Fast FIR algorithms for the continuous wavelet transform from constrained least squares

George M. Leigh, Member, IEEE

Copyright © 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work is part of AMIRA International Project P843, supported by the Australian Research Council and mining industry participants Anglo Platinum, AngloGold Ashanti, Barrick Gold, BHP Billiton, Codelco, Newcrest Mining, Newmont Mining, OZ Minerals, Peñoles, Rio Tinto, Teck Cominco, Vale, Vale Inco, Xstrata Copper, Datamine Group, Geotek, Golder Associates, ioGlobal, and Metso Minerals.

G. M. Leigh conducted this work at the University of Queensland, Sustainable Minerals Institute, Julius Kruttschnitt Mineral Research Centre, Isles Rd, Indooroopilly, Qld 4068, Australia. He is currently with the Department of Agriculture, Fisheries and Forestry, Ecociences Precinct Level 1B East, GPO Box 267, Brisbane, Qld 4001, Australia (phone: +617-3225-2550; fax: +617-3846-1207; e-mail: george.leigh@daff.qld.gov.au).

Abstract—New algorithms for the continuous wavelet transform are developed that are easy to apply, each consisting of a single-pass finite impulse response (FIR) filter, and several times faster than the fastest existing algorithms. The single-pass filter algorithm, named WT-FIR-1, is made possible by applying constraint equations to least-squares estimation of filter coefficients, which removes the need for separate low-pass and high-pass filters. Non-dyadic two-scale relations are developed and it is shown that filters based on them can work more efficiently than dyadic ones. Example applications to the Mexican hat wavelet are presented.

Index Terms—Algorithm design and analysis, Continuous wavelet transforms, Finite impulse response filter, Signal processing algorithms
I. INTRODUCTION

A. Background

Wavelet transforms have become indispensable tools in signal and image analysis over the last 25 years. Given a wavelet function $\psi(x)$, the wavelet transform of a signal $s(x)$ is ([1] p. 24)

$$Y(b, a) = \int_{-\infty}^{\infty} |a|^\frac{1}{2} \overline{\psi \left( \frac{x-b}{a} \right)} s(x) \, dx,$$

where $b$ and $a$ are location and scale parameters respectively, and the bar denotes complex conjugate.

Wavelet transforms are broadly classified into the continuous wavelet transform (CWT; [1] ch. 2), which offers great flexibility in the choice of wavelet shape and the sequence of scales of analysis, and the discrete wavelet transform (DWT; [1] ch. 3–6), which is restrictive on shape and scale but provides simple formulae for reconstruction of a signal from its wavelet transform.

The CWT for wavelets of general shape requires greater computational effort than the DWT. Fast algorithms, in which the number of arithmetic operations per wavelet coefficient depends on neither the length of the input signal nor the coarsest scale of analysis, exist for the CWT [2]–[5] but are slower and more difficult to implement than corresponding algorithms [6], [7] for the DWT. Alternatives for the CWT based on the fast Fourier transform (FFT) [8] are widely used but become slower for long signals: a signal of length $N$ combined with the real-data FFT algorithm of [9] requires roughly $\frac{12}{\pi} \log_2 N$ arithmetic operations per wavelet coefficient to perform the inverse FFT at each scale. This pure-FFT algorithm is not to be confused with the use of FFT to accelerate already-fast algorithms [2].

B. Two-scale relations

Fast algorithms for both the CWT and DWT rely on scaling functions and two-scale relations ([1] p. 78). The wavelet function $\psi(x)$ is expressed as a convolution with a scaling function $\phi(x)$ ([10] sec. II-C, slightly different notation),

$$\psi(x) = \sum_{k} \phi(x - 2^k),$$
\( \frac{1}{2} \psi(x/2) \approx \sum_{k=0}^{\infty} d_k \phi(x + k), \quad (2) \)

while the two-scale relation provides the scaling function at scale 2 from the same function at scale 1 through another convolution:

\[ \frac{1}{2} \phi(x/2) \approx \sum_{k=0}^{\infty} h_k \phi(x + k). \quad (3) \]

Equations (2) and (3) allow recursive computation of the wavelet transform using \( d_k \) and \( h_k \) as filter coefficients ([1], p. 79). The recursive equation, using (3), is

\[ S(b, 2^{n+1}) \approx \sqrt{2} \sum_{k=\infty}^{\infty} \overline{h}_k S(b - 2^n k, 2^n), \quad (4) \]

where

\[ S(b, a) = \int_{-\infty}^{\infty} |d| \frac{1}{\sqrt{a}} \phi((x-b)/a) \ s(x) \ dx, \]

and the wavelet transform at scale \( 2^{n+1} \) is given by

\[ Y(b, 2^{n+1}) \approx \sqrt{2} \sum_{k=\infty}^{\infty} \overline{d}_k S(b - 2^n k, 2^n). \quad (5) \]

Computation begins with initialization at the finest scale \( a = 1 \) and proceeds recursively through successively coarser scales using (4) and (5). Initialization is discussed in [11] and [12].

The DWT can be designed so that equations (2)–(5) work exactly and only a few of the filter coefficients \( d_k \) and \( h_k \) are nonzero [7]. For the CWT the equations are only approximate for most choices of \( \phi \) and \( \psi \), and the filter coefficients have to be chosen both to satisfy (2) and (3) as closely as possible and to allow \( d_k \) and \( h_k \) to be neglected for large \( |k| \) [3]. This problem is elucidated by taking Fourier transforms ([1] p. 132, slightly different notation): (3) becomes

\[ \hat{\phi}(2\xi) \approx m(\xi) \hat{\phi}(\xi), \quad (6) \]

where the filter function \( m(\xi) \) is given by

\[ m(\xi) = \sum_{k=\infty}^{\infty} \overline{h}_k e^{i\alpha \xi} \quad (7) \]
and is therefore periodic with period $2\pi$. The problem is to satisfy (6) as closely as possible with the restrictions that $m(\xi)$ must be periodic and that all but a few of the filter coefficients must be small enough to be neglected. A similar setup applies to the Fourier-transformed version of (2). A least-squares approach to solving this problem was undertaken in [3], which retained an infinite number of filter coefficients but showed that they decayed as $|k| \to \infty$.

C. FIR filters

When the filter coefficients $d_k$ and $h_k$ in (2)–(5) can be neglected for large $|k|$, calculation of the wavelet transform consists of applying finite impulse response (FIR) filters. This type of filter is very widely used in signal processing, and efficient implementations of it have been intensively studied.

