ is invertible if and only if H is invertible and that when this is the case, $(U\,H)^{-1}$ is bounded (according to the open mapping theorem) and

$$U(H^{-1}) = F\,H^{-1}\,F^{-1} = (U\,H)^{-1} = M_{\widehat{h}^*}^{-1} = M_{1/\widehat{h}^*}.$$

This proves (12b) and (12c). Finally, the well-known fact

$$\operatorname*{ess\,sup}_{\xi}\left|\widehat{h_1 * h_2}^*(\xi)\right| = \operatorname*{ess\,sup}_{\xi}\left|\widehat{h_1}^*(\xi) \cdot \widehat{h_2}^*(\xi)\right| < \infty,$$

implies that $H_1 H_2 \in \mathcal{C}$, so that (12d) follows directly from the fact that $U\,H = F\,H\,F^{-1}$. This completes the proof. ∎

Now we can compute all errors $\|\Phi^{-1} - \Phi_{\text{approx}}^{-1}\|$ in Section III by identifying $h$ and then applying (12a). In particular, from the above preservation of addition, multiplication and inverses, we immediately get the following simple error bounds:

*Proposition 1 (error bounds for (8a)–(8c)):* Let $m(\xi) \stackrel{\text{def}}{=} \widehat{h}^*(\xi)$ in the special case $h_k = \varphi(x_0 + k)$ corresponding to $H = \Phi$. Suppose that

$$\sum_{k \in \mathbb{Z}} |\varphi(x_0 + k)| < \infty \qquad \text{and} \qquad |m(\xi)| > 0 \qquad (13)$$

for all $\xi \in [0,1]$. Then for all integers $N \ge 0$,

$$
\begin{aligned}
\left\|\Phi_J^{-1} - 2^{-J/2} \sum_{k=0}^{N}(I - \Phi)^k\right\| &= \\
&= 2^{-J/2} \max_{\xi}\left|\frac{1}{m(\xi)} - \sum_{k=0}^{N}(1 - m(\xi))^k\right|. \quad (14a)
\end{aligned}
$$

Similarly, if $\widetilde{m}(\xi) \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}} \widetilde{\varphi}(x_0 + k)\mathrm{e}^{-\mathrm{i}2\pi k\xi}$ and $\sum_{k \in \mathbb{Z}} |\widetilde{\varphi}(x_0 + k)| < \infty$, then
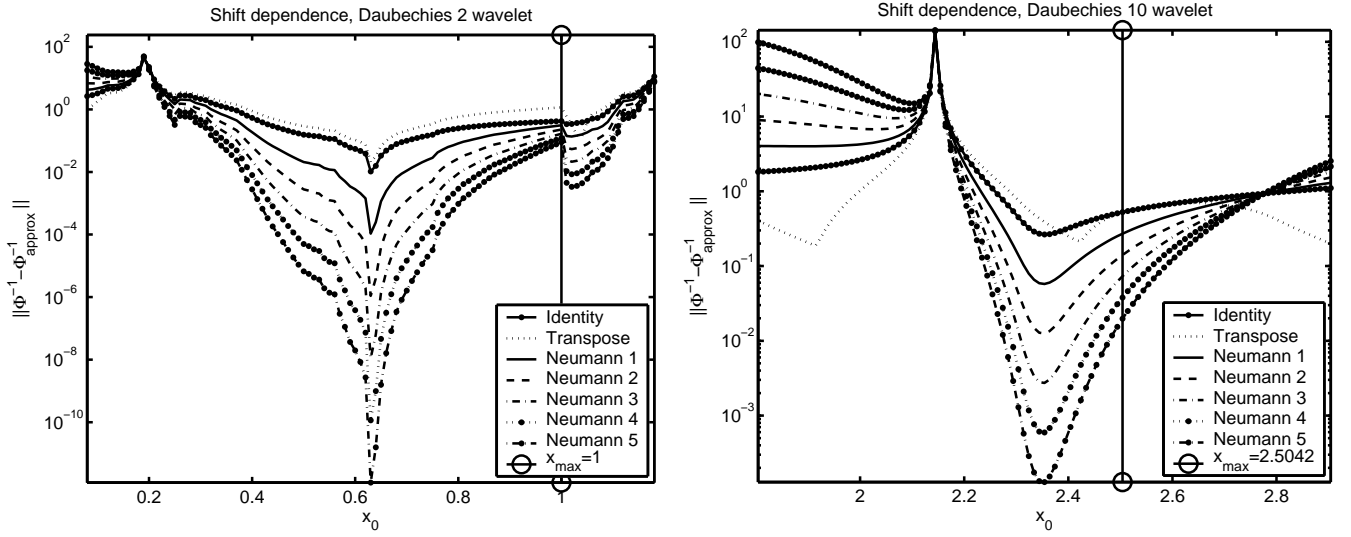
$$\left\|\Phi_J^{-1} - 2^{-J/2}\widetilde{\Phi}^*\right\| = 2^{-J/2} \max_{\xi}\left|\frac{1}{m(\xi)} - \overline{\widetilde{m}(\xi)}\right|. \quad (14b)$$

