# A generalized demodulation approach to time-frequency projections for multicomponent signals

By S. Olhede and A.T. Walden

*Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2BZ, UK.*

We introduce a flexible approach for the time-frequency analysis of multicomponent signals involving the use of analytic vectors and demodulation. The demodulated analytic signal is projected onto the time-frequency plane so that, as closely as possible, each component contributes exclusively to a different 'tile' in a wavelet packet tiling of the time-frequency plane, and at each time instant the contribution to each tile definitely comes from no more than one component. A single reverse demodulation is then applied to all projected components. The resulting instantaneous frequency of each component in each tile is not constrained to a set polynomial form in time, and is readily calculated, as is the corresponding Hilbert energy spectrum. Two examples illustrate the method.

In order to better understand the effect of additive noise, the approximate variance of the estimated instantaneous frequency in any tile has been formulated by starting with pure noise and studying its evolving covariance structure through each step of the algorithm. The validity and practical utility of the resulting expression for the variance of the estimated instantaneous frequency is demonstrated via a simulation experiment.

**Keywords: demodulation; Hilbert spectrum; instantaneous frequency; multicomponent signals; wavelet packets; wavelet transform**

## 1. Introduction

The analysis of signals with time-evolving oscillatory structure is an area of active interest. Basic representations of such signals are bilinear distributions (e.g., Cohen 1995) which are calculated by filtering the signal and its complex conjugate with a kernel function. This generally entails problems for signals with multiple components unless the kernel decays very rapidly in time-frequency, thus avoiding damaging interference effects. Our aim in this work is to introduce a simple approach to the problem of tracking the time-dependent frequency content of each component of a multicomponent signal via the undecimated discrete wavelet packet transform. It is well-known that wavelet transforms 'tile' the time-frequency plane in a proportional bandwidth or octave band manner. The extension to wavelet packets allows more irregular, but still rectangular partitioning of the time-frequency plane. Such rectangular tilings of the time-frequency plane are not best suited to representing some classes of signals (such as FM signals) which may have energy distributions which are slanted, non-linear etc in the time-frequency plane.

TEX Paper

Non-rectangular tilings can be designed. Baraniuk and Jones (1993) designed scale-shear fan bases corresponding to tiling with truncated wedges, well-matched to representing linear chirp signals with few expansion coefficients. Chirplet transforms (Mann and Haykin, 1992, 1995; Baraniuk and Jones, 1995, 1996) provide a systematic framework for designing new signal representations, and can be extended to deal with not just slant but curvature in the time-frequency plane, but a general approach to discrete chirplet transforms is not forthcoming. Gaussian chirplets do not form an orthogonal basis and adaptive matching pursuit Gaussian chirplet decompositions have thus been advocated (Yin *et al*, 2002). Philosophically related is research by Hlawatsch *et al* (1999) and Papandreou-Suppappola *et al* (2001, 2002) who explore time and frequency warping.

The fractional Fourier transform (FrFT), probably more widely used in optical research, is also closely related to wavelet and chirp transforms (Ozaktas *et al*, 1994). The Fourier transform of fractional order $a$ is defined in such a way that the standard Fourier transform is a special case with order $a = 1$ while successively transforming with fractional orders $a_1$ and $a_2$ is the same as transforming with order $a_1 + a_2$ (so that the standard Fourier transform corresponds to two successive half-transforms). Ozaktas *et al* (1994) pointed out that the effect of the FrFT is to rotate the Wigner distribution of a signal and showed that consequently linear chirp functions are simply the time domain representation of signals that appear as delta functions or harmonics in other fractional domains. Similarly, Capus *et al* (2000) noted that the FrFT is composed of a multiplication by a linear chirp in the time domain, followed by Fourier transformation, followed by multiplication by a linear chirp in the transformed domain, and finally a complex scaling. Mendlovic *et al* (1997), (see also Huang and Suter, 1996), suggested a 'fractional wavelet transform' algorithm which consists of an FrFT of order $a$ followed by continuous wavelet transformation; for reconstruction use the inverse wavelet transform followed by an FrFT of order $-a$. If the signal was dominated by a linear chirp, then the value of $a$ should match the slant in the time-frequency plane. In general the order $a$ is chosen by optimization and it is suggested that the high computational burden can be reduced by use of optical processing.

Other recent work includes Stevenson *et al* (2001) and Capus and Brown (2003).

Our approach can be viewed as a Fourier/wavelet packet transform hybrid. Instead of using undecimated wavelet packet coefficient sequences, it uses undecimated wavelet packet detail sequences, each of which can be viewed as arising by Fourier transforming the signal, multiplying it by the modulus squared of the Fourier transform of a band-pass filter, followed by inverse Fourier transformation. Our approach to generalizing this approach is to replace the standard Fourier transformation by the generalized Fourier transformation (GFT) of Detka and El-Jaroudi (1996). For a signal $x(t)$ the GFT is given by $X_G(f) = \int_{-\infty}^{\infty} x(t) \, \mathrm{e}^{-\mathrm{i}2\pi[ft+s_0(t)]} dt$, where $s_0(t)$ is a real-valued function depending on time only and specifies the evolutionary phase behaviour of the signal. Note this is the same as applying the standard Fourier transform to $x(t)\mathrm{e}^{-\mathrm{i}2\pi s_0(t)}$. The inverse GFT corresponds to applying the usual inverse Fourier transform to $X_G(f)$ followed by multiplication of the result by $\mathrm{e}^{\mathrm{i}2\pi s_0(t)}$ since

$$x(t) = \int_{-\infty}^{\infty} X_G(f)\mathrm{e}^{\mathrm{i}2\pi[ft+s_0(t)]} df = \mathrm{e}^{\mathrm{i}2\pi s_0(t)} \int_{-\infty}^{\infty} X_G(f)\mathrm{e}^{\mathrm{i}2\pi ft} df. \qquad (1.1)$$

Hence if $X_G(f) \equiv \delta(f - f_0)$ then $x(t) = \mathrm{e}^{\mathrm{i}2\pi[f_0 t + s_0(t)]}$, i.e., a signal with instantaneous frequency $\nu(t) = f_0 + s_0'(t)$ will be mapped to the point $f = f_0$, where $s_0'(t) = \frac{d}{dt}s_0(t)$. Hence if we wish to map a signal with for example a curved path in the time-frequency plane specified by $f_0 + s_0'(t)$ into a wavelet packet passband (rather than a single point), we need merely to specify a function $s(t)$ *approximating* $s_0(t)$ such that the demodulated function $x(t)\mathrm{e}^{-\mathrm{i}2\pi s(t)}$ will be mapped into the passband. Of course this result is very flexible in that $s_0'(t)$ is not constrained to be linear or even quadratic.

Many audio, acoustic and speech-type signals can be characterized as multi-component signals with time-frequency distribution consisting of several parallel, or similarly-oriented, linear or curved energy paths with low instantaneous bandwidth. Examples abound: Flandrin (1988) examined bat sonar, Cohen (1989 and 1992) considers seismic signals and whale sounds, Sucic and Boashash (2003) look at bird song, and Gribonval and Bacry (2003) consider musical recordings. Our approach is particularly well-suited to, and most easily implemented for, such signals, since with the demodulation the components should generally be shiftable into different wavelet packet frequency bands.

In order to track the time-dependent frequency content of each component of a multicomponent signal we shall first demodulate the complete signal, then project it onto the time-frequency plane in a manner such that (i) as closely as possible each component contributes exclusively to a different 'tile' in the chosen (wavelet packet) tiling of the time-frequency plane, and (ii) at each time instant the contribution to each tile definitely comes from no more than one component. The projected components are 'analytic' vectors. The single reverse demodulation is then applied to all projected components. The instantaneous frequency of each component in each tile is then well-defined and can be calculated and weighted by the energy to yield the 'Hilbert energy spectrum' for that projection. Agglomeration over projections yields the complete Hilbert spectrum (Huang et al, 1998; Olhede and Walden, 2004a).

## 2. The generalized demodulation algorithm

We give details of our algorithm using the more practically useful discrete-time notation. We assume we have sampled a continuous-time (multicomponent) signal $x(t)$ with a sample interval of unity to get an even-length vector of observations $\mathbf{x} = [x_0, \ldots, x_{N-1}]^T$ and have done likewise to $s(t)$ to get $\mathbf{s} = [s_0, \ldots, s_{N-1}]^T$. For a sample interval of unity the Nyquist frequency is $f = 1/2$.

### (*a*) *Analytic vectors and forward demodulation*

We firstly create an 'analytic' vector $\mathbf{y} = \mathbf{x} + i\mathcal{H}\mathbf{x}$ where $\mathcal{H}$ is the discrete Hilbert transform (DHT) matrix (Marple, 1999), and then apply the forward demodulation to obtain $d_l = y_l \mathrm{e}^{-\mathrm{i}2\pi s_l}$, $l = 0, \ldots, N-1$. We can write this in matrix form as $\mathbf{d} = (\mathcal{C} - \mathrm{i}\mathcal{S})\mathbf{y}$, where $(\mathcal{C})_{lm} = \cos(2\pi s_l)\delta_{lm}$, and $(\mathcal{S})_{lm} = \sin(2\pi s_l)\delta_{lm}$, where $\delta_{lm}$ is the Kronecker delta, and $m = 0, \ldots, N-1$.

The discrete Fourier transform of an analytic vector has zero coefficients for negative frequencies. Hence, by working with the analytic vector, any energy in $\mathbf{x}$ in the negative frequencies is not inadvertently moved into the positive frequencies

by the demodulation. The new complex-valued vector $\mathbf{d}$ is not analytic, and so we now create a new analytic vector $\mathbf{z} = \mathbf{d} + \mathrm{i}\mathcal{H}\mathbf{d}$.

We assume $\mathbf{x}$ to be band-limited to $\left[-\frac{1}{2} + \frac{1}{N}, -\frac{1}{N}\right] \bigcup \left[\frac{1}{N}, \frac{1}{2} - \frac{1}{N}\right]$ for discrete computation of the analytic vector $\mathbf{y}$. The effect of the demodulation is to change frequencies by $-s'_l$, where $s'_l$ denotes the discrete version of $s'(t)$. When the new analytic vector $\mathbf{z}$ is created, we must assume that $\mathbf{d}$ is band-limited to $\left[\frac{1}{N}, \frac{1}{2} - \frac{1}{N}\right]$, or, equivalently, that $\mathbf{y}$ is band-limited to $\left[\frac{1}{N} + s'_l, \frac{1}{2} - \frac{1}{N} + s'_l\right]$, for $l = 0, \dots, N-1$. This is the same as saying that the demodulation must not move signal component energy out of the frequency range $\left[\frac{1}{N}, \frac{1}{2} - \frac{1}{N}\right]$ and this is simple to check graphically, as shown by the examples of section 3.

### (b) Wavelet packet projections

The particular wavelet-type transform which we will apply to $\mathbf{z}$ is the so-called maximal overlap discrete wavelet packet transform (MODWPT) (Walden & Contreras Cristán, 1998). This is an undecimated transform. Each MODWPT is associated with a transform level $j$, $(j = 1, \dots, J_0)$, and the $j$th level decomposes the frequency interval $[0, 1/2]$ into $2^j$ equal pass-band intervals $\mathcal{I}_{j,n} = \left(\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$, where $n$, $(n = 0, \dots, 2^j - 1)$, denotes the frequency band index within level $j$. For any $j$ the $\mathcal{I}_{j,n} \times [0, (N-1)]$ partition (tile) the time-frequency plane into equal width rectangles. Of course a larger value of $j$ corresponds to a narrower pass-band, and our interest will be in the final level $J_0$ MODWPT which decomposes the frequency interval $[0, 1/2]$ into $2^{J_0}$ equal pass-band intervals; the choice of $J_0$ is considered in section 2(f).