Three types of FIR filter implementation are commonly used, depending on the length of the filter (i.e., the number of nonzero filter coefficients): (a) direct implementation using all the filter coefficients sequentially (e.g., direct use of formula (4)); (b) “divide and conquer” implementation [13]–[19], in which the filter and the input signal are first “decimated” into multiple series (e.g., odd versus even indices); and (c) FFT-FIR implementation, in which the input signal is divided into blocks of a certain length and each block undergoes FFT, multiplication by the filter’s Fourier transform, and finally inverse-FFT; use of this last technique suddenly exploded in the mid-1960s [20]. According to [19], the most efficient FIR implementation is direct for filter lengths 2, 3 and 5; divide-and-conquer for lengths 4 and 6–60; and FFT-FIR for length 64 or more. In the case of transforms applied over multiple scales, efficiency of FFT-FIR can be improved by avoiding some of the forward-FFT steps [21].

It is important to note that shorter FIR filters are always preferable to longer ones, whatever implementation method is used. An extra advantage of short FIR filters, in addition to computational efficiency, is that they allow transformations to be carried out in real-time with only a relatively short lag. This advantage continues to hold even at very large scales, because then the filters are applied to highly decimated series.
D. New developments

This paper aims to make the CWT more accessible to practitioners by implementing it with short-length FIR filters that, although underlain by complex theory, are fast and easy to use. This method is given the name \textit{WT-FIR-1}: the “1” refers to the capability to use a single filter in place of the separate low-pass and high-pass filters (4) and (5) when the wavelet \(\psi(x)\) is some order of derivative of its scaling function \(\phi(x)\). The new FIR filters can be implemented by any of the three methods described in section C above.

The major original contribution is to introduce constraint equations to the least-squares approach of Muschietti and Torrésani [3], in order to (a) explicitly limit the number of nonzero filter coefficients, rather than rely on decay as \(|k| \to \infty\), and (b) ensure regularity of adjusted scaling and wavelet functions that approximate \(\phi(x)\) and \(\psi(x)\). The regularity constraint comes from theory of iterated two-scale relations due to Daubechies and Lagarias [7], [22], [23]. The constraints used here focus on fast computation of the CWT with a given wavelet of general shape, and allow arbitrarily high precision to be achieved by the use of longer filters. Conceptually different least-squares constraints have been used in filter-bank design to ensure orthonormality similar to the DWT [24].

Non-dyadic two-scale relations, where the ratio of one scale to the previous scale is less than 2, are derived, and it is shown that these can work more efficiently than dyadic relations (in which the scale doubles each time). Non-dyadic two-scale relations have been used previously in (a) rational filter banks [24]–[26], which are similar to but lack some of the properties of wavelet transforms; (b) \(M\)-band wavelets which have scale ratios greater than 2 (see, e.g., [27]); (c) refinable functions [28]; (d) few other CWT-related applications, e.g., [29]. Usual algorithms for the CWT for non-dyadic scales still use dyadic two-scale relations [5], [30], or use pure-FFT implementation [31].

This paper also contributes methodology for initialization of the CWT, necessitated in part by the increased sensitivity of the single-pass filter technique to improper initialization.
Section II provides extra detail of existing theory. Sections III–VI respectively present the new elements: non-dyadic two-scale relations, the new algorithms, initialization techniques and example applications. Sections VII and VIII briefly discuss reconstruction of the original signal, and extensions to multidimensional wavelets.

II. PRELIMINARIES

A. Best existing fast algorithm

Many existing fast algorithms approximate the wavelet function $\psi(x)$ using two scaling functions $\phi_1(x)$ and $\phi_2(x)$ from the DWT, both of which satisfy two-scale relations of the form (3) exactly. The wavelet $\psi(x)$ is approximated by a linear combination of translates of $\phi_2(x)$, with coefficients derived from inner products of $\psi(x)$ with $\phi_1(x)$. Such algorithms are discussed in more generality in [4].

The fastest existing CWT algorithm for wavelets of general shape appears to be that of Vrhel, Lee and Unser [5], in which $\phi_1$ and $\phi_2$ are B-splines (see comparison in [32]). This algorithm is certainly faster than the pure-FFT algorithm for very long signals, but may not always be for signals of moderate length. It uses two FIR filters and one infinite impulse response (IIR) filter which, although quite efficient, make it less convenient than pure-FIR algorithms.

B. Vanishing moments

The number of vanishing moments $L$ of $\psi(x)$ is an important parameter in wavelet analysis ([1] pp. 241–247, 258–259). It is defined by $\int_{-\infty}^{\infty} x^\ell \psi(x) \, dx = 0$ for $\ell = 0, 1, 2, \ldots, L - 1$, with this integral being nonzero for $\ell = L$. A wavelet with $L = 1$ is commonly used to detect singularities such as edges, while higher values of $L$ provide analysis that is unaffected by polynomial trends in the signal.

Often the wavelet can be expressed as the $L$th derivative of its scaling function,

$$\psi(x) \propto \phi^{(L)}(x).$$

Then, by standard results on Fourier transforms of derivatives, the Fourier transform of $\psi(x)$,
\[ \hat{\psi}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi x} \psi(x) \, dx / \sqrt{2\pi}, \]
satisfies
\[ \hat{\psi}(\xi) \propto {\xi^L} \hat{\phi}(\xi). \]  
(9)

In this paper we choose the constants of proportionality so that
\[ \hat{\phi}(0) = 1. \]  
(10)

The developments in section IV assume that (8) holds, because this produces faster algorithms which do not require the dual filter (5). The new methods can still be used for wavelets that don’t satisfy (8), but they require dual filters and do not have the same margin of superiority in speed over existing algorithms.

Notable wavelets that do not meet the combination of conditions (8), (10) and (3) include wavelets of Gabor or Morlet type [33], [34] that have been made highly frequency selective by the insertion of a factor \( \exp(i\xi_0 x) \), with a large value of the modulation parameter \( \xi_0 \) (e.g., greater than half the bandwidth of \( \hat{\psi}(\xi) \)). These wavelets are not well suited to two-scale relations, and practitioners who particularly want to use them are usually better off to use pure-FFT implementation (see section I). Alternatively, a high-order derivative of the Gaussian function \( \exp(-\frac{1}{2}x^2) \), which is well suited to two-scale relations and to the algorithms developed in this paper, can be used instead.

III. NON-DYADIC TWO-SCALE RELATIONS

Scales are assumed to follow a geometric sequence with common ratio \( \gamma = 2^{1/J} \) for some positive integer \( J; J = 1 \) corresponds to dyadic scales, while larger values provide closer spacing of scales. This is the same scheme used in [5].