*Proof:* Suppose first that $J = 0$. Our assumption $\sum_k |\varphi(x_0 + k)| < \infty$ guarantees uniform convergence and therefore also continuity of the 1-periodic function $m$. Hence the assumption $|m(\xi)| > 0$ implies that there are upper and lower bounds
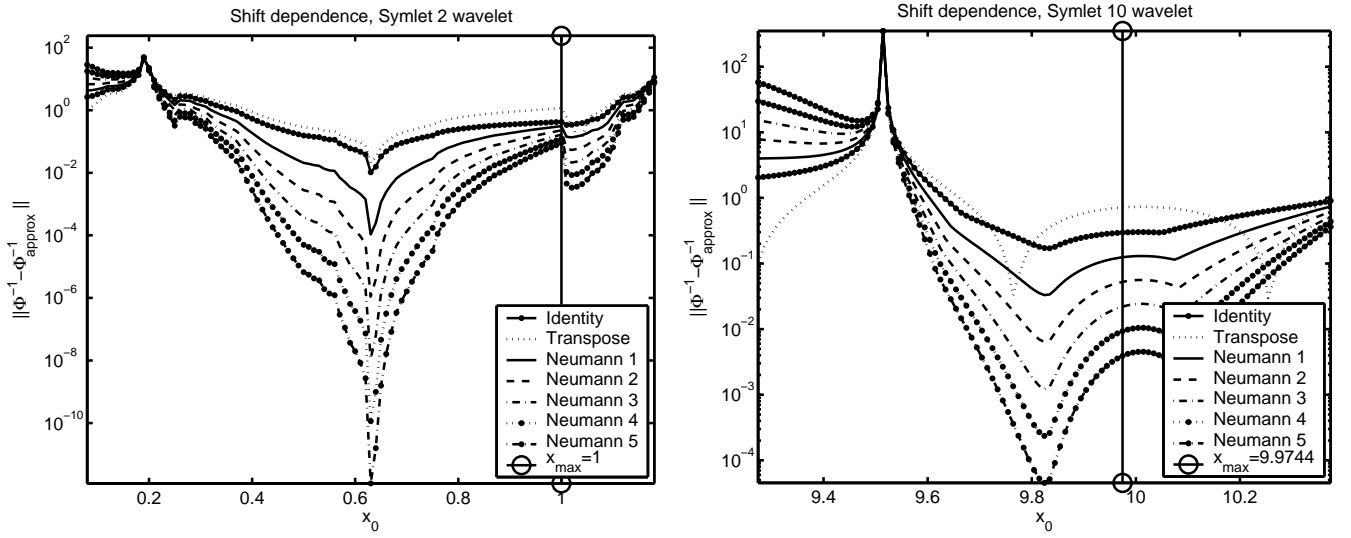
$$0 < A \le |m(\xi)| \le B < \infty. \qquad (15)$$

Consequently, we know from (12b) that $M_m$ (and thus also $\Phi$) is invertible. Thus (12c) and (12d) give
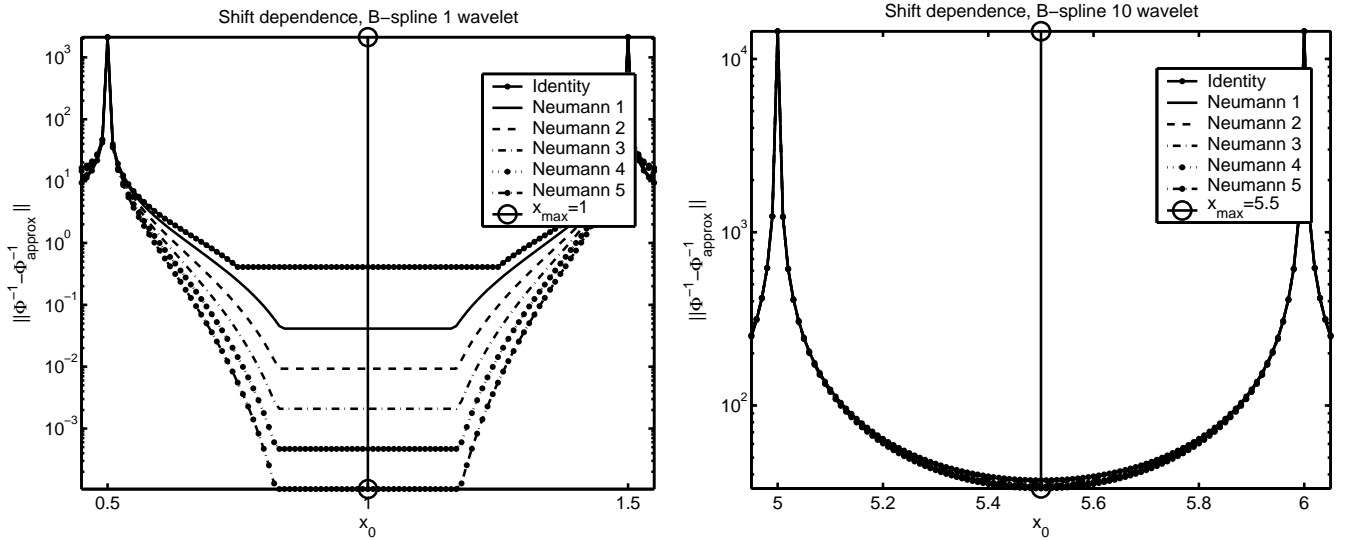
$$U\left(\Phi^{-1} - \sum_{k=0}^{N}(I - \Phi)^k\right) = M_{\frac{1}{m} - \sum_{k=0}^{N}(1-m)^k}$$

