The Hilbert spectrum via wavelet projections

BY S. OLHEDE AND A.T. WALDEN

Department of Mathematics, Imperial College of Science, Technology and Medicine, 180 Queen's Gate, London SW7 2BZ, UK.

Non-stationary signals are increasingly analysed in the time-frequency domain to determine the variation of frequency components with time. It was recently proposed in this journal that such signals could be analysed by projections onto the time-frequency plane giving a set of monocomponent signals. These could then be converted to 'analytic' signals using the Hilbert transform and their instantaneous frequency calculated, which when weighted by the energy yields the 'Hilbert energy spectrum' for that projection. Agglomeration over projections yields the complete Hilbert spectrum. We show that superior results can be obtained using wavelet-based projections. The maximal-overlap (undecimated/stationary/translation invariant) discrete wavelet transform and wavelet packet transforms are used, with the Fejér-Korovkin class of wavelet filters. These transforms produce decompositions which are conducive to statistical analysis, in particular enabling noise reduction methodology to be developed and easily and successfully applied.

Keywords: denoising; empirical mode decomposition; Hilbert spectrum; instantaneous frequency; wavelet packets; wavelet transform

1. Introduction

Most naturally occurring or man-made signals display non-stationary behaviour, i.e., some structural aspect such as their spectrum, statistical correlation or transfer function vary with time. Classical analysis techniques which assume stationarity are unsuitable, and investigators are drawn to using 'time-frequency' methods (e.g., Cohen, 1995) to analyse such signals.

Consider a generally complex-valued signal $z(t)$ which can be written in terms of a time-domain amplitude $A(t)$ and phase $\phi(t)$ as $z(t) = A(t)e^{i\phi(t)}$. Suppose its energy is highly concentrated along some trajectory in the time-frequency plane. A general representation would then be in terms of a a curve defined by the average frequency at each time — the so-called 'instantaneous frequency.' Boashash (1992a) remarks that for monocomponent signals which obey certain conditions the instantaneous frequency will have a clear physical meaning.

However, multicomponent signals are frequently encountered. Loughlin and Tacer (1997) consider the two-component signal $z(t) = A_1 e^{i2\pi f_1 t} + A_2 e^{i2\pi f_2 t}$, where A_1 and A_2 are real constants. The instantaneous frequency of the signal is

$$\nu(t) = \frac{\phi'(t)}{2\pi} = \frac{(f_1 + f_2)}{2} + \frac{(f_2 - f_1)(A_2^2 - A_1^2)}{2\{A_2^2 + A_1^2 + 2A_1 A_2 \cos[2\pi(f_2 - f_1)t]\}}. \tag{1.1}$$

TeX Paper

This is not bounded by f_1 or f_2 unless $|A_1| = |A_2|$, in which case the instantaneous frequency is consistent with the interpretation as the 'average frequency at each time' of this multicomponent signal. They also consider more general two-component signals and find in general that the instantaneous frequency of the signal exhibits large fluctuations and extends beyond the band defined by the instantaneous frequency of the components. Instantaneous frequency is generally a poor measure for multicomponent signals.

For this reason the analysis of such multicomponent signals is best done using projections onto the time-frequency plane corresponding to monocomponent signals which individually exhibit coherent frequency behaviour at any one time.

Recently in this journal a time-frequency energy spectrum estimator was proposed by Huang *et al* (1998) which projects the signal in the time-frequency plane (using the 'empirical mode decomposition (EMD)') and then calculates the instantaneous frequency for each projection. Each projection is intended to ensure coherent frequency behaviour at any time, or a monocomponent signal, so that the instantaneous frequency is well-behaved. Finally for each projection the appropriate energy is associated with the instantaneous frequency at each time, and amalgamation over projections gives the overall 'Hilbert energy spectrum.' The EMD projections are data-dependent and *ad-hoc* constructions, which vary with each signal — this is potentially a strength (adaptivity) and a weakness (resistance to statistical analysis).

The approach of Huang *et al* (1998) is aimed at analysis of both non-stationary and non-linear proceses. Our interest is directed at the analysis of non-stationary processes for which we find the EMD method produces oscillatory or poorly-defined Hilbert spectra, often with evidence of notable mode mixing. We show that superior results can be obtained by replacing the EMD projections by wavelet-based ones. In particular we utilise the maximal-overlap (undecimated/stationary/translation invariant) discrete wavelet transform and wavelet packet transforms, denoted by MODWT and MODWPT, respectively. These non-adaptive transforms produce decompositions which are conducive to statistical analysis, enabling noise reduction methodology to be developed and utilised.