We introduce \( J \) different two-scale relations for use at scales \( \beta_j = \gamma^j \) for \( j = 0, \ldots, J - 1 \): (3) becomes
\[ \gamma^{-1} \phi(x/\gamma \beta_j) \approx \sum_{k=-\infty}^{\infty} h_{jk} \phi((x+k\beta_j) / \beta_j). \]  
(11)

Taking Fourier transforms, (11) becomes
\[ \hat{\phi}(\gamma \beta_j \xi) \approx m_j(\xi) \hat{\psi}(\beta_j \xi), \] (12)

where

\[ m_j(\xi) = \sum_{k=-\infty}^{\infty} h_{jk} e^{ik\xi}. \] (13)

As in the dyadic case, the filter functions \( m_j(\xi) \) are periodic with period \( 2\pi \). In this scheme, (4) is altered to

\[ S(b, \gamma 2^n \beta_j) \approx \sqrt{\gamma} \sum_{k=-\infty}^{\infty} \tilde{h}_{jk} S(b - 2^n k, 2^n \beta_j). \] (14)

If (8) holds, combining (9) with (12) yields

\[ \hat{\psi}(\gamma \beta_j \xi) \approx \gamma^j m_j(\xi) \hat{\psi}(\beta_j \xi), \]

from which

\[ \psi \left( x / \gamma \beta_j \right) \approx \gamma^{J-1} \sum_{k=-\infty}^{\infty} \tilde{h}_{jk} \psi \left( x + k / \beta_j \right). \] (15)

Substituting (15) into (1) provides an analog of (14) which applies to \( Y(b, a) \) directly and bypasses the need to calculate \( S(b, a) \):

\[ Y(b, \gamma 2^n \beta_j) \approx \gamma^{J-1} \sum_{k=-\infty}^{\infty} \tilde{h}_{jk} Y(b - 2^n k, 2^n \beta_j). \] (16)

To carry out the CWT, the user needs only a method of initialization (see section V) and a table of filter coefficients (e.g., Table III or V below).

If (8) does not hold, dual filter functions are required and the resulting algorithms will be slower. The two-scale relation (14) with a high-pass filter of the form (5) can be used; or, as discussed in section I-D, (14) can be replaced by the dyadic version (4) with \( J \) different high-pass filters for each usage of (4).

IV. DESIGN OF THE WT-FIR-1 ALGORITHM

A. Theory

We define adjusted scaling functions \( \tilde{\phi}_j(x) \) (\( j = 0, \ldots, J - 1 \)) so that their Fourier transforms, denoted \( M_j(\xi) \), satisfy the two-scale relation (12) exactly. The filter functions
m_j(ξ) in this relation will be defined shortly. Also define M_j(ξ) = M_0(ξ). The adjusted scaling functions \( \tilde{\phi}_j(x) \) produce adjusted wavelets \( \tilde{\psi}_j(x) \) via equations (8) and (9).

Defined in this way, the functions \( \tilde{\phi}_j(x) \) and \( \tilde{\psi}_j(x) \) are much more than mere approximations to \( \phi(x) \) and \( \psi(x) \), but are genuine scaling functions and wavelets in their own right, with finite numbers of filter coefficients and exact two-scale relations. As a result, nonzero filter coefficients do not have to be neglected during computation, and approximation errors (except numerical round-off) do not accumulate as the CWT scale becomes coarser.

We set the filter coefficients \( h_{jk} \) in (11) and (13) to be zero for \( |k| > K_j \), where the filter half-length \( K_j \) is chosen to make \( \tilde{\psi}_j(x) \) sufficiently close to \( \psi_j(x) \). As stated above, a close approximation to \( \psi(x) \) is not required for \( \tilde{\psi}_j(x) \) to produce a CWT; the only pertinent question is whether the shape of \( \tilde{\psi}_j(x) \) meets the user’s needs. This property is shared by the constructions in [5], although not by [3]. Larger values of \( K_j \) provide closer matches to \( \psi_j(x) \) but longer filters. Arbitrary precision can be achieved by a combination of large \( K_j \) and substitution of \( \psi(x) \) by \( \psi_{\alpha}(x) = \psi(x/\alpha)/\alpha \) for some potentially large factor \( \alpha \) (see discussion at end of section V).

The nonzero \( h_{jk} \) are estimated by constrained minimization of

\[
\sum_{j=1}^{J} \int_{-\infty}^{\infty} \xi^{2L} |\hat{\phi}(\xi) - M_j(\xi)|^2 d\xi,
\]

where \( L \) is the number of vanishing moments as in (9). This strategy follows the approach of [3], but targets the Fourier transform of the scaling function, \( \hat{\phi}(\xi) \), instead of the filter function \( m_j(\xi) \), and applies constraints to the minimization.

The constraints come from (10) and the need for regularity of the adjusted wavelets \( \tilde{\psi}_j(x) \). For the former, we require for each \( j \) that

\[
M_j(0) = m_j(0) = 1.
\]
The regularity constraints are that $m(\xi)$ must have a zero of some order $H_j$ at $\xi = \pi$.

Daubechies ([7], sec. 3.B) showed that iteration of the dyadic two-scale relation (6) converges to a well-defined scaling function $\tilde{\phi}(x)$ whose Fourier transform is

$$
\hat{\phi}(\xi) = \prod_{n=1}^{\infty} m(2^{-n} \xi)
$$

if $m(\xi)$ contains a zero of order $H$ (denoted $N$ in [7]) at $\xi = \pi$, and $H > 1 + \log_2 B$, where $B = \sup_{\xi \in R} |m(\xi)/\{1 + \exp(i\xi)\}|^{H}$. An extension to her argument shows that such iteration also converges to a well-defined wavelet function with Fourier transform $\hat{\psi}(\xi)$ given by (9) if $H > L + \max(1 + \log_2 B, 0)$. Often the supremum defining $B$ occurs at $\xi = 0$, in which case $H \geq L + 1$ suffices to ensure regularity of $\hat{\psi}(\xi)$. The regularity of the back-transformed wavelet function $\tilde{\psi}(x)$ is fully established only for the DWT in [7], but later work [22], [23] suffices to prove it for the CWT (see Theorem 3.1 of [23]).

For the non-dyadic two-scale relations used here, the above condition $H \geq L + 1$ on the order of the zero at $\xi = \pi$ becomes

$$
\sum_{j=1}^{J} H_j \geq L + 1. \tag{19}
$$

It is desirable to use the smallest possible values of $H_j$ in order to free up the filter coefficients to fit the desired wavelet function as closely as possible. It is clear from (19) that when non-dyadic two-scale relations are used, the required number of zeros can be spread out over the $J$ different “voices per octave” (scales per doubling of scale), which reduces the filter-length penalties for inclusion of the regularity constraints.