For the MODWPT coefficients indexed by $(j, n)$ we write $\widetilde{\mathbf{W}}_{j,n} = \widetilde{\mathcal{W}}_{j,n}\mathbf{z}$ where $\widetilde{\mathbf{W}}_{j,n}$ is an $N$-length column vector, and $\widetilde{\mathcal{W}}_{j,n}$ is an $N \times N$ transformation matrix. The wavelet packet coefficients for $t = 0, \dots, N-1$ can be written as $\widetilde{W}_{j,n,t} = \sum_{l=0}^{L_j - 1} \tilde{u}_{j,n,l}z_{(t-l) \bmod N}$ where $\{\tilde{u}_{j,n,l}\}$ is a $j$th level and $n$th band MODWPT wavelet packet filter of length $L_j = (2^j - 1)(L - 1) + 1$, derived from the basic DWT scaling and wavelet filters $\{g_l, l = 0, \dots, L - 1\}$ and $\{h_l, l = 0, \dots, L - 1\}$ as described in Percival & Walden (2000, sec. 6.6). The tilde notation is used to distinguish the MODWPT filters and coefficients from the decimated (DWPT) equivalents, as in Percival & Walden (2000).

Let $\tilde{u}^\circ_{j,n,l}$ be the filter obtained by periodizing $\tilde{u}_{j,n,l}$ to length $N$, i.e., $\tilde{u}^\circ_{j,n,l} = \sum_{m=0}^{\infty} \tilde{u}_{j,n,l+mN}$. Then an equivalent representation is

$$\widetilde{W}_{j,n,t} = \sum_{l=0}^{N-1} \tilde{u}^\circ_{j,n,l}z_{(t-l) \bmod N}. \tag{2.1}$$

We will actually make use of sequences $\widetilde{\mathcal{D}}_{j,n} = \widetilde{\mathcal{W}}^T_{j,n}\widetilde{\mathcal{W}}_{j,n}\mathbf{z} = \widetilde{\mathcal{W}}^T_{j,n}\widetilde{\mathbf{W}}_{j,n}$ with $j = J_0, n = 1, \dots, 2^j - 1$. For $n > 0$ the sequence $\{\widetilde{\mathcal{D}}_{j,n,l}, l = 0, \dots, N - 1\}$ consists of level $j$, frequency band $n$, MODWPT *detail* coefficients. (The sequence $\{\widetilde{\mathcal{D}}_{j,0,l}, l = 0, \dots, N - 1\}$ involves only low-pass filters and covers low frequencies and trends).

For a level $j$ the analytic vector can be recovered via $\mathbf{z} = \sum_{n=0}^{2^j - 1} \widetilde{\mathcal{D}}_{j,n}$ so that if we define projection matrices $\mathbf{P}_{\mathcal{R}_{j,n}} = \widetilde{\mathcal{W}}^T_{j,n}\widetilde{\mathcal{W}}_{j,n}$ then (Olhede and Walden, 2004a)

$\mathbf{z} = \sum_{n=0}^{2^j-1} \mathbf{P}_{j,n}\mathbf{z} = \sum_{n=0}^{2^j-1} \widetilde{\mathcal{D}}_{j,n}$, i.e., we can decompose $\mathbf{z}$ into a linear combination of the contributions to $\mathbf{z}$ in projection sub-spaces which tile the time-frequency plane via a set of equal bandwidth rectangles. We note that at any time $l$, $l = 0, \dots, N-1$, energy in $\mathbf{x}$ in the frequency band $\left(\frac{n}{2^{j+1}} + s'_l, \frac{n+1}{2^{j+1}} + s'_l\right]$, is mapped to $\mathcal{I}_{j,n}$.

Any MODWPT detail sequence $\widetilde{\mathcal{D}}_{j,n}$ created as above is an analytic vector. To see this note that since $\mathbf{z} = \mathbf{d} + i\mathcal{H}\mathbf{d}$, we have that $\widetilde{\mathcal{D}}_{j,n} = \mathbf{P}_{j,n}\mathbf{z} = \mathbf{P}_{j,n}\mathbf{d} + i\mathbf{P}_{j,n}\mathcal{H}\mathbf{d}$. But the matrices $\mathbf{P}_{j,n} = \widetilde{\mathcal{W}}_{j,n}^T\widetilde{\mathcal{W}}_{j,n}$ and $\mathcal{H}$ commute (see Appendix A) so that $\widetilde{\mathcal{D}}_{j,n} = \mathbf{P}_{j,n}\mathbf{d} + i\mathcal{H}\mathbf{P}_{j,n}\mathbf{d}$, and hence $\widetilde{\mathcal{D}}_{j,n}$ is an analytic vector.

### (c) Reversed demodulation

Any details vector $\widetilde{\mathcal{D}}_{j,n}$ produced as above can be viewed as having been created by (discrete) GFT, followed by multiplication with the modulus squared of the Fourier transform of a band-pass filter, followed by inverse Fourier transformation. To turn this last step of inverse Fourier transformation into inverse GFT, we carry out a reverse demodulation (see (1.1)) via $\mathcal{Y}_{j,n,l} = \widetilde{\mathcal{D}}_{j,n,l}\, e^{i2\pi s_l}$, $l = 0, \dots, N-1$.

### (d) The Hilbert spectrum and instantaneous frequency

For any projection indexed by $(j, n)$ the Hilbert spectrum (Huang et al, 1998; Olhede and Walden, 2004a) can be formed by firstly calculating the amplitude and phase sequences given by

$$A_{j,n,l} = \sqrt{[\Re^2\{\mathcal{Y}_{j,n,l}\} + \Im^2\{\mathcal{Y}_{j,n,l}\}]} \text{ and } \phi_{j,n,l} = \tan^{-1}\left[\frac{\Im\{\mathcal{Y}_{j,n,l}\}}{\Re\{\mathcal{Y}_{j,n,l}\}}\right]. \qquad (2.2)$$

The derivative of the phase at discrete time $l$ may be calculated via a fourth-order generalized phase difference estimator (Boashash, 1992, p. 542) so that the instantaneous frequency is given by

$$\nu_{j,n,l} = \frac{1}{2\pi} \sum_{m=-2}^{2} r_m \phi_{j,n,l-m} \qquad (2.3)$$

with $r_{\mp 1} = \pm\frac{8}{12}, r_{\pm 2} = \pm\frac{1}{12}$.

Now, for a large $M$, define $M$ frequencies $f_k = k\Delta f$, $k = 0, \dots, M-1$, where $\Delta f = 1/[2(M-1)]$, covering the interval $[0, 1/2]$. Then the Hilbert energy spectrum of $\{\mathcal{Y}_{j,n,l}\}$ may be defined as

$$S_{j,n;l,f_k} = A_{j,n,l}^2 \delta_{k,\langle\nu_{j,n,l}/\Delta f\rangle}, \qquad l = 0, \dots, N-1; f_k = 0, \dots, 1/2, \qquad (2.4)$$

Here $\delta_{k,m}$ is the Kronecker delta, $\langle x \rangle$ denotes the integer closest to $x$; we used $M = 512$.

### (e) The algorithm summarized

1. Create the analytic vector $\mathbf{y} = \mathbf{x} + i\mathcal{H}\mathbf{x}$.

2. Forward demodulate using $\mathbf{d} = (\mathcal{C} - i\mathcal{S})\mathbf{y}$.

3. Create new analytic vector $\mathbf{z} = \mathbf{d} + i\mathcal{H}\mathbf{d}$.

4. Compute MODWPT projections $\widetilde{\mathcal{D}}_{J_0,n} = \mathbf{P}_{J_0,n}\mathbf{z}$,      $n = 1, \dots, 2^{J_0} - 1$.

5. Carry out reverse demodulation: $\mathcal{Y}_{J_0,n} = (\mathcal{C} + \mathrm{i}\mathcal{S})\widetilde{\mathcal{D}}_{J_0,n}$.

6. Compute the instantaneous frequency sequences $\{\nu_{J_0,n,l}\}$, $n = 1, \dots, 2^{J_0} - 1$ according to (2.3). Compute and plot the Hilbert energy spectrum as detailed in (2.4).

### (*f*)  *The choice of $J_0$*

In our algorithm we will compute MODWPT projections for level $j = J_0$, and so we need to choose $J_0$. Using, as in Olhede and Walden (2004a), the optimal asymptotic frequency resolution filters $\{g_l, l = 0, \dots, L - 1\}$, known as Fejér-Korovkin filters, the band-pass concentration of the MODWPT filters $\tilde{u}_{J_0,n,l}$ is very good, having the property that the only significant leakage from band $n$ is into band $n-1$ and/or $n+1$, but no further. This is true at least for $J_0 = 2, 3$ with $L = 18$ and $4$ with $L = 22$. (See figure 4 of Olhede and Walden (2004a) for $J_0 = 3$.) $J_0 = 2, 3$ and $4$ corresponds to dividing $[0, 1/2]$ into 4, 8 or 16 frequency bands. Our 'rule' is to choose $J_0$ as small as possible consistent with having at least one empty projection band between bands containing signal from different components in step 4 of the algorithm of section 2(e). In the examples which follow it is seen that this strategy can be achieved by taking $J_0 = 3$. Since the leakage properties hold at least to the 16 band case of $J_0 = 4$, the minimum frequency separation between components suitable for our algorithm is 1/16th of the Nyquist frequency.

## 3.  Examples

### (*a*)  *Application to real data*

Our first example is that of the echolocation pulse emitted by the Large Brown Bat, *Eptesicus Fuscus*, digitized with a sample interval of $\Delta t = 7\mu s$, with $N = 400$.