In §2 and §3 we give a brief summary of the EMD, and the Hilbert spectrum, respectively. The background to wavelet-based time-frequency projections via so-called MODWT 'detail' coefficients is described in §4 along with the computation of the Hilbert spectrum. Projections via the wavelet packet transform (MODWPT), a generalization of the MODWT, are developed in §5, where details are given on choosing a suitable set of wavelet packets. The importance of low time-frequency leakage for the projections is also addressed in §5, where it is shown that the Fejér-Korovkin class of wavelet filters are particularly appropriate. Examples showing that the wavelet-based approach gives excellent results, and outperforms EMD, are given in §6. Sums of intermittent and complete sinusoids are analysed, as well as a bat echolocation pulse. The theory behind Hilbert spectrum estimation in the presence of noise, via thresholding of wavelet details, is given in §7, followed by an example showing its efficacy.

Throughout the paper a t in parentheses indicates continuous time, and a subscript t indicates discrete time.

2. The Empirical Mode Decomposition

(a) Projections in time-frequency: continuous time

Given a real-valued continuous-time signal $x(t)$, consider a series of projection operators $\mathcal{P}_{\mathcal{R}_j}$ which project $x(t)$ into a time-frequency region $\mathcal{R}_j$. These operators need not be orthogonal but we assume that $\sum_{j=1}^{J_0+1}(\mathcal{P}_{\mathcal{R}_j}x)(t) = x(t)$ for some J_0. For $j = 1, \ldots, J_0$ define $d_j(t) = (\mathcal{P}_{\mathcal{R}_j}x)(t)$ and $s_{J_0}(t) = (\mathcal{P}_{\mathcal{R}_{J_0+1}}x)(t)$. Then

$$x(t) = \sum_{j=1}^{J_0} d_j(t) + s_{J_0}(t). \tag{2.1}$$

Such an additive decomposition of $x(t)$ is achieved by the empirical mode decomposition method.

(b) The EMD algorithm

Huang *et al* (1998, 1999) define an *empirical mode decomposition* (EMD) of a signal into *intrinsic mode functions* (IMFs). An IMF has the same number of zero crossings and extrema, and symmetric envelopes defined by the local maxima and minima. Each such IMF should give rise to a well-defined instantaneous frequency. The EMD of $x(t)$ proceeds as follows. Upper, $u_1(t)$, and lower, $l_1(t)$, envelopes of $x(t)$ are found by interpolation between extrema. The local mean function is defined as $m_1(t) = \frac{1}{2}[u_1(t) + l_1(t)]$. Ideally $h_{1,1}(t) = x(t) - m_1(t)$, should be an IMF, but invariably a *sifting* refinement is required: $h_{1,1}(t)$ replaces $x(t)$ and the above step is repeated to give $m_2(t)$ and $h_{1,2}(t)$. After $k-1$ such repeats $h_{1,k}(t)$ is obtained and a stopping criterion equates it to the first IMF, $d_1(t)$ say.

Various stopping criteria exist. Huang *et al* (1998, p. 920) test whether the accumulated relative squared differences between $h_{1,k}(t)$ and $h_{1,k-1}(t)$ is less than 0.3; see also Magrin-Chagnolleau and Baraniuk (1999). Huang *et al* (1999, p. 425) look for an equal number of extrema and zero crossings. A more detailed combined criterion is used by Rilling *et al* (2003) and this is the scheme used in this paper; for details see also `www.ens-lyon.fr/~flandrin/software.html`. We have also repeated our analyses using the other stopping criteria, and the basic conclusions of the paper are unchanged.

Denoting the sum of the mean functions by $m_{\text{tot}}(t)$ we note that $d_1(t) = x(t) - m_{\text{tot}}(t)$. The second IMF is found by performing the above procedure on the residual $s_1(t) = x(t) - d_1(t)$, and so on, until $s_{J_0}(t)$ satisfies a criterion indicating that no more IMFs can be extracted (no more than one extrema). $s_{J_0}(t)$ will be a trend or constant. Then (2.1) is obtained.

Locally, each IMF contains lower frequency oscillations than that previously extracted. The EMD gives an adaptive frequency subband decomposition of the signal. If the time-frequency plane has been 'tiled' ideally, each projection will contain at most one monocomponent signal.