In terms of the filter coefficients $h_{jk}$ in (13), the constraint (18) is, for $j = 1, \ldots, J$,

$$
\sum_{k=-K_j}^{K_j} h_{jk} = 1, \tag{20}
$$

and the regularity constraints are, for $j = 1, \ldots, J$ and $n = 0, \ldots, H_j - 1$, 

10
\[ \sum_{k=-K_j}^{K_j} (-1)^k k^n h_{jk} = 0. \]  \hspace{1cm} (21)

The last equation is found by differentiating \( m_j(\xi) \) \( n \) times with respect to \( \xi \) and setting \( \xi = \pi \).

It is also used in [23].

**B. Evaluation of filter coefficients**

Replacing \( \gamma \beta \zeta \) by \( \zeta \) in (12) yields, for \( j = 1, 2, \ldots, J \),

\[ M_j(\xi) = m_{j-1}(\xi/2^{1/j}) M_{j-1}(\xi/2^{1/j}). \]  \hspace{1cm} (22)

Using (13) and (22), the quantity \( \hat{\phi}(\xi) - M_j(\xi) \) in (17) is

\[ \delta_j(\xi) = \hat{\phi}(\xi) - M_j(\xi/2^{1/j}) \sum_{k=-K_j}^{K_j} h_{jk} \exp(ik\xi/2^{(j+1)/j}). \]

We ignore the dependence of \( M_j \) on \( h_{jk} \) and differentiate (17) with respect to the real and imaginary parts of \( h_{jk} \); for \( j = 0, \ldots, J-1 \) and \( k = -K_j, \ldots, K_j \),

\[
\int_{-\infty}^{\infty} \xi^{2jL} M_j(\xi/2^{1/j}) \exp(-ik\xi/2^{(j+1)/j}) \delta_j(\xi) d\xi = \sqrt{2\pi} \left( \lambda_j + (-1)^j \sum_{n=0}^{H_j-1} k^n \mu_{jn} \right), \]  \hspace{1cm} (23)

where \( \lambda_j \) and \( \mu_{jn} \) are Lagrange multipliers that account for constraints (20) and (21); the scaling by \( \sqrt{2\pi} \) is for convenience. The left-hand side of (23) can be written as a sum of inverse Fourier transforms to provide the following equation:

\[
2^{(2j+1)/j} \sum_{\ell=-K_j}^{K_j} h_{j \ell} f_j(\ell-k)/2^{(j+1)/j} + \lambda_j + (-1)^j \sum_{n=0}^{H_j-1} k^n \mu_{jn} = \bar{g}_j(k/2^{(j+1)/j}), \]  \hspace{1cm} (24)

where \( f_j(x) \) and \( g_j(x) \) are the inverse Fourier transforms of \( \varepsilon^{2jL} \left| M_j(\xi) \right|^2 \) and \( \varepsilon^{2jL} \bar{\phi}(\xi) M_j(\xi/2^{1/j}) \) respectively. This takes the form of a linear system, except that the functions \( f_j \) and \( g_j \) themselves depend on \( h_{jf} \) because \( M_j \) does. To solve it, we can either replace these functions by others which don’t depend on \( h_{jf} \), or solve iteratively, using the previous values of \( h_{jf} \) in \( f_j \) and \( g_j \) to solve for new values until the two converge. The latter
method requires some functional form to be assumed for $M_j$ in order to evaluate the inverse Fourier transforms.

A straightforward way to solve (24) is to recall that $M_j(\xi)$ is designed to estimate $\hat{\phi}(\xi)$, and then approximate $f_j(x)$ and $g_j(x)$ by their values when $M_j(\xi) = \hat{\phi}(\xi)$. These are functions $F(x)$ and $G(x)$ respectively, whose Fourier transforms are

$$
\hat{F}(\xi) = \xi^{2j} \left| \hat{\phi}(\xi) \right|^2, \quad \hat{G}(\xi) = \xi^{2j} \overline{\hat{\phi}(\xi)} \hat{\phi}(\xi / 2^{1/2}).
$$

(25)

Then $h_{jk}$ can be found by solving a linear system for each value of $j$. In many cases, expressions for $F(x)$ and $G(x)$ can be found analytically.

Alternatively, a more precise, iterative method of solution is outlined in Appendix A. In the examples presented in section VI, the approximation of $f_j(x)$ and $g_j(x)$ by $F(x)$ and $G(x)$ as above yields integrated squared errors and maximum absolute errors that are both between 100% and 200% of the values for the precise method. These differences are small relative to the precision gained from extra filter coefficients. Therefore approximation by $F(x)$ and $G(x)$ is adequate for practical purposes. We emphasize that lower accuracy in the solution of (24) affects only the closeness of the adjusted wavelet $\tilde{\psi}(x)$ to the desired wavelet $\psi(x)$, and does not propagate errors or affect the validity of $\tilde{\psi}(x)$ as a wavelet in its own right, because (20) and (21) are still satisfied exactly.

V. Initialization

Wavelet transform algorithms that rely on two-scale relations have to be initialized at some fine scale to provide a starting point for the two-scale relations. Fortunately, at fine scales the wavelet is well localized in space or time, and usually initialization is not a computational burden: direct use of formula (1) involves only a few values of $x$ in calculating each $Y(b, a)$ at small values of $a$.

It is important to ensure that the values assigned at initialization do not later cause trouble during iteration of the two-scale relations. It is also desirable to avoid spurious high frequencies by interpolating the signal $s(x)$ from its sampled values at integer points $x$. 


Because of the single-pass nature of the WT-FIR-1 filters, it is essential that the first \( L \) moments in the version of \( \psi(x) \) used for initialization must vanish. If initialization is carried out by direct use of (1), this means that the numerical sums representing the integral must be evaluated to the limit of numerical precision; small terms must not be neglected until their contributions are numerically zero. An FFT version of (1) could be used if \( \hat{\psi}(\xi) \) has a simple closed form but \( \psi(x) \) does not; initialization can still take advantage of the spatial localization of \( \psi(x) \) at fine scales to apply the FFT only over moderate-sized blocks of the signal rather than the whole signal [2]. Alternatively, \( \psi(x) \) can be approximated by, e.g., a cubic B-spline in which the first \( L \) moments vanish and the \( L \)th moment matches the \( L \)th moment of \( \psi(x) \).

Interpolation of the signal \( s(x) \) for use in integration can be handled by a reconstruction filter (see, e.g., [35]). For example, if \( \psi(x) \) is based on the Gaussian scaling function \( \exp(-\frac{1}{2}x^2) \), interpolation by a Gaussian reconstruction filter kernel makes the integration straightforward.