Fig. 1(a) shows the standard Wigner-Ville distribution for the analytic signal vector $\mathbf{y}$. The $p = 3$ significant components generate $p(p - 1)/2 = 3$ 'outer' inferences (i.e., interferences due to interaction of components — Flandrin (1999, p. 232)), and there is also visible 'inner' interference where the time-frequency energy distributions are noticeably nonlinear. Fig. 1(b) shows the ordinary Hilbert energy spectrum via MODWPT projections, *excluding any demodulation*, i.e., steps 2, 3 and 5 are omitted in the algorithm of section 2(e). Here $j = J_0 = 3$ so that $\mathcal{I}_{j,n} = \left(\frac{n}{2^{j+1}\Delta t}, \frac{n+1}{2^{j+1}\Delta t}\right] = \left(\frac{n}{16\Delta t}, \frac{n+1}{16\Delta t}\right]$, for $n = 0, \dots, 7$; these 8 frequency passbands are delineated by the horizontal dotted lines; the band between zero and the first horizontal dotted line corresponds to $n = 0$, with increasing $n$ until the top interval between the uppermost horizontal line and the top of the plot which covers the band $n = 7$. Although, as required in multicomponent analysis, *at each time instant* the different frequency chirps have been successfully separated into different frequency subbands, each component spans at least two subbands and is noticeably affected by the band-edge imperfections in the (albeit extremely good) band-pass filters generated using Fejér-Korovkin filter coefficients — see Olhede and Walden (2004a).
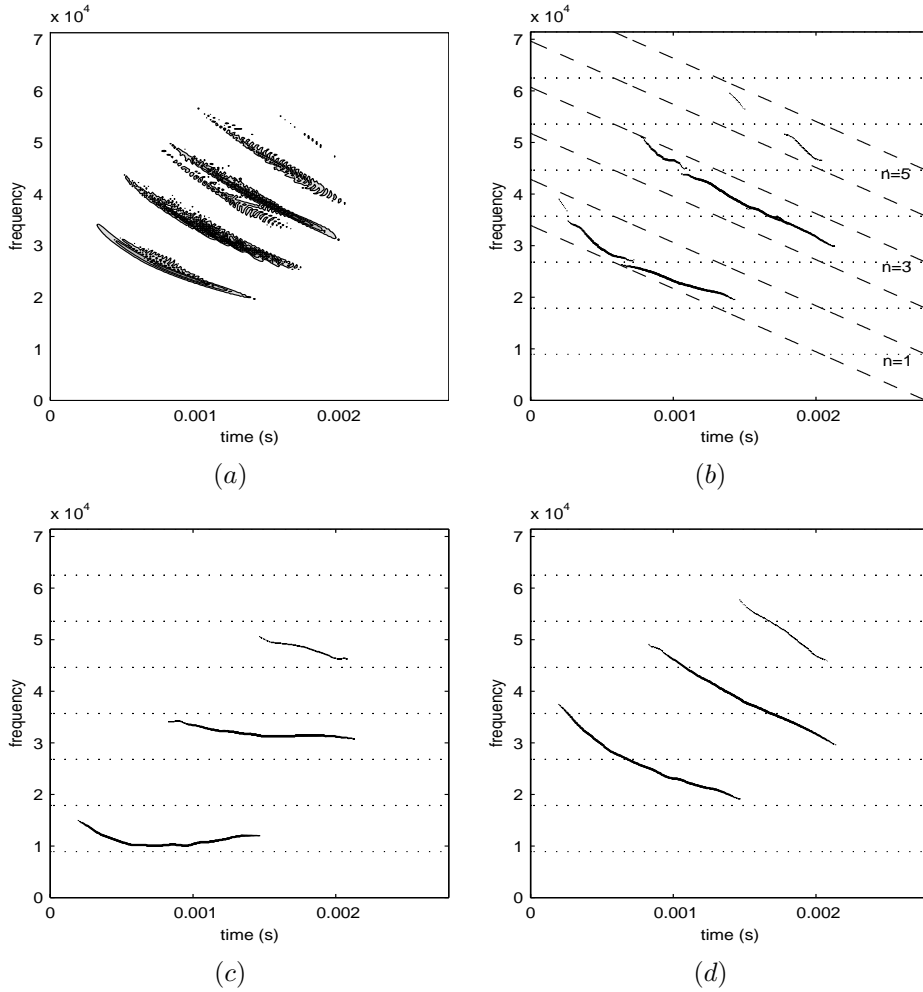
Figure 1. (a) Wigner-Ville distribution corresponding to analytic version **y** of the bat echolocation pulse. (b) The Hilbert spectrum using the MODWPT, 18-point Fejér-Korovkin wavelet filters, and $J_0 = 3$. (c) The Hilbert spectrum incorporating only forward demodulation. (d) The Hilbert spectrum using the complete algorithm. Details of marked lines etc are given in the text.

Turning now to demodulation, the form of $s_l$ used here was $s_l = b_2 l^2 + b_1 l$, $l = 0, \ldots, N - 1$, where $b_2 = -0.0003$ and $b_1 = 0.175$; hence $s'_l = (2b_2 l + b_1)/\Delta t$. The value of $b_2$ is easily 'guesstimated' from the Hilbert spectrum results of Fig. 1(b) by following the middle component, and $b_1$ can then be chosen to suitably align the signal components within the passbands as shown via the diagonal dashed lines in Fig. 1(b). The bottom dashed line shows $\frac{n}{2^{j+1}\Delta t} + s'_l$, $l = 0, \ldots, N - 1$, for $n = 1$, i.e., $\frac{1}{16\Delta t} + s'_l$, $l = 0, \ldots, N - 1$. The next to bottom line shows $\frac{n+1}{2^{j+1}\Delta t} + s'_l$, $l = 0, \ldots, N-1$, for $n = 1$, i.e., $\frac{2}{16\Delta t} + s'_l$, $l = 0, \ldots, N-1$. So the interval delineated by the two bottom lines is $\left(\frac{n}{2^{j+1}\Delta t} + s'_l, \frac{n+1}{2^{j+1}\Delta t} + s'_l\right]$, $l = 0, \ldots, N - 1$, for $n = 1$ and $j = J_0 = 3$, and as pointed out above, energy in such an interval will be mapped to $\mathcal{I}_{j,n} = \mathcal{I}_{3,1} = \left(\frac{1}{16\Delta t}, \frac{2}{16\Delta t}\right]$ following the demodulation in step 2 of the algorithm

of section 2(e). This interval is marked with '$n = 1$' on the right of the plot. The corresponding intervals for $n = 3$ and $n = 5$ are also marked on the right of Fig. 1(b) and, following the demodulation, energy in these intervals will be mapped to $\mathcal{I}_{3,3}$ and $\mathcal{I}_{3,5}$.

Fig. 1(c) shows the Hilbert spectrum via MODWPT projections *incorporating forward demodulation only*, i.e., only step 5 is omitted from the algorithm in section 2(e). We note as predicted from Fig. 1(b), the energy in the signal has been mapped to the pass-bands $\mathcal{I}_{3,1}, \mathcal{I}_{3,3}$ and $\mathcal{I}_{3,5}$. The demodulation has succeeded in realigning the time-frequency energy to (almost perfectly) lie within the frequency subbands. Fig. 1(d) shows the extremely good results obtained using the full algorithm *incorporating both the forward and reverse demodulation steps*. The curvature of the lower frequency component is well-preserved — there is no assumption of linear chirp components.

Our chosen demodulation has kept signal component energy in the frequency range $\left[ \frac{1}{N\Delta t}, \frac{1}{2\Delta t} - \frac{1}{N\Delta t} \right]$, as required — see section 2(a). We note that a poor choice of $s_l$ could cause energy loss by not satisfying this requirement, and/or result in one or more demodulated components crossing band boundaries and suffering band edge-effects as seen in the basic Hilbert spectrum without demodulation in Fig. 1(b).

### (b)  *Application to synthetic data*

Here we consider a synthetic signal which closely resembles a whale vocalisation in shape. It consists of two quadratic chirps, one having a shorter duration than the other:

$$x_l = \begin{cases} \sin(a_1 l^3 + a_2 l^2 + a_3 l), & l = 0, \dots, 249 \\ \sin(a_1 l^3 + a_2 l^2 + a_3 l) + \sin(a_4(l - 512)^3 + a_5 l), & l = 250, \dots, 1023 \end{cases}$$

with $a_1 = -1.2566 \ 10^{-6}, a_2 = 0.0019, a_3 = 1.1983, a_4 = -1.2783 \ 10^{-6}, a_5 = 1.3823$. Fig. 2(a) shows the standard Wigner-Ville distribution for the analytic signal vector $\mathbf{y}$. The two components generate a single case of outer inteference (between the components), while the notable curvature of the components generates severe inner interference. Fig. 2(b) shows the Hilbert spectrum via MODWPT projections, excluding any demodulation. Here $j = J_0 = 3$ again so that, with $\Delta t = 1$, $\mathcal{I}_{j,n} = \left( \frac{n}{16}, \frac{n+1}{16} \right]$, for $n = 0, \dots, 7$; these 8 frequency pass-bands are delineated by the horizontal dotted lines. Although, as with the bat signal, at any time the different frequency chirps have been successfully separated into different frequency subbands, here both components cross three subbands leading to band-edge imperfections. To eliminate these artifacts we can apply the demodulation method. To find $\{s_l\}$ we defined the quadratic $s'_l = 3b_3 l^2 + 2b_2 l + c$ and found the parameter values $b_3, b_2, c$ by matching of its values at $l = 0, 500$ and $1000$ to $f = 0.125, 0.28125$ and $0.14$, read off the plot by eye; this choice of points correspond roughly to an arc halfway between the two components. Then we let $s_l = b_3 l^3 + b_2 l^2 + b_1 l$ so that $s'_l = 3b_3 l^2 + 2b_2 l + b_1$, where $b_1$ is chosen to align the signal components sensibly within the passbands as shown via the curved dashed lines in Fig. 2(b). The values used were $b_1 = -8.5611 \ 10^{-2}, b_2 = 3.0592 \ 10^{-4}$ and $b_3 = -1.9874 \ 10^{-7}$. We have not used the known simulation parameters $a_1, \dots, a_5$ to obtain $s_l$, but worked entirely from the Hilbert spectrum of Fig. 2(b), as would be the case in practice.

Fig. 2(c) shows the Hilbert spectrum via MODWPT projections incorporating forward demodulation only. The energy in the signal has been mapped to the pass-bands $\mathcal{I}_{3,2}$ and $\mathcal{I}_{3,4}$. The time-frequency energy lies comfortably within the frequency subbands. Also shown by the dash-dot lines in Fig. 2(c) are the instantaneous frequencies calculated directly from the two separate components of the synthetic whale vocalization signal.

Our chosen demodulation has kept signal component energy in the frequency range $\left[\frac{1}{N}, \frac{1}{2} - \frac{1}{N}\right]$, as required.

Fig. 2(d) shows the final result using both forward and reverse demodulation steps. Most of the artifacts visible in Fig. 2(b) have been completely eliminated, and the result is excellent. Moreover the instantaneous frequencies are also plotted as dash-dot lines in this figure, but cannot be distinguished from the Hilbert spectrum calculated with our algorithm.

## 4. Standard deviation of the instantaneous frequency estimate in the presence of noise

In section 2(e) we gave the steps in the demodulation algorithm, ending with the calculation of instantaneous frequency (and thence the Hilbert spectrum). Hypothetically some bias could be introduced in the calculation of the instantaneous frequency of a component, through the steps of the deterministic algorithm of section 2(e). However, the whale vocalisation was shown to exhibit no significant bias, and so it is not inevitable. Minimization of wavelet packet filter leakage effects by choosing $J_0$ as described in section 2(f), and choosing $s_l'$ to align the signal components comfortably within (i.e., not on the edge of) the passbands, will always act to reduce bias.

We will now derive an expression for the variance of the *estimated* instantaneous frequency when the signal is observed in the presence of additive white noise. It does not refer to any formal uncertainty principle associated with our algorithm of section 2(e). Starting with white noise, we examine the effects of the various steps of the algorithm on the correlation structure of the real and imaginary parts of the resulting complex vectors.