3. The Hilbert spectrum and instantaneous frequency

Given a real-valued $x(t)$, and hence real-valued $d_j(t)$, we can create the analytic signal $d_j^a(t) = d_j(t) + \mathrm{i}\mathcal{H}\{d_j(t)\} = d_j(t) + \mathrm{i}d_j^{(H)}(t)$, where $\mathcal{H}\{\cdot\}$ denotes

the Hilbert transform. Then the phase is $\phi_j(t) = \tan^{-1}(\Im\{d_j(t)\}/\Re\{d_j(t)\}) = \tan^{-1}(d_j^{(H)}(t)/d_j(t))$, and the amplitude is $A_j(t) = |d_j^a(t)| = \sqrt{\{d_j^2(t) + [d_j^{(H)}(t)]^2\}}$. The corresponding instantaneous frequency is $\nu_j(t) = \phi_j'(t)/(2\pi)$. The Hilbert energy spectrum of $d_j(t)$ is given by

$$S_j(t,f) = A_j^2(t)\delta(f - \nu_j(t)),$$

where $\delta(\cdot)$ is the Dirac delta function. The overall Hilbert energy spectrum for $x(t)$ is defined (Huang *et al* 1998) as the amalgamation of the spectra of each of the $d_j(t)$'s. We can also define a time-varying mean frequency for the entire signal,

$$\nu(t) = \frac{\sum_{j=1}^{J_0} A_j^2(t)\,\nu_j(t)}{\sum_{j=1}^{J_0} A_j^2(t)},$$

a useful 'summary' curve.

4. Discrete time wavelet decompositions

(a) Wavelet projections

Analogous to section 2(a) we consider projections of a real-valued signal which attempt to produce monocomponent separation, only here we use discrete-time methods. Assume we have sampled a continuous-time signal at intervals $\Delta t = 1$ to get a vector of observations $\mathbf{X} = [X_0, \ldots, X_{N-1}]^T$. Assume also we have a series of projection matrices $\mathbf{P}_{\mathcal{R}_j}$, which project a vector $\mathbf{X}$ into a particular time-frequency region $\mathcal{R}_j$. These operators need not be orthogonal but we shall assume $\sum_{j=1}^{J_0+1} \mathbf{P}_{\mathcal{R}_j}\mathbf{X} = \mathbf{X}$ for some particular J_0. For $j = 1, \ldots, J_0$ define $\tilde{\mathcal{D}}_j = \mathbf{P}_{\mathcal{R}_j}\mathbf{X}$ and $\tilde{\mathcal{S}}_{J_0} = \mathbf{P}_{\mathcal{R}_{J_0+1}}\mathbf{X}$. Then

$$\mathbf{X} = \sum_{j=1}^{J_0} \tilde{\mathcal{D}}_j + \tilde{\mathcal{S}}_{J_0}. \tag{4.1}$$

Decomposition (4.1) is the discrete equivalent of (2.1). So we can decompose the sampled signal into a linear combination of the contribution of $\mathbf{X}$ in projected spaces. If the time-frequency plane has been 'tiled' ideally, each projection will contain at most one monocomponent signal for which the concept of instantaneous frequency is well defined.

The maximal overlap discrete wavelet transform (MODWT) is a version of the discrete wavelet transform which does away with the downsampling step, thus ensuring the important property of circular shift equivariance. Transforms that are essentially the same include the undecimated discrete wavelet transform (Shensa, 1992) and translation invariant discrete wavelet transform (Liang and Parks, 1996). Full details of the MODWT may be found in Percival and Walden (2000, chapter 5); we adopt their notation here. For $t = 0, \ldots, N-1$ the jth level wavelet and scaling coefficients are given by $\widetilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{(t-l) \bmod N}$ and $\widetilde{V}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{(t-l) \bmod N}$, where $\{\tilde{h}_{j,l}\}$ and $\{\tilde{g}_{j,l}\}$ are level-j filters of length $L_j = (2^j - 1)(L - 1) + 1$ derived from the basic DWT wavelet and scaling filters $\{h_l\}$