(a) Daubechies wavelets with 2 and 10 vanishing moments.



(b) Symlet wavelets with 2 and 10 vanishing moments.



(c) B-spline wavelets of polynomial degree 1 and 10.

Fig. 3.   Example plots showing the $x_0$-dependence of the asymptotic errors (14) for $J = 0$. Multiplication with $2^{J/2}$ gives the actual error.

and (14a) follows. For (14b), the matrix $H = \widetilde{\Phi}^*$ corresponds to a convolution with $h = (h_k)_k = \left(\overline{\widetilde{\varphi}(x_0 - k)}\right)_k$. Hence,

$$\widehat{h}^*(\xi) = \sum_{k \in \mathbb{Z}} \overline{\widetilde{\varphi}(x_0 - k)} \mathrm{e}^{-\mathrm{i}2\pi k \xi}$$
$$= \overline{\sum_{n \in \mathbb{Z}} \widetilde{\varphi}(x_0 + n) \mathrm{e}^{-\mathrm{i}2\pi n\xi}} = \overline{\widetilde{m}(\xi)},$$

so that $U\widetilde{\Phi}^* = M_{\overline{\widetilde{m}}}$ and (14b) follows. We have thus proved (14) for $J = 0$. The extension to arbitrary $J \in \mathbb{Z}$ now follows from a multiplication with $2^{-J/2}$, since

$$\Phi_J^{-1} = (2^{J/2}\Phi)^{-1} = 2^{-J/2}\Phi^{-1}.$$

∎

A striking advantage of these bounds is that for compact support scaling functions (like those studied in this paper), $m$ is an easily computed trigonometric polynomial.

*Remark 3:* We used the condition (13) to guarantee that $A < |m(\xi)| < B$ for some $A, B > 0$. For convergence of the Neumann series approximation (when $N \to \infty$), it is sufficient to use the stronger condition $|1 - m(\xi)| < 1$, or equivalently $\|I - \Phi\| < 1$, which is known to hold, for example, for Shannon wavelets and for all B-spline wavelets [6]. It is easy to check (for example, in MATLAB) that $|1 - m(\xi)| < 1$ also for Daubechies wavelets with 1–31 vanishing moments and at least for all Symlet wavelets with up to 36 vanishing moments.

## V. COMPUTATIONAL RESULTS, CONCLUSIONS & REMARKS

Recall from (5) and Remark 1 that throughout the paper, we assume the input to be *integer translated* sample values $f_J(2^{-J}(x_0 + k))$ of some $f_J \in V_J$. We also assume $x_0$ to be known or possible to choose, so that the prefilter $\Phi_J^{-1}$ and its approximations are uniquely defined by (7) and (8). This is a reasonable assumption in several applications (like DMT in Example 1), but others (e.g., due to synchronization problems, OFDM in Example 1) might require an additional analysis of the error caused by choosing an incorrect $x_0$ in (7) and (8). Both this and analysis of errors caused by irregular sampling is out of the scope of this paper but a planned topic for future papers. For now, we refer to, e.g., [6], [18] for some related error bounds and irregular sampling theorems.

Hence the approximation errors $\left\|\Phi^{-1} - \Phi_{\text{approx}}^{-1}\right\|$ depend on $\varphi$ and $x_0$. For Daubechies and Symlet wavelets with 2–31 and 2–36 vanishing moments, as well as for B-spline wavelets with polynomial degree 1–30, plots like the example plots in Figure 3 show a dramatic dependence on $x_0$. Hence, if the application allows for it, $x_0$ should be chosen with care.