### (a) General correlation structure

We can write the DHT matrix defined by the algorithm in Marple(1999) as the circulant

$$\mathcal{H} = \begin{bmatrix} 0 & q_{N-1}^{\circ} & 0 & q_{N-3}^{\circ} & 0 & \ldots & 0 & q_1^{\circ} \\ q_1^{\circ} & 0 & q_{N-1}^{\circ} & 0 & q_{N-3}^{\circ} & \ldots & q_3^{\circ} & 0 \\ 0 & q_1^{\circ} & 0 & q_{N-1}^{\circ} & 0 & \ldots & 0 & q_3^{\circ} \\ q_3^{\circ} & 0 & q_1^{\circ} & 0 & q_{N-1}^{\circ} & \ldots & q_5^{\circ} & 0 \\ 0 & q_3^{\circ} & 0 & q_1^{\circ} & 0 & & 0 & q_5^{\circ} \\ \vdots & & & & & \ddots & \vdots & \vdots \\ \vdots & & & & & & \ddots & \vdots \\ q_{N-1}^{\circ} & 0 & & & & & q_1^{\circ} & 0 \end{bmatrix} \quad (4.1)$$
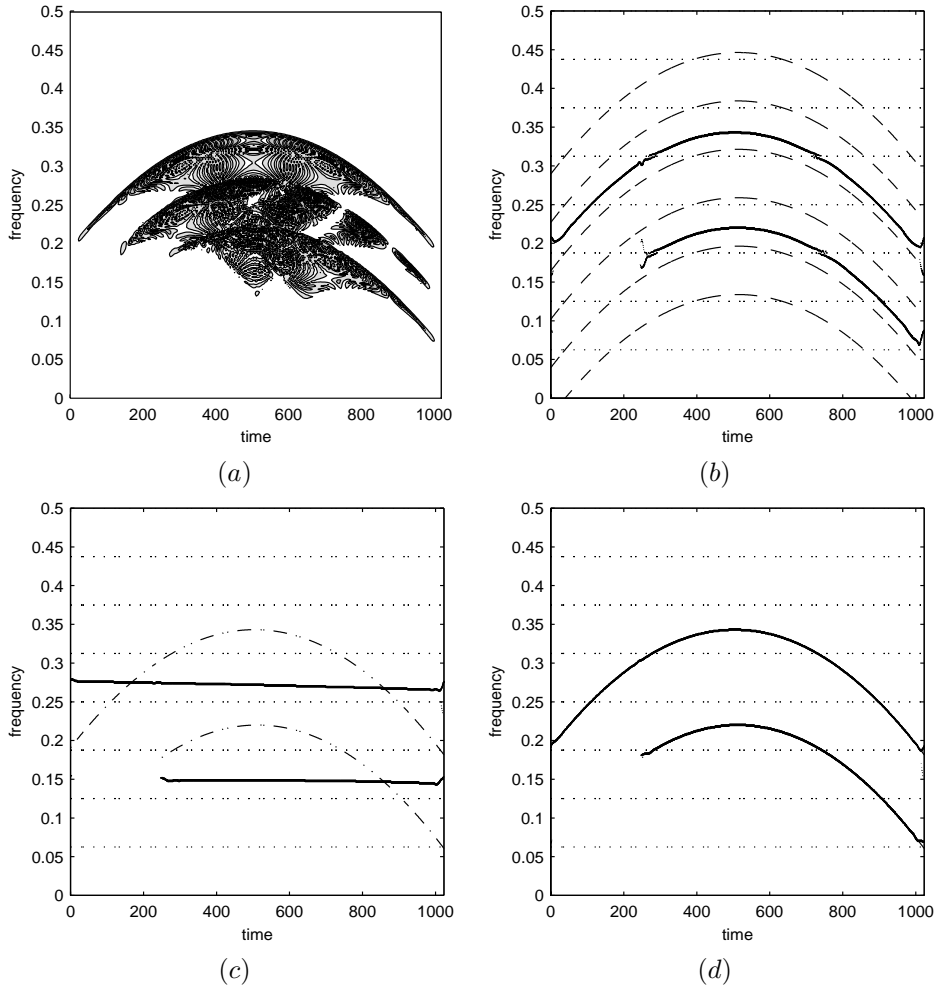
Figure 2. (a) Wigner-Ville distribution corresponding to analytic version **y** of the artificial whale vocalisation. (b) The Hilbert spectrum using the MODWPT, 18-point Fejér-Korovkin wavelet filters, and $J_0 = 3$. (c) The Hilbert spectrum incorporating only forward demodulation. (d) The Hilbert spectrum using the complete algorithm. Details of marked lines etc are given in the text.

where (e.g., Hahn, 1996, p. 143) the periodized filter $\{q_l^\circ\}$ is given by

$$q_l^\circ = \frac{2}{N} \sin^2[l\pi/2] \cot[l\pi/N] = \begin{cases} \frac{2}{N}\cot[l\pi/N], & l \text{ odd}, \\ 0, & l \text{ even}, \end{cases} \qquad (4.2)$$

for $l = 0, \ldots, N-1$. But

$$\{q_l^\circ : l = 0, \ldots, N-1\} \longleftrightarrow \{Q_k^\circ : k = 0, \ldots, N-1\},$$

where $\longleftrightarrow$ denotes 'Fourier transform pair' and $Q_k^\circ$ is given by

$$Q_k^\circ = \sum_{l=0}^{N-1} q_l^\circ e^{-\mathrm{i}2\pi kl/N} = \begin{cases} 0, & k = 0, \frac{N}{2}, \\ -\mathrm{i}, & k = 1, \dots, \frac{N}{2}-1, \\ \mathrm{i}, & k = \frac{N}{2}+1, \dots, N-1, \end{cases} \quad (4.3)$$

where $k/N$ denotes a Fourier frequency. The definition of $q_l$ means that $\mathcal{H}$ is, importantly, skew-symmetric, i.e., $\mathcal{H} = -\mathcal{H}^T$.

Setting $\mathbf{x} = \boldsymbol{\epsilon}$, an $N$-length vector of independent and identically distributed (IID) Gaussian noise (mean zero, variance $\sigma_\epsilon^2$), we firstly create the analytic vector $\mathbf{y} = \boldsymbol{\epsilon} + \mathrm{i}\mathcal{H}\boldsymbol{\epsilon} = \boldsymbol{\epsilon} + \mathrm{i}\,\breve{\boldsymbol{\epsilon}}$, say, where $\breve{\boldsymbol{\epsilon}} \equiv \mathcal{H}\boldsymbol{\epsilon}$. Then

$$\mathrm{var}\{\boldsymbol{\epsilon}\} = \sigma_\epsilon^2 \mathbf{I}, \quad \mathrm{var}\{\breve{\boldsymbol{\epsilon}}\} \approx \sigma_\epsilon^2 \mathbf{I}, \quad \mathrm{cov}\{\boldsymbol{\epsilon}, \breve{\boldsymbol{\epsilon}}\} \equiv E\{\boldsymbol{\epsilon}, \breve{\boldsymbol{\epsilon}}^T\} = \sigma_\epsilon^2 \mathcal{H}^T, \quad \mathrm{cov}\{\breve{\boldsymbol{\epsilon}}, \boldsymbol{\epsilon}\} = \sigma_\epsilon^2 \mathcal{H}.$$

The approximation for the second variance follows because of the transfer function structure in (4.3) — energy at Fourier frequencies 0 and 1/2 are removed by the DHT. (Since zero-mean white noise will have no energy at frequency zero, only the latter frequency will matter). In fact (Olhede and Walden, 2004b),

$$E\{\breve{\epsilon}_l \breve{\epsilon}_m\} = \begin{cases} \sigma_\epsilon^2 \frac{N-2}{N}, & \text{if } l = m, \\ -\sigma_\epsilon^2 \frac{[1+(-1)^{l-m}]}{N}, & \text{otherwise.} \end{cases} \quad (4.4)$$

So provided $N$ is large, then for practical purposes we can replace the approximation by an equality, and this is done hereafter. Other useful results are

$$\begin{aligned} \mathrm{cov}\{\epsilon_l, \breve{\epsilon}_m\} &= \sigma_\epsilon^2 (\mathcal{H}^T)_{lm} = -\sigma_\epsilon^2 (\mathcal{H}^T)_{ml} = -\mathrm{cov}\{\epsilon_m, \breve{\epsilon}_l\} \\ \text{and } \mathrm{cov}\{\epsilon_l, \breve{\epsilon}_l\} &= \sigma_\epsilon^2 (\mathcal{H}^T)_{ll} = 0. \end{aligned}$$

Step 2 of the algorithm in section 2(e) consists of the demodulation $d_l = y_l e^{-\mathrm{i}2\pi s_l}$, $l = 0, \dots, N-1$. Since $\mathcal{C}$ and $\mathcal{S}$ are both diagonal matrices they are symmetric (so that, e.g., $\mathcal{C} = \mathcal{C}^T$) and commute with each other. Also $\mathcal{C}$ will commute with any matrix which has all diagonal entries of zero, and similarly for $\mathcal{S}$; hence both $\mathcal{C}$ and $\mathcal{S}$ will commute with $\mathcal{H}$. Then with $\mathbf{y} = \boldsymbol{\epsilon} + \mathrm{i}\,\breve{\boldsymbol{\epsilon}}$, we obtain

$$\mathbf{d} = (\mathcal{C} - \mathrm{i}\mathcal{S})(\boldsymbol{\epsilon} + \mathrm{i}\,\breve{\boldsymbol{\epsilon}}) = (\mathcal{C}\boldsymbol{\epsilon} + \mathcal{S}\breve{\boldsymbol{\epsilon}}) + \mathrm{i}(\mathcal{C}\breve{\boldsymbol{\epsilon}} - \mathcal{S}\boldsymbol{\epsilon}).$$

If we now write $\mathbf{d} = \Re(\mathbf{d}) + \mathrm{i}\Im(\mathbf{d})$ we get

$$\begin{aligned} \mathrm{var}\{\Re(\mathbf{d})\} &= E\{(\mathcal{C}\boldsymbol{\epsilon} + \mathcal{S}\breve{\boldsymbol{\epsilon}})(\mathcal{C}\boldsymbol{\epsilon} + \mathcal{S}\breve{\boldsymbol{\epsilon}})^T\} \\ &= E\{\mathcal{C}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\mathcal{C}^T + \mathcal{C}\boldsymbol{\epsilon}\breve{\boldsymbol{\epsilon}}^T\mathcal{S}^T + \mathcal{S}\breve{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}^T\mathcal{C}^T + \mathcal{S}\breve{\boldsymbol{\epsilon}}\breve{\boldsymbol{\epsilon}}^T\mathcal{S}^T\} \\ &= \sigma_\epsilon^2[\mathcal{C}\mathcal{C}^T + \mathcal{C}\mathcal{H}^T\mathcal{S}^T + \mathcal{S}\mathcal{H}\mathcal{C}^T + \mathcal{S}\mathcal{S}^T] \\ &= \sigma_\epsilon^2[\mathcal{C}\mathcal{C}^T + \mathcal{S}\mathcal{S}^T] + \sigma_\epsilon^2[\mathcal{C}(\mathcal{H}^T + \mathcal{H})\mathcal{S}^T], \end{aligned}$$

where we have made use of the symmetry and commutative properties of $\mathcal{C}$ and $\mathcal{S}$. Now $\mathcal{C}\mathcal{C}^T + \mathcal{S}\mathcal{S}^T = \mathbf{I}$ and $\mathcal{H}^T + \mathcal{H} = \mathbf{0}$ since $\mathcal{H}$ is skew-symmetric, so we see that $\mathrm{var}\{\Re(\mathbf{d})\} = \sigma_\epsilon^2 \mathbf{I}$. In a similar way we obtain $\mathrm{var}\{\Im(\mathbf{d})\} = \sigma_\epsilon^2 \mathbf{I}$ and