Related to initialization of the CWT is the matter of scales below which two-scale relations cannot be made sufficiently accurate. The practitioner is welcome to choose any range of scales of analysis, but very fine ones are not amenable to two-scale relations. The initialization scale, nominally \( a = 1 \), can be changed to effectively \( a = \alpha \) by replacing \( \psi(x) \) by \( \psi_{\alpha}(x) = \psi(x/\alpha)/\alpha \) for some factor \( \alpha > 0 \) (see examples below). This technique has been used by previous researchers [3], [5]. However, small values of \( \alpha \) reduce the accuracy achievable in a two-scale relation (see increasing values of \( \alpha \) in Tables I and IV below). As noted at the beginning of this section, accurate results for small \( a \) can be calculated rapidly without two-scale relations, by the same means as initialization.

VI. Example Applications to the Mexican Hat Wavelet

A. Dyadic scales

The Mexican hat or Laplacian of Gaussian wavelet has the form:

\[
\psi_{\alpha}(x) = \left(1 - x^2/\alpha^2\right)e^{-x^2/\alpha^2}/\alpha, \quad \hat{\psi}_{\alpha}(\xi) = \alpha^2 \xi^2 e^{-\alpha^2\xi^2},
\]

(26)
where \( \alpha \) is a scaling parameter whose value we can choose to obtain adequate CWT precision (see sections IV-A and V).

The Mexican hat wavelet has two vanishing moments; i.e., \( L = 2 \) in (9). The scaling function is

\[
\phi_\alpha(x) = e^{-\frac{x^2}{\alpha^2}} / \alpha, \quad \hat{\phi}_\alpha(\xi) = e^{-\frac{\xi^2}{\alpha^2}}.
\]

The functions \( F(x) \) and \( G(x) \) in (25) can be found analytically:

\[
F(x) = \left\{ 3 - 3 \frac{x^2}{\alpha^2} + \frac{x^4}{(4\alpha^4)} \right\} e^{-\frac{x^2}{\alpha^2}} / \left(4\sqrt{2}\alpha^3\right),
\]

\[
G(x) = 32 \left\{ 3 - 24 \frac{x^2}{(5\alpha^2)} + 16 \frac{x^4}{(25\alpha^4)} \right\} e^{-\frac{x^2}{\alpha^2}} / \left(25\sqrt{5}\alpha^5\right).
\]

Because the scales are dyadic, i.e., \( J = 1 \), we omit the subscript \( j \): there is only one filter function \( m(\xi) \), one adjusted scaling function \( \tilde{\phi}_\alpha(x) \) whose Fourier transform \( M(\xi) \) approximates \( \hat{\phi}_\alpha(\xi) \), and one adjusted wavelet \( \tilde{\psi}_\alpha(x) \) derived from \( \tilde{\phi}_\alpha(x) \).

The value \( H = 3 \), as recommended by (19), was used for this wavelet, forcing \( m(\xi) \) to have a zero of order 3 at \( \xi = \pi \).

For various values of \( K \) providing different levels of precision, filter coefficients \( h_{-K}, \ldots, h_K \) and Lagrange multipliers \( \lambda, \mu_0, \mu_1 \) and \( \mu_2 \) were estimated from (24), with \( M(\xi) \) estimated as in Appendix A. The scaling parameter \( \alpha \) was chosen by trial and error to minimize the integrated squared error (17).

A low-accuracy WT-FIR-1 approximation to the Mexican hat wavelet is graphed in Fig. 1 in the frequency domain. This approximation uses only seven filter coefficients. The adjusted wavelet \( \tilde{\psi}_\alpha(x) \) satisfies the two-scale relation (15) exactly. Even for this very rough approximation, the only cause for concern is the bulge of high frequency around \( \xi = \pm 10 \).

The accuracies of the WT-FIR-1 algorithms and the best existing fast algorithms [5] are compared in Table I. The numbers of arithmetic operations listed there can also be compared to those of the pure-FFT method discussed in section I, listed in Table II; these depend on
signal length rather than filter length. The principal measure of accuracy used is the integrated squared error (ISE),

\[ \alpha^2 \int_{-\infty}^{\infty} \xi^4 |\hat{\phi}_\alpha(\xi) - M(\xi)|^2 d\xi = \int_{-\infty}^{\infty} \xi^4 |\hat{\phi}_\alpha(\xi) - M(\xi/\alpha)|^2 d\xi, \]

where \( \hat{\phi}_\alpha(\xi) = e^{i\xi^2} \), which does not depend on the scaling parameter \( \alpha \): the inclusion of \( \alpha \) in the ISE in this way allows ISEs with different values of \( \alpha \) to be validly compared. By Parseval’s theorem from Fourier analysis, it is also the integrated squared error of estimation of the wavelet function \( \psi_t(x) = (1-x^2)e^{i\xi x^2} \). Table I also lists the maximum absolute difference between \( \hat{\psi}_\alpha(\xi) \) and \( \hat{\psi}_\alpha(\xi) \).

The algorithm of [5] involves three separate filters: two FIR filters and one infinite impulse response (IIR) filter. The number of arithmetic operations listed in Table I takes this into account, and is larger than the corresponding value for a single FIR filter with the same number of coefficients.

It can be seen from Tables I and II that the WT-FIR-1 algorithm is roughly three times as fast as that of [5], and (for moderate accuracy) between two and three times as fast as the pure-FFT method. The WT-FIR-1 algorithm is also simpler to apply than [5], consisting of a single FIR filter, and in most cases uses smaller values of the scaling parameter \( \alpha \), which allow use of the two-scale relation at smaller scales.

Filter coefficients for various levels of accuracy are listed for reference in Table III.

For the case of 11 filter coefficients, the difference between the wavelet Fourier transform \( \hat{\psi}_\alpha(\xi) \) and its estimate \( \hat{\psi}_\alpha(\xi) \) is plotted in Fig. 2. For \(-15 \leq \xi \leq 15\) this difference oscillates with roughly constant amplitude, but outside that range the peaks decrease rapidly, resulting in regularity of \( \psi_\alpha(\xi) \).

Such regularity contrasts with that in Fig. 3, which plots the difference when the constraint (21) is dropped, allowing \( M(\pi) \) to be nonzero. If the integrated squared error over any finite range (here \(-40 \leq \xi \leq 40\)) is minimized, the peaks increase exponentially in magnitude.
integrated squared error over \(-\infty < \xi < \infty\) is infinite. Thus the condition (21) is critical to correct calculation of filter coefficients.

Filter coefficients \(h_k\) in Table III decay fairly rapidly as the index \(k\) increases, before they become zero for \(|k| > K\). This is related to the decay of an infinite sequence of such coefficients proven in [3]. Because the filters developed here are finite and short, decay is not essential to stability of the algorithms.

**B. Four voices per octave**

At four voices per octave, the scale increases by a factor of only \(\sqrt{2} \approx 1.189\) with each application of the two-scale relation, providing finer resolution than the dyadic sequence.