In these and the following plots $J = 0$, so the *actual* error is obtained by division with a factor $2^{J/2}$, which is asymptotically proportional to the square root of the signal length $|I_J| = 2^J(|I| + 2^{-J}M)$ (with notation as in (1)). This shows how to choose $J$ for a given target error bound. It is also a strong argument against decreasing the approximation error (with a factor $1/n$) with oversampling (with a factor $n^2$).

From each $x_0$-dependence plot, we have also found the $x_0$ (with precision $\pm 0.005$) that minimizes the approximation error and plotted the results in Figure 4. Note, for example,

that for Daubechies wavelets with 2–5 vanishing moments the Neumann 1 error is about 10–100 times smaller than for the computationally comparable approximations $\Phi^{-1} \approx I$ and $\Phi^{-1} \approx \widetilde{\Phi}^* = \Phi^T$. Hence, *to obtain the same improvement by oversampling, 100–10 000 times more samples are needed.*

The right-hand plots in Figure 4 shows the nearly optimal values of $x_0$ that we used in the left-hand plots. Note that the minimum is obtained for some $x_0 \approx x_{\max}$, with $x_{\max}$ such that

$$|\varphi(x_{\max})| = \max_x |\varphi(x)|. \tag{16}$$

This was predicted in [6], but Figure 3 shows that even if $x_0 \approx x_{\max}$ is a good rule of thumb for quick decisions, an optimal $x_0$ can give much lower approximation errors.

In figure 4 it is also notable that for scaling functions with short support, the rectangle integral approximation ($\Phi^{-1} \approx \widetilde{\Phi}^* = \widetilde{\Phi}^T$) is *even worse* than doing nothing ($\Phi^{-1} \approx I$).

Due to the complexity of the algorithm for computing Sard-optimal filter coefficients, our current MATLAB implementation can compute prefilter errors reasonably fast and with high precision only for up to 10 vanishing moments or polynomial degree 10. It reproduced the filter coefficients in [14, Section 4.1] correctly but the resulting approximation errors are larger than those plotted in Figure 5. Thus Sard-optimal prefilters put some more demand on an efficient implementation and for good performance, we also think that they should be generalized (if possible) to arbitrary $x_0$.

Figure 5, finally, is a comparison of the Neumann $N$ approximations with the truncated inverse $N$ and Lagrange $N$ approximations (with equal $N$ denoting identical number of nonzero diagonals). Note that the Lagrange $N$ error *increases* with increasing $N$. The most likely reason for both this and the Lagrange 1 behavior in Figure 5 (a) is that due to the fixed sampling density, the higher degree Lagrange interpolants are more likely to give high amplitude ripples and a bad local approximation, as illustrated in Figure 5 (d), which shows Lagrange 1 interpolation of the function $f_0$ given by $a_{0,k} = (-1)^k$ in (1) and with the Daubechies 22 optimal shift used in Figure 5 (a).

As a final conclusion, we note that for short support scaling functions, the Neumann filters (which also are the easiest ones to compute) often have the smallest (or nearly smallest) approximation errors and that for longer supports, the truncated inverse prefilter is clearly superior to the others.

*Remark 4:* Recall from Section III that $\Phi^{-1}$ can be relatively fast computed and applied using the fast Fourier transform. A much faster *computation* of $\Phi^{-1}$ is possible if we only consider $l_2$ with real scalars or if $\Phi$ is Hermitian, that is, if $\Phi = \Phi^*$, like for any B-spline $\varphi$ with $x_0 = x_{\max}$. In both these cases, it follows from the invertibility condition $\|I - \Phi\| < 1$ in Remark 3 that $\Phi$ is positive definite. In fact, under these assumptions, $\langle x, (I - \Phi)x \rangle \in \mathbb{R}$, so that (by the Schwarz inequality and the definition of the operator norm)

$$\langle x, \Phi x \rangle = \langle x, Ix - (I - \Phi)x \rangle$$
$$\geq \|x\|^2 - |\langle x, (I - \Phi)x \rangle| \geq \|x\|^2 - \|x\| \cdot \|(I - \Phi)x\|$$
$$\geq \|x\|^2 \left(1 - \|(I - \Phi)\|\right) > 0, \qquad \forall x \in l_2.$$

(a) Daubechies wavelets with 2–31 vanishing moments.



(b) Symlet wavelets with 2–36 vanishing moments.