$$\begin{aligned} \mathrm{cov}\{\Re(\mathbf{d}), \Im(\mathbf{d})\} &= E\{(\mathcal{C}\boldsymbol{\epsilon} + \mathcal{S}\breve{\boldsymbol{\epsilon}})(\mathcal{C}\breve{\boldsymbol{\epsilon}} - \mathcal{S}\boldsymbol{\epsilon})^T\} \\ &= E\{\mathcal{C}\boldsymbol{\epsilon}\breve{\boldsymbol{\epsilon}}^T\mathcal{C}^T - \mathcal{C}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\mathcal{S}^T + \mathcal{S}\breve{\boldsymbol{\epsilon}}\breve{\boldsymbol{\epsilon}}^T\mathcal{C}^T - \mathcal{S}\breve{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}^T\mathcal{S}^T\} \\ &= \sigma_\epsilon^2[\mathcal{C}\mathcal{H}^T\mathcal{C}^T - \mathcal{C}\mathcal{S}^T + \mathcal{S}\mathcal{C}^T - \mathcal{S}\mathcal{H}\mathcal{S}^T] \\ &= \sigma_\epsilon^2[\mathcal{C}\mathcal{H}^T\mathcal{C}^T + \mathcal{S}\mathcal{H}^T\mathcal{S}^T] = \sigma_\epsilon^2\mathcal{H}^T[\mathcal{C}\mathcal{C}^T + \mathcal{S}\mathcal{S}^T] = \sigma_\epsilon^2\mathcal{H}^T. \end{aligned}$$

So we have shown the covariance structure is unaffected by the demodulation, as would be expected.

Step 3 of the algorithm in section 2(e) creates a new analytic vector $\mathbf{z} = \mathbf{d} + i\mathcal{H}\mathbf{d} = \mathbf{d} + i\breve{\mathbf{d}}$, say, where $\breve{\mathbf{d}} \equiv \mathcal{H}\mathbf{d}$. We might expect the real and imaginary parts of $\mathbf{z}$ to have the same covariance properties as the real and imaginary parts of $\mathbf{y}$, since the demodulation has not changed these second-order properties. There is an important difference however: $\mathbf{z}$ is the analytic version of a complex-valued vector, while $\mathbf{y}$ is the analytic version of a real-valued vector. Now $\mathbf{z} = \mathbf{d} + i\mathcal{H}\mathbf{d} = \Re(\mathbf{d}) - \mathcal{H}\Im(\mathbf{d}) + i[\Im(\mathbf{d}) + \mathcal{H}\Re(\mathbf{d})]$. So

$$
\begin{aligned}
\mathrm{cov}\{\Re(\mathbf{z}), \Im(\mathbf{z})\} &= E\{(\Re(\mathbf{d}) - \mathcal{H}\Im(\mathbf{d}))(\Im(\mathbf{d}) + \mathcal{H}\Re(\mathbf{d}))^T\} \\
&= E\{\Re(\mathbf{d})\Im(\mathbf{d})^T + \Re(\mathbf{d})\Re(\mathbf{d})^T\mathcal{H}^T \\
&\quad -\mathcal{H}\Im(\mathbf{d})\Im(\mathbf{d})^T - \mathcal{H}\Im(\mathbf{d})\Re(\mathbf{d})^T\mathcal{H}^T\} \\
&= \sigma_\epsilon^2[\mathcal{H}^T + \mathcal{H}^T - \mathcal{H} - \mathcal{H}(\mathcal{H}\mathcal{H}^T)] \\
&= \sigma_\epsilon^2\mathcal{H}^T[3\mathbf{I} + (\mathcal{H}\mathcal{H}^T)].
\end{aligned}
$$

The final line follows from the skew-symmetry of $\mathcal{H}$. Now, as shown in Appendix B, $\mathcal{H}\mathcal{H}^T$ is given by the symmetric Toeplitz matrix with first row

$$
[1 - \frac{2}{N}, 0, -\frac{2}{N}, 0, -\frac{2}{N}, \dots, -\frac{2}{N}, 0] \tag{4.5}
$$

and as $N \to \infty$ we see that $\mathcal{H}\mathcal{H}^T \to \mathbf{I}$. Hence, for large $N$,

$$
\mathrm{cov}\{\Re(\mathbf{z}), \Im(\mathbf{z})\} = 4\sigma_\epsilon^2\mathcal{H}^T, \tag{4.6}
$$

and similarly, for large $N$,

$$
\mathrm{var}\{\Re(\mathbf{z})\} = 4\sigma_\epsilon^2\mathbf{I} \quad \text{and} \quad \mathrm{var}\{\Im(\mathbf{z})\} = 4\sigma_\epsilon^2\mathbf{I}. \tag{4.7}
$$

Importantly, we see that, apart from the factor of 4, the variances and covariances of the real and imaginary parts of $\mathbf{z}$ are the same as for the real and imaginary parts of $\mathbf{y}$.

In step 4 of the algorithm we compute MODWPT projections $\widetilde{\mathcal{D}}_{j,n} = \mathbf{P}_{j,n}\mathbf{z}$. Since $\mathbf{P}_{j,n}$ is real-valued,

$$
\widetilde{\mathcal{D}}_{j,n} = \mathbf{P}_{j,n}\mathbf{z} = \mathbf{P}_{j,n}\Re(\mathbf{z}) + i\mathbf{P}_{j,n}\Im(\mathbf{z}) = \Re(\widetilde{\mathcal{D}}_{j,n}) + i\Im(\widetilde{\mathcal{D}}_{j,n}),
$$

say. Then

$$
\mathrm{var}\{\Re(\widetilde{\mathcal{D}}_{j,n})\} = 4\sigma_\epsilon^2\mathbf{P}_{j,n}\mathbf{P}_{j,n}^T \quad \text{and} \quad \mathrm{var}\{\Im(\widetilde{\mathcal{D}}_{j,n})\} = 4\sigma_\epsilon^2\mathbf{P}_{j,n}\mathbf{P}_{j,n}^T,
$$

and

$$
\mathrm{cov}\{\Re(\widetilde{\mathcal{D}}_{j,n}), \Im(\widetilde{\mathcal{D}}_{j,n})\} = 4\sigma_\epsilon^2\mathbf{P}_{j,n}\mathcal{H}^T\mathbf{P}_{j,n}^T.
$$

It was shown in Olhede and Walden (2004b) that $\mathbf{x}^T\mathcal{H}^T\mathbf{x} = 0$ for any real-valued vector $\mathbf{x}$ so that $(\mathrm{cov}\{\Re(\widetilde{\mathcal{D}}_{j,n}), \Im(\widetilde{\mathcal{D}}_{j,n})\})_{ll} = 0$, $l = 1, \dots, N$, i.e., all diagonal entries of the covariance matrix are zero.

A simple approximate expression for the elements of $\mathrm{var}\{\Re(\widetilde{\mathcal{D}}_{j,n})\}$ is derived in Appendix C, equation (C 3), with the same expression for $\mathrm{var}\{\Im(\widetilde{\mathcal{D}}_{j,n})\}$. Likewise

in Appendix C, equation (C 2), a simple approximate expression for the elements of $\mathrm{cov}\{\Re(\widetilde{\mathcal{D}}_{j,n}), \Im(\widetilde{\mathcal{D}}_{j,n})\}$ is derived.

Step 5 of the algorithm consists of the reverse demodulation: $\mathcal{Y}_{j,n} = (\mathcal{C}+\mathrm{i}\mathcal{S})\widetilde{\mathcal{D}}_{j,n}$, which, as we showed for the forward step, will not alter the covariance structure. Hence if we write $\mathcal{Y}_{j,n} = \Re\{\mathcal{Y}_{j,n}\} + \mathrm{i}\Im\{\mathcal{Y}_{j,n}\}$, then the variances and covariances of $\Re\{\mathcal{Y}_{j,n}\}$ and $\Im\{\mathcal{Y}_{j,n}\}$ are the same as for $\Re(\widetilde{\mathcal{D}}_{j,n})$ and $\Im(\widetilde{\mathcal{D}}_{j,n})$, above.

Hence, starting with $\mathbf{x} = \boldsymbol{\epsilon}$, a zero-mean IID Gaussian noise vector, we have found the (approximate) covariance structure of the complex-valued random vector $\mathcal{Y}_{j,n}$ resulting from the five steps of the algorithm.

### (b) *Instantaneous frequency*

Since all steps of the algorithm up to and including reverse demodulation consist of linear matrix operations, we know that if our original input signal $\mathbf{x}$ consists of both deterministic and zero-mean white noise components, then the result of the algorithm at step 5 is given by $\mathcal{Y}_{j,n} = \mathcal{Y}_{j,n}^{(\mathbf{D})} + \mathcal{Y}_{j,n}^{(\boldsymbol{\epsilon})}$, where henceforth the superscripts $\mathbf{D}$ and $\boldsymbol{\epsilon}$ indicate the deterministic and noise components, respectively, and $\mathcal{Y}_{j,n}^{(\boldsymbol{\epsilon})}$ has a mean of zero. We can express a component of the deterministic part $\mathcal{Y}_{j,n}^{(\mathbf{D})}$ in terms of its amplitude and phase:

$$\mathcal{Y}_{j,n,l}^{(\mathbf{D})} = A_{j,n,l}^{(\mathbf{D})} \mathrm{e}^{\mathrm{i}\phi_{j,n,l}^{(\mathbf{D})}}. \tag{4.8}$$

But from (2.2) the $l$th component of the estimated instantaneous phase vector at level $j$ frequency band index $n$ is

$$\hat{\phi}_{j,n,l} = \tan^{-1}\left[\frac{\Im\{\mathcal{Y}_{j,n,l}\}}{\Re\{\mathcal{Y}_{j,n,l}\}}\right] = \tan^{-1}\left[\frac{\Im\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\} + \Im\{\mathcal{Y}_{j,n,l}^{(\boldsymbol{\epsilon})}\}}{\Re\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\} + \Re\{\mathcal{Y}_{j,n,l}^{(\boldsymbol{\epsilon})}\}}\right],$$

and a first-order Taylor expansion about $(\Re\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}, \Im\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\})$ gives

$$\hat{\phi}_{j,n,l} \approx \tan^{-1}\left[\frac{\Im\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}}{\Re\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}}\right] + \left[\frac{\Re\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}\Im\{\mathcal{Y}_{j,n,l}^{(\boldsymbol{\epsilon})}\} - \Im\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}\Re\{\mathcal{Y}_{j,n,l}^{(\boldsymbol{\epsilon})}\}}{\Re^2\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\} + \Im^2\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}}\right]. \tag{4.9}$$