and $\{g_l\}$ of length L. (Note the tilde notation is used for MODWT filters and coefficients to distinguish them from the DWT equivalents). In matrix form we can write $\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j \mathbf{X}$ and $\widetilde{\mathbf{V}}_j = \widetilde{\mathcal{V}}_j \mathbf{X}$, where $\widetilde{\mathbf{W}}_j$ and $\widetilde{\mathbf{V}}_j$ are N-length column vectors, and $\widetilde{\mathcal{W}}_j$ and $\widetilde{\mathcal{V}}_j$ are $N \times N$ matrices composed of elements from $\{\tilde{h}_{j,l}\}$ and $\{\tilde{g}_{j,l}\}$ respectively, and zeros. Furthermore, the signal can be recovered as in (4.1), where $\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{W}}_j^T \widetilde{\mathcal{W}}_j \mathbf{X}$ and $\widetilde{\mathcal{S}}_j = \widetilde{\mathcal{V}}_j^T \widetilde{\mathcal{V}}_j \mathbf{X}$ (Percival and Walden 2000, p. 173). The corresponding projection matrices are $\mathbf{P}_{\mathcal{R}_j} = \widetilde{\mathcal{W}}_j^T \widetilde{\mathcal{W}}_j$ and $\mathbf{P}_{\mathcal{R}_{J_0+1}} = \widetilde{\mathcal{V}}_{J_0}^T \widetilde{\mathcal{V}}_{J_0}$, where $\mathbf{P}_{\mathcal{R}_j}$ is approximately the identity matrix on time-frequency region $\mathcal{R}_j$, and zero elsewhere, where

$$\mathcal{R}_j = \begin{cases} [0, (N-1)] \times \left\{ -\left[\frac{1}{2^j}, \frac{1}{2^{j+1}}\right) \cup \left(\frac{1}{2^{j+1}}, \frac{1}{2^j}\right] \right\} & \text{if } 1 \le j \le J_0 \\ [0, (N-1)] \times \left\{ \left[-\frac{1}{2^{J_0+1}}, \frac{1}{2^{J_0+1}} \right] \right\} & \text{if } j = J_0 + 1. \end{cases}$$

This is then an example of a set of projection operators as described above. $\widetilde{\mathcal{S}}_j$ represents the over-all time-varying mean trend, while the $\widetilde{\mathcal{D}}_j$ represent local time-frequency information. The MODWT of course tiles the time-frequency plane in octave bands.

The elements of $\widetilde{\mathcal{S}}_j$ and $\widetilde{\mathcal{D}}_j$ are called jth level smooth and detail coefficients. We can write (Percival and Walden 2000, p. 199)

$$\widetilde{\mathcal{D}}_{j,t} = \sum_{l=-(L_j-1)}^{L_j-1} \tilde{h}_j \star \tilde{h}_{j,l} X_{(t-l) \bmod N},$$

where $\tilde{h}_j \star \tilde{h}_{j,l}$ denotes the cross-correlation of $\{\tilde{h}_{j,l}\}$ with itself, i.e., the autocorrelation of $\{\tilde{h}_{j,l}\}$. The effective filter is symmetric, zero-phase, and considerably smoother than $\{\tilde{h}_{j,l}\}$ itself. Analogous results hold for $\widetilde{\mathcal{S}}_j$. So while the utility of wavelet methods in efficiently representing discontinuities is well-known, in fact our projections are well-suited to represent smooth, possibly sinusoidal, sequences.

(*b*) *The Hilbert spectrum via wavelets*

The MODWT detail coefficients for level j, $\{\widetilde{\mathcal{D}}_{j,t}, t = 0, \dots, N-1\}$, will be real for any real-valued wavelet filters. We can create analytic wavelet details via

$$\widetilde{\mathcal{D}}_{j,t} + \mathrm{i}(\widetilde{\mathcal{D}}_j * q)_t = \widetilde{\mathcal{D}}_{j,t} + \mathrm{i}\widetilde{\mathcal{D}}_{j,t}^{(H)}, \tag{4.2}$$

where $\{q_l\}$ is the Hilbert transform filter and $\widetilde{\mathcal{D}}_{j,t}^{(H)}$ denotes the discrete Hilbert transformed version of $\widetilde{\mathcal{D}}_{j,t}$.

The amplitude and phase sequences are given by

$$A_{j,t} = \sqrt{\{\widetilde{\mathcal{D}}_{j,t}^2 + [\widetilde{\mathcal{D}}_{j,t}^{(H)}]^2\}}; \ \phi_{j,t} = \tan^{-1}(\widetilde{\mathcal{D}}_{j,t}^{(H)}/\widetilde{\mathcal{D}}_{j,t}). \tag{4.3}$$

We compute the derivative of the phase at discrete time t via a fourth-order generalized phase difference estimator (Boashash, 1992b, p. 542),