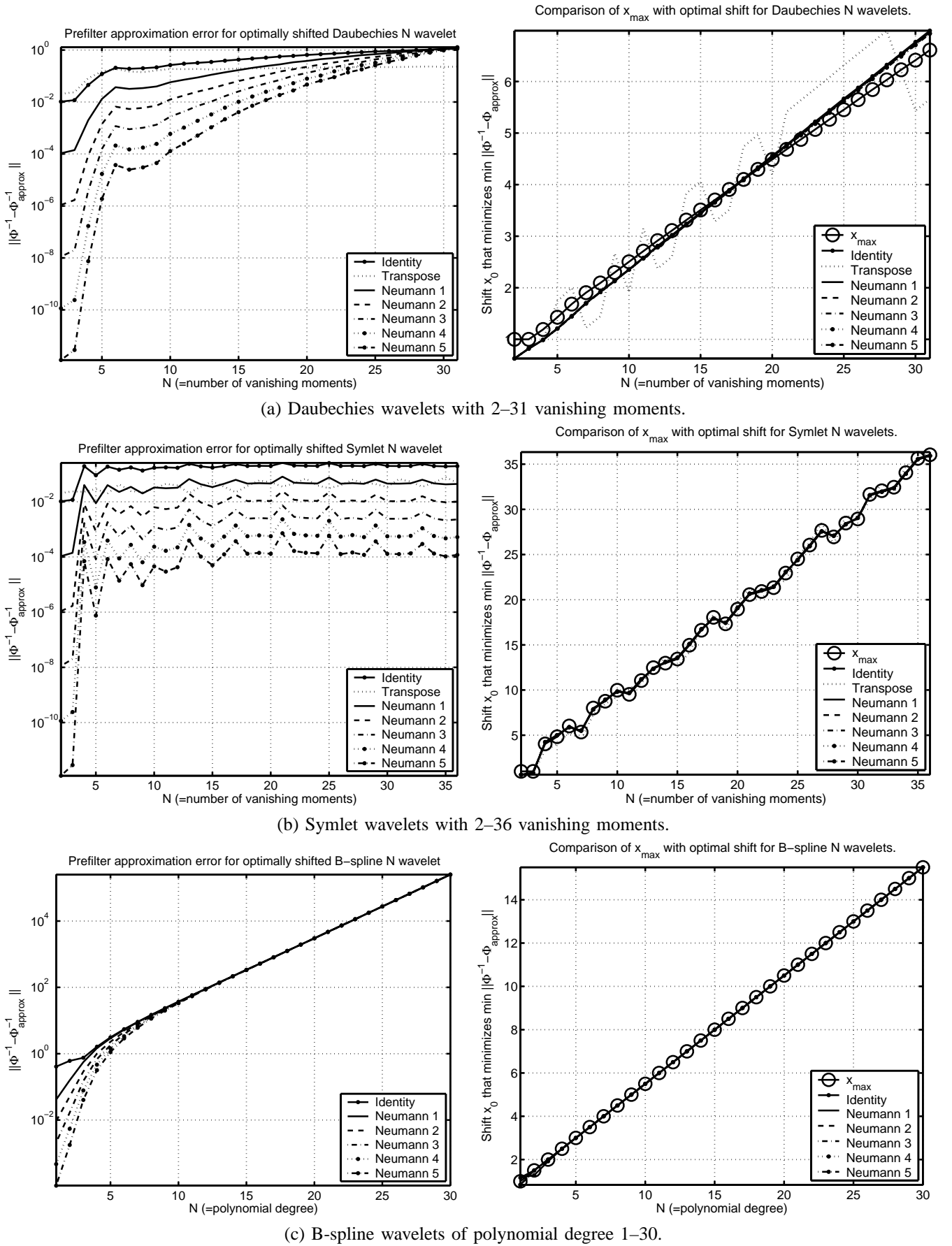(c) B-spline wavelets of polynomial degree 1–30.

Fig. 4.   The left-hand column plots show the minimum operator norm errors of the prefilter approximations (8) when $J = 0$ and with $x_0$ chosen (with maximum error $\pm 0.005$) for minimum error. Multiplication with $2^{J/2}$ gives the actual error. The minimizing $x_0$ used in the lefthand plots are plotted in the right-hand column plots. In accordance with an observation in [6], the minimum errors occur for some $x_0 \approx x_{\max}$ with $x_{\max}$ defined in (16).

(a) Daubechies wavelets with 2–31 vanishing moments.



(b) Symlet wavelets with 2–36 vanishing moments.