From (4.8) we know $[A_{j,n,l}^{(\mathbf{D})}]^2 = \Re^2\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\} + \Im^2\{\mathcal{Y}_{j,n,l}^{(\mathbf{D})}\}$, then since the variances and covariances of $\Re\{\mathcal{Y}_{j,n}^{(\boldsymbol{\epsilon})}\}$ and $\Im\{\mathcal{Y}_{j,n}^{(\boldsymbol{\epsilon})}\}$ are the same as for $\Re(\widetilde{\mathcal{D}}_{j,n}^{(\boldsymbol{\epsilon})})$ and $\Im(\widetilde{\mathcal{D}}_{j,n}^{(\boldsymbol{\epsilon})})$, we use (C 2) and (C 3) to see that the variance of the *estimated* phase — since there is now additive noise — is

$$\mathrm{var}\{\hat{\phi}_{j,n,l}\} \approx \frac{4\sigma_\epsilon^2}{2^j[A_{j,n,l}^{(\mathbf{D})}]^2}.$$

The estimated instantaneous frequency $\hat{\nu}_{j,n,l}$ is got by substituting the estimated phase into (2.3). Hence its variance takes the form

$$\begin{aligned}
\mathrm{var}\{\hat{\nu}_{j,n,l}\} &= \frac{1}{4\pi^2}\mathrm{cov}\left\{\sum_{m=-2}^{2} r_m\hat{\phi}_{j,n,l-m}, \sum_{k=-2}^{2} r_k\hat{\phi}_{j,n,l-k}\right\} \\
&= \frac{1}{4\pi^2}\sum_{m=-2}^{2}\sum_{k=-2}^{2} r_m r_k\, \mathrm{cov}\left\{\hat{\phi}_{j,n,l-m}, \hat{\phi}_{j,n,l-k}\right\}. \tag{4.10}
\end{aligned}$$

Now, using (4.9), $\text{cov}\left\{\hat{\phi}_{j,n,l-m}, \hat{\phi}_{j,n,l-k}\right\}$ can be expressed as

$$E\left\{\left[\Re\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Im\{\mathcal{Y}_{j,n,l-m}^{(\epsilon)}\} - \Im\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Re\{\mathcal{Y}_{j,n,l-m}^{(\epsilon)}\}\right] \times \right.$$
$$\left.\left[\Re\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\}\Im\{\mathcal{Y}_{j,n,l-k}^{(\epsilon)}\} - \Im\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\}\Re\{\mathcal{Y}_{j,n,l-k}^{(\epsilon)}\}\right]\right\} \Big/ \left\{[A_{j,n,l-m}^{(\mathbf{D})}]^2[A_{j,n,l-k}^{(\mathbf{D})}]^2\right\}.$$

So $\text{cov}\left\{\hat{\phi}_{j,n,l-m}, \hat{\phi}_{j,n,l-k}\right\}$ is approximately

$$\left\{\Re\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Re\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\}\Xi_{j,n,m-k} - \Re\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Im\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\}\Upsilon_{j,n,m-k}\right.$$
$$\left.-\Im\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Re\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\}\Upsilon_{j,n,k-m} + \Im\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Im\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\}\Xi_{j,n,m-k}\right\} \Big/$$
$$\left\{[A_{j,n,l-m}^{(\mathbf{D})}]^2[A_{j,n,l-k}^{(\mathbf{D})}]^2\right\},$$

where the symmetric autocovariance sequence, $\{\Xi_{j,n,\tau}\}$, and approximate skew-symmetric cross-covariance sequence, $\{\Upsilon_{j,n,\tau}\}$, are defined in (C 3) and (C 2). But

$$\Re\{\mathcal{Y}_{j,n,l-m}^{(\mathbf{D})}\}\Re\{\mathcal{Y}_{j,n,l-k}^{(\mathbf{D})}\} = A_{j,n,l-m}^{(\mathbf{D})}A_{j,n,l-k}^{(\mathbf{D})}\cos(\phi_{j,n,l-m}^{(\mathbf{D})})\cos(\phi_{j,n,l-k}^{(\mathbf{D})}),$$

and similarly for the other terms. Our approximate form for $\text{cov}\left\{\hat{\phi}_{j,n,l-m}, \hat{\phi}_{j,n,l-k}\right\}$ becomes

$$\left[\left\{\cos(\phi_{j,n,l-m}^{(\mathbf{D})})\cos(\phi_{j,n,l-k}^{(\mathbf{D})}) + \sin(\phi_{j,n,l-m}^{(\mathbf{D})})\sin(\phi_{j,n,l-k}^{(\mathbf{D})})\right\}\Xi_{j,n,m-k}\right.$$
$$+\left.\left\{\sin(\phi_{j,n,l-m}^{(\mathbf{D})})\cos(\phi_{j,n,l-k}^{(\mathbf{D})}) - \cos(\phi_{j,n,l-m}^{(\mathbf{D})})\sin(\phi_{j,n,l-k}^{(\mathbf{D})})\right\}\Upsilon_{j,n,m-k}\right] \Big/$$
$$\left\{[A_{j,n,l-m}^{(\mathbf{D})}]^2[A_{j,n,l-k}^{(\mathbf{D})}]^2\right\}$$
$$= \left[\{\cos(\phi_{j,n,l-m}^{(\mathbf{D})} - \phi_{j,n,l-k}^{(\mathbf{D})})\}\Xi_{j,n,m-k} + \{\sin(\phi_{j,n,l-m}^{(\mathbf{D})} - \phi_{j,n,l-k}^{(\mathbf{D})})\}\Upsilon_{j,n,m-k}\right] \Big/$$
$$\left\{[A_{j,n,l-m}^{(\mathbf{D})}]^2[A_{j,n,l-k}^{(\mathbf{D})}]^2\right\}.$$

So, finally, our approximation for $\text{var}\{\hat{\nu}_{j,n,l}\}$ is given by

$$\sum_{m,k=-2}^{2} r_m r_k \frac{\left[\{\cos(\phi_{j,n,l-m}^{(\mathbf{D})} - \phi_{j,n,l-k}^{(\mathbf{D})})\}\Xi_{j,n,m-k} + \{\sin(\phi_{j,n,l-m}^{(\mathbf{D})} - \phi_{j,n,l-k}^{(\mathbf{D})})\}\Upsilon_{j,n,m-k}\right]}{4\pi^2 A_{j,n,l-m}^{(\mathbf{D})}A_{j,n,l-k}^{(\mathbf{D})}}. \quad (4.11)$$

When calculating this expression with real data, the sequences of amplitudes, $\{A_{j,n,l}^{(\mathbf{D})}\}$, and phases, $\{\phi_{j,n,l}^{(\mathbf{D})}\}$, are unknown, but these can be replaced by their estimates.

In order to show the validity of the above variance expression, consider the linear chirp $x_l = \sin(a_1 l^2 + a_2 l)$, $l = 0, \ldots, 1023$, with $a_1 = 0.02\pi/1024$ and $a_2 = 0.36\pi$. This has an instantaneous frequency increasing from 0.18 to 0.2 over the 1024 time points. White Gaussian noise was added to achieve a signal-to-noise variance ratio of 10. A level $J_0 = 2$ MODWPT was used; the signal lies in band $n = 1$, frequency band [0.125, 0.25]. $\text{var}\{\hat{\nu}_{2,1,l}\}$ was calculated using (4.11) and the known amplitudes $\{A_{2,1,l}^{(\mathbf{D})}\}$ and phases $\{\phi_{2,1,l}^{(\mathbf{D})}\}$ and is shown in Figure 3(a). Figure 3(b) shows the mean

standard deviation obtained at each time from 200 independent repeat simulations, again using (4.11), but with the amplitudes $\{A^{(\mathbf{D})}_{2,1,l}\}$ and phases $\{\phi^{(\mathbf{D})}_{2,1,l}\}$ estimated from the noisy data of each simulation. Finally Figure 3(c) compares the result shown in Figure 3(b) with the the empirical standard deviation computed from the 200 independent repeat simulations of the estimation of instantaneous frequency at each time. It can be seen that there is good agreement between all three cases, supporting the validity of expression (4.11).

When calculating (4.11) with real data, not only must the sequences of amplitudes, $\{A^{(\mathbf{D})}_{j,n,l}\}$, and phases, $\{\phi^{(\mathbf{D})}_{j,n,l}\}$, be estimated, but so too must the variance of the white noise, $\sigma^2_\epsilon$, occurring in the formula for $\{\Xi_{j,n,\tau}\}$ in (C 3) and in the formula for $\{\Upsilon_{j,n,\tau}\}$ in (C 2). This variance can be estimated from the MODWPT details in an appropriate signal-free level and band. For example for Figure 3(b) we used $j = J_0 = 2$ and $n = 3$, which corresponds to the highest frequency band. When incorporating the demodulation and reverse demodulation steps we must proceed with caution: energy moved out of the positive frequency region of time-frequency into the negative region (via wrap-around) by the demodulation in step 2 of the algorithm in section 2(e) will be eliminated in the creation of the new analytic vector in step 3. We assume of course that the demodulation is chosen so that signal is not eliminated in this way, but some noise will inevitably be, so that when estimating $\sigma^2_\epsilon$ in this case, care should be taken not to choose MODWPT details in a signal-free region where the noise has been affected by the demodulation.

## 5. Summary

We have described a flexible approach for the time-frequency analysis of multi-component signals. It involves the use of analytic vectors and demodulation. The demodulated analytic signal is projected onto the time-frequency plane so that, as closely as possible, each component contributes exclusively to a different 'tile' in a wavelet packet tiling of the time-frequency plane, and at any time the contribution to each tile definitely comes from no more than one component. A single reverse demodulation is then applied to all projected components. Two examples of how to choose the form of the demodulation sequence $s_l$ were presented in some detail within the context of the real and simulated data examples. Following the above procedures, the instantaneous frequency of each component in each tile is well-defined and may be calculated, followed by the Hilbert energy spectrum.

In order to better understand the effect of additive noise, the approximate variance of the estimated instantaneous frequency in any tile has been formulated by starting with pure noise and studying its evolving covariance structure through each step of the algorithm. The validity and practical utility of the resulting expression was demonstrated by comparing the theoretical standard deviation with all parameters taken as known, with its mean over 200 simulations with unknown quantities estimated from the noisy data, and with the empirical standard deviation computed from 200 simulations of the estimation of instantaneous frequency.

The site `http://www.ma.imperial.ac.uk/statistics/research/wavelets/` contains MATLAB code for our algorithm.
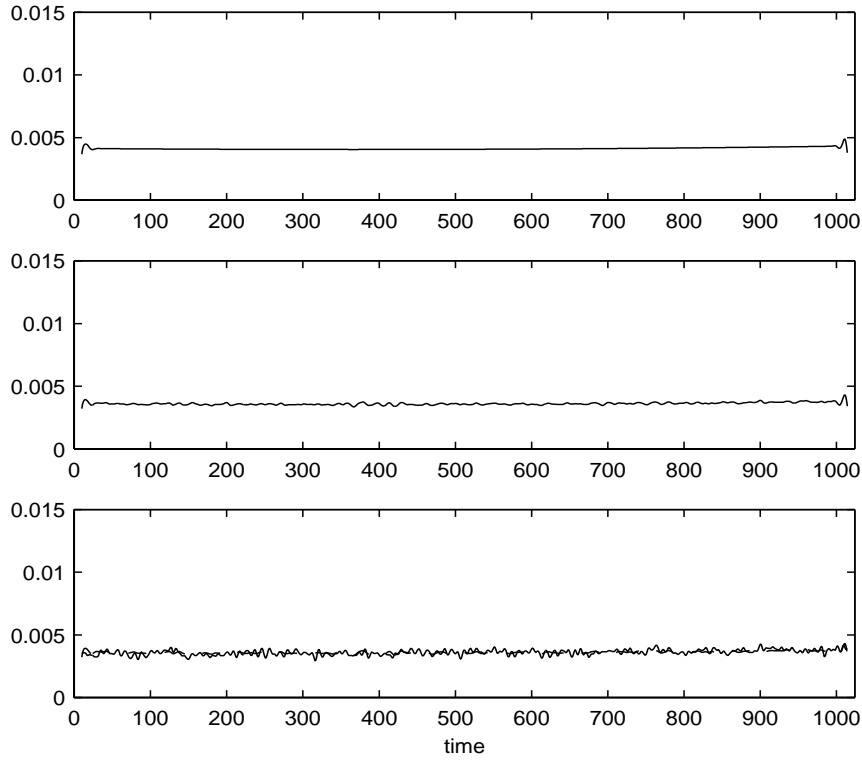
Figure 3. (a) Theoretical standard deviation from (4.11), of estimated instantaneous frequency. (b) The mean standard deviation from (4.11) from 200 independent repeat simulations using quantities estimated from the noisy data. (c) Comparison of (b) (dashed) with the empirical standard deviation computed from 200 independent repeat simulations of the estimation of instantaneous frequency. Calculations used the MODWPT with $J_0 = 2$ and 18-point Fejér-Korovkin wavelet filters.