The function \(F(x)\) is the same as in section A above. Now \(J = 4\) and \(G(x)\) is given by

\[
G(x) = \left(6 - 4\sqrt{2}\right)\sqrt{2 - \sqrt{2}} \left(3 - 6\left(2 - \sqrt{2}\right)x^2/\alpha^2 + \left(6 - 4\sqrt{2}\right)x^4/\alpha^4\right)e^{-\left(|\xi|/\alpha\right)^x}/\alpha^4.
\]

There are four filter functions \(m_j(\xi)\) \((j = 0, 1, 2, 3)\) and four functions \(M_j(\xi)\) that approximate \(\hat{\phi}_0(\xi)\) (see (13) and (17)).

For regularity, (19) recommends three zeros per octave as a safe number. Judging from plots similar to Figs 2 and 3, however, two zeros per octave sufficed in most cases: \(H_j\) was set to 1 for \(j = 1\) and 2, and zero for \(j = 0\) and 3, except at the lowest level of precision (first row of Table IV) when \(H_j\) was set to 1. For various filter half-lengths \(K_j\) \((j = 0, 1, 2, 3)\), the coefficients \(h_{j-\cdot\cdot\cdot\cdot}, \ldots, h_{j\cdot\cdot\cdot\cdot}\) and Lagrange multipliers \(\lambda_j\) and \((\text{for } j = 1, 2)\) \(\mu_{j\cdot\cdot\cdot\cdot}\) were estimated from (24), with \(M_j(\xi)\) estimated as in Appendix A. Numbers of filter coefficients used in the WT-FIR-1 algorithm and that of [5] are compared in Table IV.

From Table IV it can be seen that the WT-FIR-1 algorithm is again roughly three times as fast as that of [5]. As was also the case in section A above, the new method has smaller values of the scaling parameter \(\alpha\), allowing use of two-scale relations at smaller scales.

Non-dyadic two-scale relations have provided extra speed. Comparison of Table I with Table IV shows that, for similar accuracy, non-dyadic two-scale relations save about one-
third in the number of filter coefficients per scale. The WT-FIR-1 algorithm is about three times faster than the pure-FFT method for accuracies of practical interest (see Table II).

Filter coefficients for the case of six coefficients per scale are listed for reference in Table V.

VII. RECONSTRUCTION OF THE SIGNAL

If reconstruction of the original signal from its CWT is required, it is convenient to use wavelet packets [36], which are versions of the CWT partially integrated over scale. The packet wavelet function is

\[ \Psi(x) = \int_{a}^{\infty} \psi(x/a) \, da/a^2. \]

Low frequencies are integrated into the function

\[ \Phi(x) = \int_{a}^{\infty} \psi(x/a) \, da/a^2. \]

To facilitate the algorithms developed in this paper, the function

\[ \Phi_j(x) = \left\{ L/(1 - \gamma^{-L}) \right\} \int_{a}^{\infty} a^j \phi(x/a) \, da/a^2 \]

can be used as a scaling function, where \( \gamma = 2^{1/j} \) as in section III; \( \Phi_j(x) \) is defined so that \( \hat{\Psi}(\xi) \propto \xi^j \hat{\Phi}_j(\xi) \) and \( \hat{\Phi}_j(0) = 1 \). Implementation of wavelet packets is then identical to the raw CWT with \( \Psi \) and \( \Phi_j \) in place of \( \psi \) and \( \phi \).

VIII. MULTIDIMENSIONAL WAVELETS

The WT-FIR-1 algorithms are well suited to some multidimensional wavelets. A feature of the single-pass filters is that in \( d \) dimensions they are very nearly isotropic when the scaling function \( \phi(\vec{x}) \) is isotropic, but they are also applicable to anisotropic wavelets whose Fourier transforms take the form

\[ \hat{\psi}(\vec{\xi}) = \hat{\phi}(\vec{\xi}) \prod_{n=1}^{d} P_{n,j}(\xi_n), \]

in which \( \vec{\xi} \) are the frequency coordinates corresponding to Cartesian coordinates \( x_j \), and \( P_{n,j} \) are all polynomials of degree \( L \) or less. The isotropic nature of the single-pass filters allows
them to be applied to rotated versions of the wavelet, with no need for matching rotation of
the filter.

A simple example is provided by the two-dimensional derivative-of-Gaussian (DOG)
wavelet

\[ \psi(x_1, x_2) = x_1 \exp \left\{ -\frac{1}{2}(x_1^2 + x_2^2) \right\} \]

with dyadic scales. A 1-D single-pass filter \( \{h_k\} \) can be derived in the same way as in section
VI.A, but with \( L = 1 \); or indeed the filters from there (Table III) can be used unaltered because
they apply to any wavelet based on the Gaussian scaling function with derivatives up to order
\( L = 2 \). In two dimensions the two-scale relation can be implemented by applying the filter in
the \( x_1 \) direction, and applying the same filter in the \( x_2 \) direction. If \( \psi(\bar{x}) \) is rotated by an angle
\( \theta \) to become

\[ \psi_\theta(x_1, x_2) = (x_1 \cos \theta + x_2 \sin \theta) \exp \left\{ -\frac{1}{2}(x_1^2 + x_2^2) \right\}, \]

the same filters can still be used for \( \psi_\theta(\bar{x}) \), i.e., applying the 1-D filter in the \( x_1 \) and \( x_2 \)
directions, with no need to use a rotated version of the filter.

The new algorithms can find application to the plethora of extensions of the CWT,
colloquially known as “X-lets” (see review in [37]). The application of isotropic filters to
anisotropic X-lets can be pursued provided that the scaling functions of the X-lets remain
isotropic. An X-let to which the isotropic filter approach described above is not applicable is
the curvelet transform [38], because it has parabolic scaling, expanding the length and width
scales by different ratios, which results in anisotropic scaling functions. Curvelets can make
use of the non-dyadic two-scale relations of section III to provide different two-scale relations
for length and width, but the filters would have to be rotated to be aligned to the curvelet’s
length and width rather than to Cartesian axes.

IX. CONCLUSION

The single-pass form of the WT-FIR-1 filters raises the question of how far they might be
developed for use in the DWT. A point of interest is that, when single-pass filters are used, all
the information present in the wavelet transform over all scales must be contained in the
wavelet transform at the finest scale, because the two-scale relations use no other source.
Single-pass filters for the DWT would therefore have to involve redundancy of information,
but it is still possible that computation of the single-pass DWT could be made faster than the
standard, non-redundant DWT that uses dual filters.

The computations for this paper were performed in the software R [39]. The R code, and
Fortran code for computing the CWT, are available as supplemental material and from
www.mathematicaldiscovery.net. Suggestions of additional wavelets for which practitioners
require filter coefficients (as in Tables III and V) are welcome on the latter site.