(c) B-spline wavelets of polynomial degree 1–30.
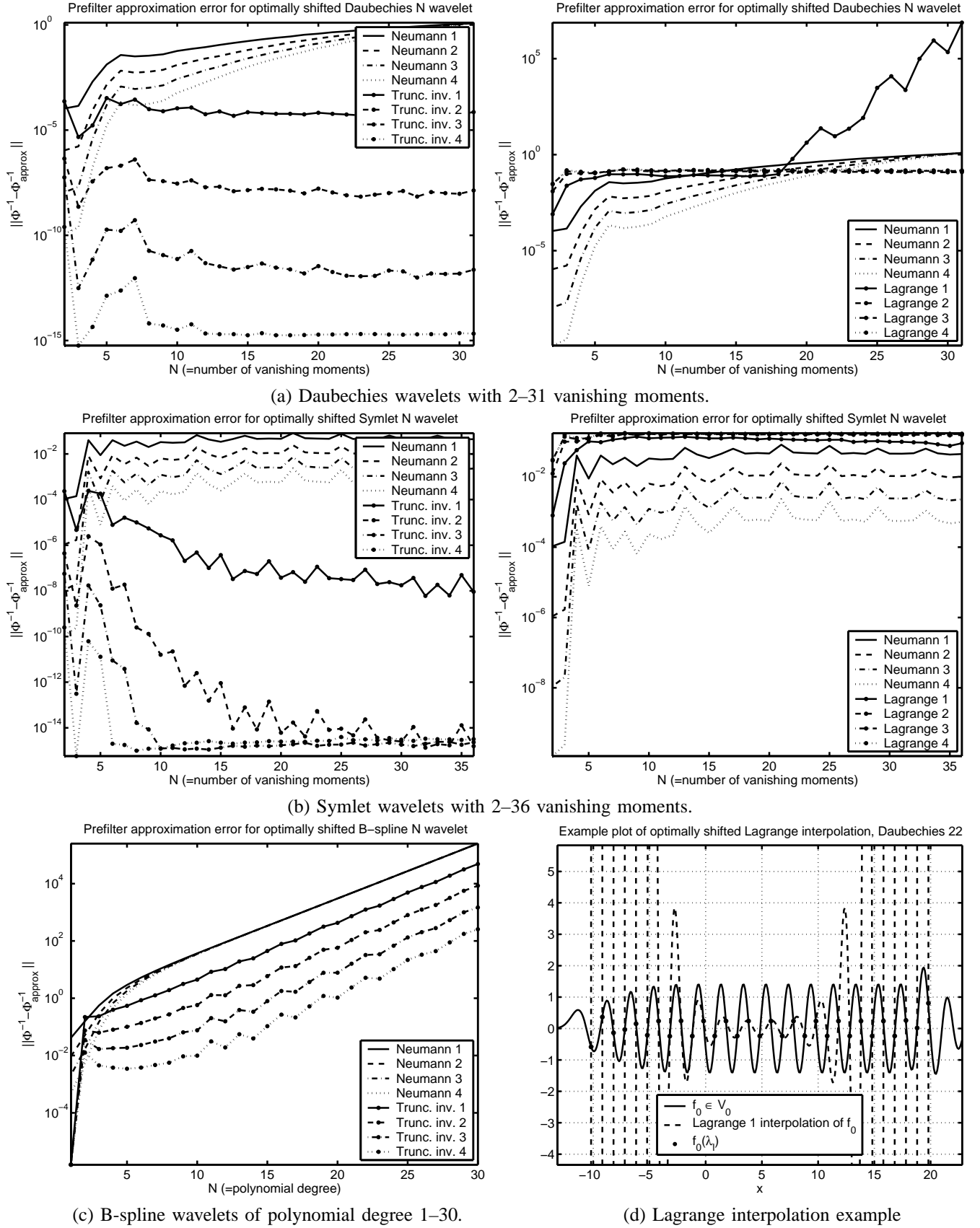
(d) Lagrange interpolation example

Fig. 5. Plots (a)–(c) show a comparison of the Neumann, truncated inverse and Lagrange prefilter approximation errors. Neumann prefilters can generally be computed with the smallest number of arithmetic operations and also performs best for several short support wavelets, whereas the truncated inverse prefilter is superior for longer support scaling functions. Plot (d) shows an example where Lagrange interpolation goes wrong.

An important property of real, symmetric and positive definite matrices is that inverses can be computed with the very fast *conjugate gradient method* (see, e.g., [19]). Note also that it is unlikely that $\Phi = \Phi^*$ for any real-valued and compactly supported *orthonormal* wavelet basis except for the Haar basis, which is the only symmetric such basis [3, page 47].

*Remark 5:* Both the prefilter matrix $\Phi^{-1}$ and the invertibility condition (13) appear in a few different variations and disguises (equations (15) and (17)) in different books and papers. We give a brief review and refer to, for example, [6], [7], [8] for details, convergence issues and generalizations to irregular sampling. With notation as in Proposition 1, it follows from the Poisson summation formula that

$$m(\xi) = \sum_{k \in \mathbb{Z}} \varphi(x_0 + k) e^{-i2\pi k\xi} = \sum_{k \in \mathbb{Z}} \widehat{\varphi(x_0 + \cdot)}(\xi - k).$$

Thus, by the proof of Proposition 1, (13) can be replaced with the weaker assumption (15), which sometimes appears in the equivalent form

$$0 < A \le \left| \sum_{k \in \mathbb{Z}} \widehat{\varphi(x_0 + \cdot)}(\xi - k) \right| \le B < \infty. \qquad (17a)$$