## Appendix A.

We wish to show that the matrices $\mathbf{P}_{j,n} = \widetilde{\mathcal{W}}_{j,n}^T \widetilde{\mathcal{W}}_{j,n}$ and $\mathcal{H}$ commute,

From (2.1) we can write $\widetilde{\mathcal{W}}_{j,n}$ as

$$
\widetilde{\mathcal{W}}_{j,n} =
\begin{bmatrix}
\tilde{u}_{j,n,0}^{\circ} & \tilde{u}_{j,n,N-1}^{\circ} & \tilde{u}_{j,n,N-2}^{\circ} & \cdots & \tilde{u}_{j,n,3}^{\circ} & \tilde{u}_{j,n,2}^{\circ} & \tilde{u}_{j,n,1}^{\circ} \\
\tilde{u}_{j,n,1}^{\circ} & \tilde{u}_{j,n,0}^{\circ} & \tilde{u}_{j,n,N-1}^{\circ} & \cdots & \tilde{u}_{j,n,4}^{\circ} & \tilde{u}_{j,n,3}^{\circ} & \tilde{u}_{j,n,2}^{\circ} \\
\tilde{u}_{j,n,2}^{\circ} & \tilde{u}_{j,n,1}^{\circ} & \tilde{u}_{j,n,0}^{\circ} & \cdots & \tilde{u}_{j,n,5}^{\circ} & \tilde{u}_{j,n,4}^{\circ} & \tilde{u}_{j,n,3}^{\circ} \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
\tilde{u}_{j,n,N-2}^{\circ} & \tilde{u}_{j,n,N-3}^{\circ} & \tilde{u}_{j,n,N-4}^{\circ} & \cdots & \tilde{u}_{j,n,1}^{\circ} & \tilde{u}_{j,n,0}^{\circ} & \tilde{u}_{j,n,N-1}^{\circ} \\
\tilde{u}_{j,n,N-1}^{\circ} & \tilde{u}_{j,n,N-2}^{\circ} & \tilde{u}_{j,n,N-3}^{\circ} & \cdots & \tilde{u}_{j,n,2}^{\circ} & \tilde{u}_{j,n,1}^{\circ} & \tilde{u}_{j,n,0}^{\circ}
\end{bmatrix} ; \quad (A\,1)
$$

see e.g., Percival and Walden (2000, p. 171).

To show that the matrices $\mathbf{P}_{j,n} = \widetilde{\mathcal{W}}_{j,n}^T \widetilde{\mathcal{W}}_{j,n}$ and $\mathcal{H}$ commute, we firstly define the circular shift (orthonormal) transform matrix $\mathcal{T}$. Then with $\mathbf{x} = [x_0, \ldots, x_{N-1}]^T$ we have $\mathcal{T}\mathbf{x} = [x_{N-1}, x_0, \ldots, x_{N-2}]^T$. Now $\mathcal{T}^2 \equiv \mathcal{T}\mathcal{T}$ etc, and we can write $\mathcal{H}$ in (4.1) as $\mathcal{H} = \sum_{m=0}^{N-1} q_m^\circ \mathcal{T}^m$ where $q_m^\circ = 0$ for $m$ even. Similarly $\widetilde{\mathcal{W}}_j$ in (A 1) can be written as $\widetilde{\mathcal{W}}_j = \sum_{l=0}^{N-1} \tilde{u}_{j,n,l}^\circ \mathcal{T}^l$ and $\widetilde{\mathcal{W}}_j^T = \sum_{k=0}^{N-1} \tilde{u}_{j,n,k}^\circ (\mathcal{T}^k)^T = \sum_{k=0}^{N-1} \tilde{u}_{j,n,k}^\circ (\mathcal{T}^T)^k = \sum_{k=0}^{N-1} \tilde{u}_{j,n,k}^\circ \mathcal{T}^{-k}$, since $\mathcal{T}^T = \mathcal{T}^{-1}$ because $\mathcal{T}$ is orthonormal. Hence

$$
\begin{aligned}
\mathcal{H}\widetilde{\mathcal{W}}_{j,n}^T \widetilde{\mathcal{W}}_{j,n} &= \sum_{m=0}^{N-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} q_m^\circ \, \tilde{u}_{j,n,k}^\circ \, \tilde{u}_{j,n,l}^\circ \mathcal{T}^{m-k+l} \\
&= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} \tilde{u}_{j,n,k}^\circ \, \tilde{u}_{j,n,l}^\circ \, q_m^\circ \mathcal{T}^{-k+l+m} = \widetilde{\mathcal{W}}_{j,n}^T \widetilde{\mathcal{W}}_{j,n}\mathcal{H},
\end{aligned}
$$

as required.

## Appendix B.

To show that $\mathcal{H}\mathcal{H}^T$ has the form stated in (4.5) it is sufficient to consider the dot product of the first row of $\mathcal{H}$ with circularly shifted versions of other rows. Now $q^\circ \star q_l^\circ \equiv \sum_{n=0}^{N-1} q_n^\circ q_{(n+l)\,\mathrm{mod}\,N}^\circ$ and $\{q^\circ \star q_l^\circ\} \longleftrightarrow \{|Q_k^\circ|^2 \equiv \mathcal{Q}_k^\circ\}$. The inverse Fourier transform relationship yields $q^\circ \star q_l^\circ = \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{Q}_k^\circ e^{\mathrm{i}2\pi lk/N}$. But from (4.3),

$$
\mathcal{Q}_k^\circ = \begin{cases} 0, & k = 0, \frac{N}{2}, \\ 1, & k = 1, \ldots, \frac{N}{2} - 1, \text{ and } k = \frac{N}{2} + 1, \ldots, N - 1. \end{cases}
$$

We know that

$$
\frac{1}{N} \sum_{k=0}^{N-1} e^{\mathrm{i}2\pi lk/N} = \begin{cases} 1, & l = 0, \\ 0, & l = 1, \ldots, N - 1, \end{cases}
$$

so that

$$
\sum_{n=0}^{N-1} q_n^\circ q_{(n+l)\,\mathrm{mod}\,N}^\circ + \frac{1}{N}(1 + [-1]^l) = \begin{cases} 1, & l = 0, \\ 0, & l = 1, \ldots, N - 1, \end{cases}
$$

i.e.,

$$
\sum_{n=0}^{N-1} q_n^\circ q_{(n+l)\,\mathrm{mod}\,N}^\circ = \begin{cases} 1 - \frac{2}{N}, & l = 0, \\ 0, & l = 1, 3, 5, \ldots, N - 1, \\ -\frac{2}{N}, & l = 2, 4, 6, \ldots, N - 2. \end{cases}
$$

The form of $\mathcal{H}\mathcal{H}^T$ in (4.5) then follows immediately.

## Appendix C.

Here we derive simple approximate expressions for the elements of $\mathrm{var}\{\Re(\widetilde{\mathcal{D}}_{j,n})\}$, $\mathrm{var}\{\Im(\widetilde{\mathcal{D}}_{j,n})\}$, and $\mathrm{cov}\{\Re(\widetilde{\mathcal{D}}_{j,n}), \Im(\widetilde{\mathcal{D}}_{j,n})\}$. Since $\mathbf{P}_{j,n}$ is real-valued we know that,

$$
\widetilde{\mathcal{D}}_{j,n} = \mathbf{P}_{j,n}\mathbf{z} = \mathbf{P}_{j,n}\Re(\mathbf{z}) + i\mathbf{P}_{j,n}\Im(\mathbf{z}) = \Re(\widetilde{\mathcal{D}}_{j,n}) + i\Im(\widetilde{\mathcal{D}}_{j,n}).
$$

We shall make use of the shorthand $\mathbf{a}_{j,n} \equiv \Re(\widetilde{\mathcal{D}}_{j,n})$; $\mathbf{b}_{j,n} \equiv \Im(\widetilde{\mathcal{D}}_{j,n})$. The variances and covariances of $\Re(\mathbf{z})$ and $\Im(\mathbf{z})$ are specified in (4.7) and (4.6). Now consider $\text{cov}\{a_{j,n,l}, b_{j,n,m}\}$. Let $\mathbf{r}^T$ be the row vector of $\mathbf{P}_{j,n}$ that yields $a_{j,n,l}$ in $\mathbf{a}_{j,n} = \mathbf{P}_{j,n}\Re(\mathbf{z})$ and let $\mathbf{v}^T$ be the row vector of $\mathbf{P}_{j,n}$ that yields $b_{j,n,m}$ in $\mathbf{b}_{j,n} = \mathbf{P}_{j,n}\Im(\mathbf{z})$. Now $\text{cov}\{\Re(\mathbf{z}), \Im(\mathbf{z})\} = 4\sigma_\epsilon^2 \mathcal{H}^T$ — see equation (4.6) — and since $\mathcal{H}^T$ is a circulant matrix, $\mathcal{H}^T = \mathcal{F}^H \boldsymbol{\Lambda} \mathcal{F}$, (e.g., Percival and Walden, 2000, p. 441). The $(l,m)$th element of the orthonormal DFT (ODFT) matrix $\mathcal{F}$ is given by $\exp(-\mathrm{i}2\pi lm/N)/\sqrt{N}$, for $0 \le l, m \le N-1$, 'H' denotes Hermitian transpose, and $\boldsymbol{\Lambda}$ is a diagonal matrix with diagonal elements $\{Q_k^\circ : k = 0, \ldots, N-1\}$, the DFT of the first row of $\mathcal{H}^T$. Then

$$\text{cov}\{a_{j,n,l}, b_{j,n,m}\} = 4\sigma_\epsilon^2 \mathbf{r}^T \mathcal{H}^T \mathbf{v} = 4\sigma_\epsilon^2 \mathbf{r}^T \mathcal{F}^H \boldsymbol{\Lambda} \mathcal{F} \mathbf{v} = 4\sigma_\epsilon^2 (\mathcal{F}\mathbf{r})^H \boldsymbol{\Lambda} \mathcal{F} \mathbf{v}.$$