APPENDIX A. PRECISE EVALUATION OF FILTER COEFFICIENTS

A. Another periodic estimating function

In this section we estimate \( M_j(\xi) \) using a periodic function with period \( 2^{1+j/2} \pi \), which
matches the periodicity of \( m_j(\xi/2^J) \) in (22). For \( j = 1, \ldots, J \), we set

\[
M_j(\xi) = \hat{\phi}(\xi) \sum_{r=-R}^{R} a_j, \exp(-ir\xi/U_j)
\]

for some large integer \( R \), where

\[
U_j = 2^{j/2}.
\]

We also define \( a_0 = a_J \) and \( U_0 = U_J \), so that (27) holds for \( j = 0 \), but (28) does not.

The coefficients \( a_j \) in (27) are redefined each time (22) is applied in updating \( M_j(\xi) \).
They are evaluated by a similar method to that for \( h_{jk} \) in section IV-B. We choose \( a_j \) to
minimize the integrated squared difference

\[
\int_{-\infty}^{\infty} \xi^2 \left| M_j(\xi) - \hat{\phi}(\xi) \sum_{r=-R}^{R} a_j, \exp(-ir\xi/U_j) \right|^2 d\xi,
\]

subject to the constraint

\[
\sum_{r=-R}^{R} a_j = 1
\]
which ensures that (18) is satisfied and gives rise to a Lagrange multiplier denoted $\nu_j$. With substantial labor, using (22), and (13) and (27), with $j$ replaced by $j - 1$, the following recursive linear system is found which, in conjunction with (29), can be solved for $\{a_{jr}\}$ and $\nu_j$ in terms of $\{a_{j-1r}\}$: For $j = 1, \ldots, J$, and $r = -R, \ldots, R$,

$$
\sum_{r=-R}^{R} a_{jr} F\left( (r-t) / U_j \right) + \nu_j = \sum_{k=-K_j}^{K_j-1} \sum_{r=-R}^{R} h_{j-1k} a_{j-1r} G\left( (k+r) / U_j - t / (2^{1/J} U_{j-1}) \right),
$$

where $F(x)$ and $G(x)$ are defined by their Fourier transforms (25). Solution is iterated, starting again at $j = 1$ after solving for $j = J$, until convergence is achieved.

The value $R = 15$ was used for most of the applications in section VI, but had to be reduced, down to 6 in two cases, for some of the larger values of $K$ because the linear system for $a_{jr}$ was very close to singular.

B. Equations for filter coefficients

We use the above periodic estimate of $M_j(\xi)$. From (27), the function $g_j(x)$ in (24) is:

$$
g_j(x) = \int_{-\infty}^{\infty} q^2 e^{j2\pi \xi x} \tilde{\phi}(\xi) M_j(\xi / 2^{1/J}) d\xi / \sqrt{2\pi} = \sum_{r=-R}^{R} a_{jr} G(x - r / \{2^{1/J} U_j\}).
$$

Then the equation (24) for $h_{j\ell}$ becomes, for $j = 0, \ldots, J - 1$ and $k = -K_j, \ldots, K_j$,

$$
2^{(2L+1)/J} \sum_{r=-R}^{R} h_{j\ell} \sum_{r=-R}^{R} \sum_{r=-R}^{R} \tilde{a}_{jr} a_{jr} F\left( (\ell - k) / 2^{1/J} + (r-t) / U_j \right) + \lambda_j + (-1)^{J-1} \sum_{h=0}^{J-1} k^h \mu_{jh}
$$

$$
= \sum_{r=-R}^{R} \tilde{a}_{jr} \tilde{G}(k / U_{j+1} - r / \{2^{1/J} U_j\}).
$$

ACKNOWLEDGMENT

Dr Stephen Gay and two anonymous reviewers provided comments on the style and structure of this paper, which substantially improved its readability and breadth of reference to related work.

REFERENCES


## TABLE I

**ACCURACY OF WT-FIR-1 ALGORITHMS FOR THE MEXICAN HAT WAVELET WITH DYADIC SCALES, AND COMPARISON WITH BEST EXISTING METHOD**

<table>
<thead>
<tr>
<th></th>
<th>WT-FIR-1</th>
<th>Best existing$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td># filter coef.$^b$</td>
<td>$\alpha^c$</td>
<td>Integrated squared error$^d$</td>
</tr>
<tr>
<td>7 (9.8)</td>
<td>0.71</td>
<td>$3.6 \times 10^{-3}$</td>
</tr>
<tr>
<td>9 (12.3)</td>
<td>0.80</td>
<td>$9.7 \times 10^{-5}$</td>
</tr>
<tr>
<td>11 (14.2)</td>
<td>0.87</td>
<td>$4.2 \times 10^{-6}$</td>
</tr>
<tr>
<td>13 (16.4)</td>
<td>0.94</td>
<td>$2.1 \times 10^{-7}$</td>
</tr>
<tr>
<td>15 (18.3)</td>
<td>1.00</td>
<td>$6.7 \times 10^{-9}$</td>
</tr>
<tr>
<td>17 (20.3)</td>
<td>1.06</td>
<td>$3.6 \times 10^{-10}$</td>
</tr>
<tr>
<td>19 (22.5)</td>
<td>1.11</td>
<td>$1.1 \times 10^{-11}$</td>
</tr>
<tr>
<td>21 (23.9)</td>
<td>1.16</td>
<td>$6.2 \times 10^{-13}$</td>
</tr>
<tr>
<td>23 (24.8)</td>
<td>1.21</td>
<td>$2.0 \times 10^{-14}$</td>
</tr>
</tbody>
</table>

$^a$Method of Vrhel, Lee and Unser [5]; settings are chosen to provide the same accuracy as WT-FIR-1.

$^b$Numbers of arithmetic operations to implement are given in parentheses, using the best of the FIR implementation methods discussed in section I, and symmetry of coefficients.

$^c$Scaling factor, chosen to minimize the integrated squared error; smaller values allow application of the two-scale relation down to finer scales.