*The prefilter*, finally, is sometimes "hidden" in a so-called *interpolating Riesz basis* $(q_{J,x_0+k})_k$ for $V_J$, which is constructed so that the reconstruction formula $f = \sum_{k \in \mathbb{Z}} a_{J,k} \varphi_{J,k}$ can be replaced with a reconstruction from sample values

$$f = \sum_{k \in \mathbb{Z}} 2^{-J} f(2^{-J}(x_0 + k)) \widetilde{q}_{J,x_0+k}, \qquad \forall f \in V_J, \quad (17b)$$

where $(\widetilde{q}_{J,k})_k$ is some biorthogonal Riesz basis for $V_J$. Roughly speaking, this is done by building the prefilter (or equivalently, the factor $1/m$) into $\widetilde{q}_{J,x_0+k} \stackrel{\text{def}}{=} 2^{J/2} \widetilde{q}(2^J \cdot -(x_0 + k)) \stackrel{\text{def}}{=} 2^{J/2} \widetilde{q}_{x_0}(2^J \cdot -k)$ in the following way: Set

$$\widehat{\widetilde{q}_{x_0}}(\xi) \stackrel{\text{def}}{=} \frac{\widehat{\varphi}(\xi)}{m(\xi)},$$

that is,

$$\widehat{\varphi}(\xi) = m(\xi) \widehat{\widetilde{q}_{x_0}}(\xi) = \sum_{l \in \mathbb{Z}} \varphi(x_0 + l) e^{-i2\pi l\xi} \widehat{\widetilde{q}_{x_0}}(\xi).$$

Then (at least formally) $\varphi = \sum_{l \in \mathbb{Z}} \varphi(x_0 + l) \widetilde{q}_{x_0}(\cdot - l)$ and

$$\varphi_{J,k} = \sum_{l \in \mathbb{Z}} \varphi(x_0 + l) \widetilde{q}_{J,x_0+k+l}. \qquad (18)$$

Under certain mild decay and continuity conditions on $\varphi$ (see [6], [8] for details), there is a unique family of functions $q_x \stackrel{\text{def}}{=} q(\cdot - x) \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}} \overline{\varphi(x - k)} \widetilde{\varphi}(\cdot - k)$ with the special property that $\langle f, q_x \rangle = f(x)$ for all $f \in V_0$. (It also follows from the second last line of (19) that $(q_{J,k})_k$ and $(\widetilde{q}_{J,k})_k$ are dual.) Thus we get the reconstruction formula (17b) from (18)

and the following change of basis:

$$\begin{aligned}
f &= \sum_{k \in \mathbb{Z}} \langle f, \widetilde{\varphi}_{J,k} \rangle \varphi_{J,k} = \sum_{k \in \mathbb{Z}} \langle f, \widetilde{\varphi}_{J,k} \rangle \sum_{l \in \mathbb{Z}} \varphi(x_0 + l) \widetilde{q}_{J,x_0+k+l} \\
&= \sum_{k \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \widetilde{\varphi}_{J,k} \rangle \varphi(x_0 + n - k) \widetilde{q}_{J,x_0+n} \\
&= \sum_{n \in \mathbb{Z}} \left\langle 2^{-J/2} f(2^{-J} \cdot), \sum_{k \in \mathbb{Z}} \overline{\varphi(x_0 + n - k)} \widetilde{\varphi}(\cdot - k) \right\rangle \widetilde{q}_{J,x_0+n} \\
&= \sum_{n \in \mathbb{Z}} \left\langle 2^{-J/2} f(2^{-J} \cdot), q_{x_0+n} \right\rangle \widetilde{q}_{J,x_0+n} \\
&= \sum_{n \in \mathbb{Z}} 2^{-J/2} f(2^{-J}(x_0 + n)) \widetilde{q}_{J,x_0+n}(x), \qquad \forall f \in V_J.
\end{aligned}$$

$$(19)$$

We mention also that [20] contain estimates of the error caused by truncation of the series expansion (17b) to a finite sum.

*Remark 6:* Most commonly used wavelets satisfy the mild decay and continuity conditions of Remark 5. Thus the Lagrange prefilter approximation (8d) can usually be seen as the result of a polynomial approximation $f_J(x) \approx \sum_{l \in I_k} f_J(\lambda_l) L_{I_k,l}(x)$ of (19), thus (and once again with $\widetilde{q}_{J,x} \stackrel{\text{def}}{=} 2^{J/2} \widetilde{q}(2^J \cdot -x))$ making (8d) an approximation of the fact that

$$\begin{aligned}
a_{J,k} &= \langle f_J, \widetilde{\varphi}_{J,k} \rangle = \left\langle \sum_{l \in \mathbb{Z}} 2^{-\frac{J}{2}} f_J(\lambda_l) \widetilde{q}_{J,x_0+l}, \widetilde{\varphi}_{J,k} \right\rangle \\
&= 2^{-\frac{J}{2}} \sum_{l \in \mathbb{Z}} f_J(\lambda_l) \langle \widetilde{q}_{x_0+l}, \widetilde{\varphi}_k \rangle
\end{aligned}$$