Now $\mathcal{F}\mathbf{r}$ is the ODFT of $\mathbf{r}$ which is related to the DFT, say $\{R_k\}$ of $\mathbf{r}$ via $\mathcal{F}\mathbf{r} = (1/\sqrt{N})[R_0, \ldots, R_{N-1}]^T$, and similarly for $\mathcal{F}\mathbf{v}$. Hence,

$$\text{cov}\{a_{j,n,l}, b_{j,n,m}\} = \frac{4\sigma_\epsilon^2}{N} [R_0^*, R_1^*, \ldots, R_{N-1}^*] \begin{bmatrix} Q_0^\circ & 0 & \cdots & 0 \\ 0 & Q_1^\circ & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_{N-1}^\circ \end{bmatrix} \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_{N-1} \end{bmatrix}$$

$$= \frac{4\sigma_\epsilon^2}{N} \sum_{k=0}^{N-1} R_k^* Q_k^\circ V_k.$$

With $U_{j,n,k}^\circ = \sum_{l=0}^{N-1} \tilde{u}_{j,n,l}^\circ e^{-\mathrm{i}2\pi lk/N}$, we have $R_k^* = |U_{j,n,k}^\circ|^2 e^{\mathrm{i}2\pi kl/N}$ and $V_k = |U_{j,n,k}^\circ|^2 e^{-\mathrm{i}2\pi km/N}$, so that

$$\text{cov}\{a_{j,n,l}, b_{j,n,m}\} = \frac{4\sigma_\epsilon^2}{N} \sum_{k=0}^{N-1} |U_{j,n,k}^\circ|^4 Q_k^\circ e^{\mathrm{i}2\pi k(l-m)/N}. \tag{C 1}$$

But $U_{j,n,k}^\circ = U_{j,n}(\frac{k}{N})$ where $U_{j,n}(f) = \sum_{l=0}^{L_j-1} \tilde{u}_{j,n,l} e^{-\mathrm{i}2\pi lf}$, where we recall that $\{\tilde{u}_{j,n,l}\}$ is a $j$th level and $n$th band MODWPT wavelet packet filter of length $L_j = (2^j - 1)(L - 1) + 1$ while $\tilde{u}_{j,n,l}^\circ$ is its periodized version. Also $Q_k^\circ = Q(\frac{k}{N})$ (Percival and Walden, 2000, p. 506) where $Q(f) = \sum_{l=-\infty}^{\infty} q_l e^{-\mathrm{i}2\pi lf}$ is the periodic function

$$Q(f) = -\mathrm{i}\,\text{sgn}_p(f) = \begin{cases} 0, & \text{if } f = 0, 1/2; \\ -\mathrm{i}, & \text{if } 0 < f < 1/2; \\ \mathrm{i}, & \text{if } -1/2 < f < 0. \end{cases}$$

and $\{q_l\}$ is the non-periodized Hilbert transform filter given by

$$q_l = \int_{-1/2}^{1/2} -\mathrm{i}\,\text{sgn}_p(f) e^{\mathrm{i}2\pi fl} df = \begin{cases} \frac{2\sin^2(l\pi/2)}{l\pi}, & \text{if } l \text{ odd}; \\ 0, & \text{if } l \text{ even}, \end{cases}$$

so that $q_l^\circ = \sum_{n=-\infty}^{\infty} q_{l+nN}$. It then follows from (C 1) that $\text{cov}\{a_{j,n,l}, b_{j,n,m}\} \approx 4\sigma_\epsilon^2 \int_{-1/2}^{1/2} |U_{j,n}(f)|^4 Q(f) e^{\mathrm{i}2\pi f(l-m)} df$. But,

$$|U_{j,n}(f)|^2 \approx \begin{cases} 1, & \text{if } f \in \left\{ -\left[\frac{n+1}{2^{j+1}}, \frac{n}{2^{j+1}}\right) \cup \left(\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right] \right\} \\ 0, & \text{otherwise}, \end{cases}$$

so that

$$
\begin{aligned}
\operatorname{cov}\{a_{j,n,l}, b_{j,n,m}\} &\approx \frac{4\sigma_\epsilon^2}{2\pi(l-m)}\left[\mathrm{e}^{\mathrm{i}2\pi f(l-m)}\Big|_{-(n+1)/2^{j+1}}^{-n/2^{j+1}} - \mathrm{e}^{\mathrm{i}2\pi f(l-m)}\Big|_{n/2^{j+1}}^{(n+1)/2^{j+1}}\right] \\
&= \begin{cases} 0, & l = m, \\ 4\sigma_\epsilon^2\frac{\{\cos[n\pi(l-m)/2^j]-\cos[(n+1)\pi(l-m)/2^j]\}}{[\pi(l-m)]}, & l \neq m, \end{cases} \quad \text{(C\,2)} \\
&\equiv \Upsilon_{j,n,l-m}.
\end{aligned}
$$

This is skew-symmetric in $l$ and $m$.

We can determine $\operatorname{cov}\{a_{j,n,l}, a_{j,n,m}\}$ (or $\operatorname{cov}\{b_{j,n,l}, b_{j,n,m}\}$) by noting from (4.7) that the appropriate circulant matrix in this case is simply $\mathbf{I}$; the DFT of its first row is unity, so that

$$
\begin{aligned}
\operatorname{cov}\{a_{j,n,l}, a_{j,n,m}\} &\approx 4\sigma_\epsilon^2\int_{-1/2}^{1/2}|U_{j,n}(f)|^4 e^{\mathrm{i}2\pi f(l-m)}df \\
&= \begin{cases} 4\sigma_\epsilon^2/2^j, & l = m, \\ 4\sigma_\epsilon^2\frac{\{\sin[(n+1)\pi(l-m)/2^j]-\sin[n\pi(l-m)/2^j]\}}{[\pi(l-m)]}, & l \neq m, \end{cases} \text{(C\,3)} \\
&\equiv \Xi_{j,n,l-m}.
\end{aligned}
$$

with an identical result for $\operatorname{cov}\{b_{j,n,l}, b_{j,n,m}\}$; note this is symmetric in $l$ and $m$.

# References

Baraniuk, R. G. & Jones, D. L. 1993 Shear madness: new orthogonal bases and frames using chirp functions. *IEEE Trans. Signal Process.* **41**, 3543–3549.

Baraniuk, R. G. & Jones, D. L. 1995 Unitary equivalence: a new twist on signal processing. *IEEE Trans. Signal Process.* **43**, 2269–2282.

Baraniuk, R. G. & Jones, D. L. 1996 Wigner-based formulation of the chirplet transform. *IEEE Trans. Signal Process.* **44**, 3129–3135.

Boashash, B. 1992 Estimating and interpreting the instantaneous frequency of a signal-Part 2: algorithms and applications. *Proc. IEEE* **80**, 540–568.

Capus, C., Rzhanov, Y. & Linnett, L. 2000 The analysis of multiple linear chirp signals. In *Time-Scale and Time-Frequency Analysis and Applications*, London: IEE, 4/1–4/7.

Capus, C. & Brown, K. 2003 Short-time fractional Fourier methods for the time-frequency representation of chirp signals. *J. Accoust. Soc. Am.* **113**, 3253–3263.

Cohen, L. 1989, Time-frequency distributions - a review. *Proc. IEEE*, **77**, 941–981.

Cohen, L. 1995, *Time-frequency analysis.* Upper Saddle River, New Jersey: Prentice-Hall.

Cohen, L. 1992 What is a multicomponent signal? In *IEEE Int. Conf. Acoust., Speech and Signal Processing*, San Francisco, **5**, 113–116.

Detka C. S. & El-Jaroudi, A. 1996 The generalised evolutionary spectrum. *IEEE Trans. Signal Process.* **44**, 2877–2881.

Flandrin, P. 1988 Time-frequency processing of bat sonar signals. In *Animal Sonar – Process and Performance* (eds. P. Nachtigall and P. Moore), pp. 797–802. Plenum.

Flandrin, P. 1999 *Time-frequency/time-scale analysis.* San Diego: Academic Press.

Gribonval, R. & Bacry, E. 2003 Harmonic decomposition of audio signals with matching pursuit. *IEEE Trans. Signal Process.* **51**, 101–111.

Hahn, S. L. 1996 *Hilbert transforms in signal processing.* Boston: Artech House.

Hlawatsch, F., Papandreou-Suppappola, A. & Boudreaux-Bartels, G. F. 1999 The power classes - quadratic time-frequency representations with scale covariance and dispersive time-shift covariance. *IEEE Trans. Signal Process.* **47**, 3067–3083.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N-C., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **454**, 903–95.

Huang, Y. & Suter, B. 1996 Fractional wavelet packet transform. In *Proc. IEEE Digital Signal Processing Workshop*, Loen, Norway, 413-415.

Mann, S. & Haykin, S. 1992 Time-frequency perspectives: the "Chirplet" transform. In *IEEE Int. Conf. Acoust., Speech and Signal Processing*, San Francisco, **3**, 417–420.

Mann, S. & Haykin, S. 1995 The chirplet transform: physical considerations. *IEEE Trans. Signal Process.* **43**, 2745–2761.

Marple, S. L. Jr. 1999 Computing the discrete-time "analytic" signal via FFT. *IEEE Trans. Signal Process.* **47**, 2600–2603.

Mendlovic, D., Zalevsky, Z., Mas, D., Garcia, J. & Ferreira, C. 1997 Fractional wavelet transform. *Applied Optics* **36**, 4801–4805.

Olhede, S.C. & Walden, A.T. 2004a The Hilbert spectrum via wavelet projections. *Proc. R. Soc. A* **460**, 955–975.

Olhede, S.C. & Walden, A.T. 2004b 'Analytic' wavelet thresholding. *Biometrika*, to appear in December.

Ozaktas, H. M., Barshan, B., Mendlovic, D. & Onural, L. 1994 Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms. *J. Opt. Soc. Amer. A.* **11**, 547–559.

Papandreou-Suppappola, A., Murray, R. L., Iem, B.-G. & Boudreaux-Bartels, G. F. 2001 Group delay shift covariant quadratic time-frequency representations. *IEEE Trans. Signal Process.* **49**, 2549–2564.

Papandreou-Suppappola, A. & Suppappola, S. B. 2002 Analysis and classification of time-varying signals with multiple time-frequency structures. *IEEE Signal Process. Letters* **9**, 92–95.

Percival, D. B. & Walden, A. T. 2000 *Wavelet methods for time series analysis.* Cambridge: Cambridge University Press.

Sucic, V. & Boashash, B. 2002 Selecting the optimal time-frequency distribution for real-life multicomponent signals under given constraints. In *Proc. XI Eur. Sig. Proc. Conf., Toulouse*, vol. 1, 141–144.

Stevenson, I. R., Nicholson, P., Linnett, L. M. & Morrison, S. 2001 A method for the analysis of chirp signals insonifying layered media for sub-bottom profiling. In *OCEANS 2001, MTS/IEEE Conference and Exhibition, 5-8 Nov. 2001,* **4**, 2608–2615.

Walden, A. T. & Contreras Cristán, A. 1998 The phase-corrected undecimated discrete wavelet packet transform and its application to interpreting the timing of events. *Proc. R. Soc. Lond. A* **454**, 2243–2266.

Yin, Q, Qian, S. & Feng, A. 2002 A fast refinement for adaptive Gaussian chirplet decomposition. *IEEE Trans. Signal Process.* **50**, 1298–1306.