$^d$Precision of approximation of the Fourier transform of the wavelet.
### TABLE II

**NUMBERS OF ARITHMETIC OPERATIONS FOR THE PURE-FFT IMPLEMENTATION OF THE CWT**

<table>
<thead>
<tr>
<th>(N^a)</th>
<th>1024</th>
<th>32,768</th>
<th>1,048,576</th>
<th>32M</th>
<th>1024M</th>
</tr>
</thead>
<tbody>
<tr>
<td># ops</td>
<td>20.0</td>
<td>29.6</td>
<td>39.1</td>
<td>48.6</td>
<td>58.1</td>
</tr>
</tbody>
</table>

\(^a\)Length of signal \((1M = 1,048,576)\)

\(^b\)Number of operations per wavelet coefficient, to compare to numbers in parentheses in Tables I and IV; algorithm of [9]

### TABLE III

**FILTER COEFFICIENTS FOR THE MEXICAN HAT WAVELET, DYADIC SCALES\(^a\)**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>9 coef.</th>
<th>11 coef.</th>
<th>13 coef.</th>
<th>15 coef.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>0.80</td>
<td>0.87</td>
<td>0.94</td>
<td>1.00</td>
</tr>
<tr>
<td>(h_0)</td>
<td>0.2880</td>
<td>0.264688</td>
<td>0.245040</td>
<td>0.230326</td>
</tr>
<tr>
<td>(h_{+1})</td>
<td>0.2221</td>
<td>0.212511</td>
<td>0.202908</td>
<td>0.194973</td>
</tr>
<tr>
<td>(h_{+2})</td>
<td>0.1019</td>
<td>0.109786</td>
<td>0.115239</td>
<td>0.118256</td>
</tr>
<tr>
<td>(h_{+3})</td>
<td>0.0279</td>
<td>0.036542</td>
<td>0.044891</td>
<td>0.051397</td>
</tr>
<tr>
<td>(h_{+4})</td>
<td>0.0041</td>
<td>0.007870</td>
<td>0.011984</td>
<td>0.016009</td>
</tr>
<tr>
<td>(h_{+5})</td>
<td>0.000947</td>
<td>0.002201</td>
<td>0.003569</td>
<td></td>
</tr>
<tr>
<td>(h_{+6})</td>
<td>0.000257</td>
<td>0.000572</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(h_{+7})</td>
<td></td>
<td>0.000061</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)For use in the two-scale relation (16) with \(\gamma = 2\) and \(\beta_j = 1\); rounding of the final digits has been chosen to satisfy constraints (20) and (21) exactly.
## Table IV

### Accuracy of WT-FIR-1 Algorithms for the Mexican Hat Wavelet with Four Scales per Octave

<table>
<thead>
<tr>
<th># filter coef.</th>
<th>$\alpha$</th>
<th>Integrated squared error</th>
<th>Maximum error</th>
<th># filter coef.</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5 (6.3)</td>
<td>1.02</td>
<td>$1.1 \times 10^{-3}$</td>
<td>$2.1 \times 10^{-2}$</td>
<td>12.8 (18.3)</td>
<td>1.15</td>
</tr>
<tr>
<td>6 (8.4)</td>
<td>1.10</td>
<td>$4.4 \times 10^{-5}$</td>
<td>$3.5 \times 10^{-3}$</td>
<td>18.8 (24.2)</td>
<td>1.58</td>
</tr>
<tr>
<td>8 (11.0)</td>
<td>1.21</td>
<td>$6.9 \times 10^{-7}$</td>
<td>$5.2 \times 10^{-4}$</td>
<td>29.8 (32.3)</td>
<td>2.29</td>
</tr>
<tr>
<td>10 (13.2)</td>
<td>1.37</td>
<td>$8.4 \times 10^{-9}$</td>
<td>$5.6 \times 10^{-5}$</td>
<td>52.3 (37.3)</td>
<td>3.81</td>
</tr>
<tr>
<td>12 (15.3)</td>
<td>1.50</td>
<td>$7.9 \times 10^{-11}$</td>
<td>$5.7 \times 10^{-6}$</td>
<td>96.3 (41.5)</td>
<td>6.6</td>
</tr>
<tr>
<td>14 (17.3)</td>
<td>1.63</td>
<td>$1.0 \times 10^{-12}$</td>
<td>$6.5 \times 10^{-7}$</td>
<td>172.3 (45.3)</td>
<td>11.1</td>
</tr>
<tr>
<td>16 (19.3)</td>
<td>1.74</td>
<td>$1.0 \times 10^{-14}$</td>
<td>$5.6 \times 10^{-8}$</td>
<td>326.8 (49.4)</td>
<td>19.9</td>
</tr>
</tbody>
</table>

*Method of [5]; settings for the same accuracy as WT-FIR-1

$^b$Average number per scale; average numbers of arithmetic operations to implement are given in parentheses.

$^c$Scaling factor, chosen to minimize the integrated squared error; smaller values allow application of the two-scale relation down to finer scales.

$^d$Precision of approximation of the Fourier transform of the wavelet function; maximum over the four scales in the octave
digits has been chosen to satisfy constraints (20) and (21) exactly.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$j = 0$</th>
<th>$j = 1$</th>
<th>$j = 2$</th>
<th>$j = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>1.10</td>
<td>1.10</td>
<td>1.10</td>
<td>1.10</td>
</tr>
<tr>
<td>$h_{j,0}$</td>
<td>0.5552</td>
<td>0.4598</td>
<td>0.3946</td>
<td>0.4628</td>
</tr>
<tr>
<td>$h_{j,1}$</td>
<td>0.2128</td>
<td>0.2452</td>
<td>0.2453</td>
<td>0.1311</td>
</tr>
<tr>
<td>$h_{j,2}$</td>
<td>0.0096</td>
<td>0.0201</td>
<td>0.0527</td>
<td>0.1375</td>
</tr>
<tr>
<td>$h_{j,3}$</td>
<td>0.0048</td>
<td>0.0047</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*For use in the two-scale relation (16) with $\gamma = \sqrt{2}$ and $\beta_j = 2^j$; rounding of the final digits has been chosen to satisfy constraints (20) and (21) exactly.

Fig. 1. Fourier transforms of the Mexican hat wavelet, $\hat{\phi}(\xi)$, and a low-accuracy (seven filter coefficients) adjusted wavelet, $\hat{\psi}(\xi)$, which has an exact dyadic two-scale relation and a zero of order 3 at $\xi = 2\pi$. 
Fig. 2. Difference between the Fourier transforms of the adjusted and true Mexican hat wavelet functions, for dyadic scales with 11 filter coefficients: (a) overview; (b) right-hand tail, showing two dyadically-spaced series of peaks $A_n$ and $B_n$ with exponentially decreasing sizes, confirming that the adjusted wavelet function is well-behaved.
Fig. 3. Difference between the Fourier transforms of the adjusted and true Mexican hat wavelet functions, for dyadic scales with 11 filter coefficients, without the constraint (21) on the filter coefficients. In contrast to the good behavior of Fig. 2, the peaks are now exponentially increasing.

George M. Leigh (S’07–M’10) received the B.Sc. (Hons) degree in mathematics in 1984 and the Ph.D. degree in mineral processing science in 2009, both from the University of Queensland, Australia.

He has worked as a mathematician since 1985, specializing in applications to mineral processing and fisheries. Since 2009 he has been a member of the Fisheries Resource Assessment team at the Queensland Department of Agriculture, Fisheries and Forestry. His research interests include image analysis, stereology, statistics, and fish population dynamics.