gives $(\Phi_J^{-1})_{k,l} = 2^{-J/2} \langle \widetilde{q}_{x_0+l}, \widetilde{\varphi}_k \rangle$. Moreover, if $V_0$ possesses an *orthonormal* interpolation function (such spaces are derived in [21]), then $q = \widetilde{q}$ and the above computation reduces to $a_{J,k} = \sum_{l \in \mathbb{Z}} f_J(\lambda_l) \overline{\widetilde{\varphi}(x_0 + l - k)}$. Thus, for such spaces, $\Phi_J^{-1} = 2^{-J/2} \Phi^*$ and approximation (8b) becomes an equality.

REFERENCES

[1] S. Mallat, *A wavelet tour of signal processing*, 2nd ed. London: Academic Press, 1999.

[2] E. Hernández and G. Weiss, *A First Course on Wavelets*. CRC Press, 1996.

[3] J. Bergh, F. Ekstedt, and M. Lindberg, *Wavelets*. Lund, Sweden: Studentlitteratur, 1999.

[4] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 247–268, 1998.

[5] N. Grip, "Wavelet and Gabor frames and bases: Approximation, sampling and applications," Doctoral Thesis 2002:49, Luleå University of Technology, SE-971 87 Luleå, 2002, WWW: http://www.sm.luth.se/~grip/Research/publications.php.

[6] S. Ericsson and N. Grip, "An analysis method for sampling in shift-invariant spaces," *Int. J. Wavelets Multiresolut. Inf. Process.*, 2004, to appear. WWW: http://www.sm.luth.se/~grip/Research/publications.php.

[7] Y. Liu and G. G. Walter, "Irregular sampling in wavelet subspaces," *J. Fourier Anal. Appl.*, vol. 2, no. 2, pp. 181–189, 1995.

[8] G. G. Walter, "A sampling theorem for wavelet subspaces," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 881–884, Mar. 1992.

[9] D. Veitch, M. S. Taqqu, and P. Abry, "Meaningful MRA initialization for discrete time series," *Signal Process.*, vol. 80, no. 9, pp. 1971–1983, 2000.

[10] L. T. Dechevsky, N. Grip, and L.-E. Persson, "Sharp error estimates for approximation by wavelet frames in Lebesgue spaces," *J. Anal. Appl.*, vol. 1, 2003.

[11] ——, "Sharp error estimates for multivariate wavelet approximation in Besov and Triebel-Lizorkin spaces," 2004, in preparation.

[12] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Box 812060, Wellesley MA 02181 USA: Wellesley-Cambridge Press, 1996.

[13] J. Zhang and Z. Bao, "Initialization of orthogonal discrete wavelet transforms," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1474–1477, May 2000.

[14] S. Ehrich, "Sard-optimal prefilters for the fast wavelet transform," *Numer. Algorithms*, vol. 16, no. 3–4, pp. 303–319, 1997.

[15] B. R. Johnson and J. L. Kinsey, "Quadrature prefilters for the discrete wavelet transform," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 873–875, Mar. 2000.

[16] P. G. Lemarié-Rieusset, "On the existence of compactly supported dual wavelets," *Appl. Comput. Harmon. Anal.*, vol. 4, no. 1, pp. 117–118, Jan. 1997.

[17] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill, 1987.

[18] W. Sun and X. Zhou, "Sampling theorem for wavelet subspaces: Error estimate and irregular sampling," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 223–226, 2000.

[19] L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*. Siam, 1997.

[20] W. Chen, J. Chen, and S. Itoh, "The error estimation of sampling in wavelet subspaces," *IEICE T. Fund. Electr.*, vol. E82–A, no. 5, pp. 835–841, May 1999.

[21] N. Kaiblinger and W. Madych, "Orthonormal sampling functions," *Preprint*, 2004, WWW: http://www.mat.univie.ac.at/~kaib/preprints/sinc.ps.

PLACE PHOTO HERE

**Stefan Ericsson** was born in Ängelholm, Sweden in 1966. He received his Master of Science degree in Computer Science and Engineering from Luleå University of Technology, Sweden in 1991 and the Ph.D. degree in mathematics at the same university in 1997, with a thesis considering real interpolation theory. His major scientific interests are interpolation, function space, wavelets and applications to signal processing.

PLACE PHOTO HERE

**Niklas Grip** was born in Luleå, Sweden in 1970. He received his Master of Science degree in Computer Science and Engineering from Luleå University of Technology, Sweden in 1997 and the Ph.D. degree in mathematics at the same university in 2003. His major scientific interests are wavelet and Gabor theory with applications to signal processing.