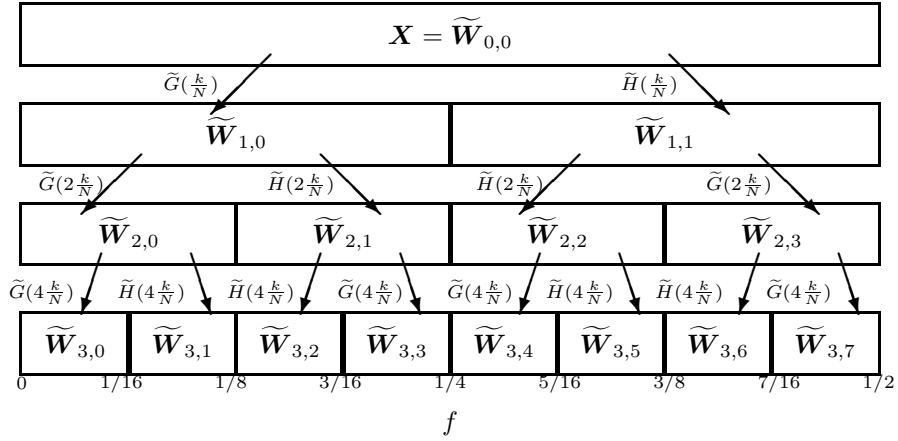$$\phi'_{j,t} = \frac{1}{12}(\phi_{j,t-2} - 8\phi_{j,t-1} + 8\phi_{j,t+1} - \phi_{j,t+2}). \tag{4.4}$$

Figure 1. First three levels of the MODWPT of **X**. All coefficient vectors together form a (maximal overlap) wavelet packet table.

This formula is skew-symmetric about time t. At the end-points we cannot use this centred five-point formula but use instead

$$\phi'_{j,t} = \pm\frac{1}{12}(-25\phi_{j,t} + 48\phi_{j,t\pm1} - 36\phi_{j,t\pm2} + 16\phi_{j,t\pm3} - 3\phi_{j,t\pm4}). \qquad (4.5)$$

Both formulae can be found in Burden and Faires (1993, p. 161). The instantaneous frequency is then computed from $\nu_{j,t} = \phi'_{j,t}/(2\pi)$.

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

$$S_{j;t,f_l} = A^2_{j,t}\delta_{l,\langle\nu_{j,t}/\Delta f\rangle}, \qquad t = 0, \ldots, N-1; \ f_l = 0, \ldots, 1/2, \qquad (4.6)$$

where $\delta_{l,m}$ is the Kronecker delta, and $\langle x \rangle$ denotes the integer closest to x; we used $M = 512$. The overall Hilbert energy spectrum is the amalgamation of the $S_{j;t,f_l}$'s over the j. The time-varying mean frequency for the entire signal follows as

$$\nu_t = \frac{\sum_{j=1}^{J_0} A^2_{j,t}\nu_{j,t}}{\sum_{j=1}^{J_0} A^2_{j,t}}. \qquad (4.7)$$

5. Discrete time wavelet packet decompositions

(a) Wavelet packet projections

The MODWT has the disadvantage of imposing a fixed octave band tiling on the time-frequency plane. The maximal overlap discrete wavelet packet transform (MODWPT) (Walden & Contreras Cristán, 1998) is a generalization of the MODWT which at level j of the transform partitions the frequency axis into 2^j equal width frequency bands, often labelled $n = 0, \ldots, 2^j - 1$. This gives better frequency resolution. Let $\widetilde{G}(f) = \sum_{l=0}^{L-1}\tilde{g}_l e^{-i2\pi fl}$ be the transfer function of $\{\tilde{g}_l\}$ where
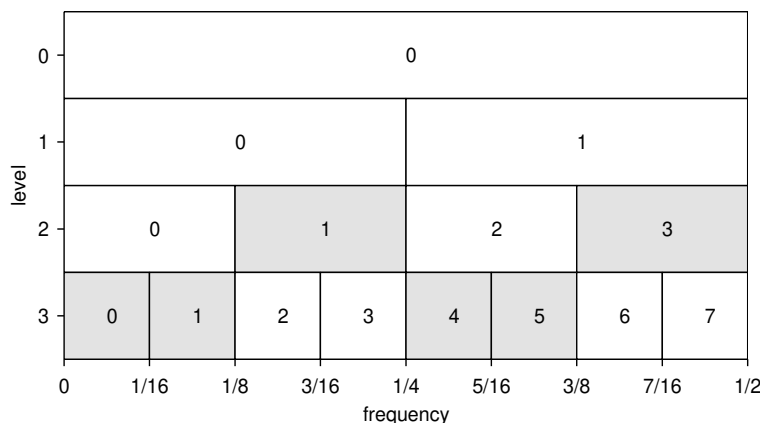
Figure 2. The shaded segments determine a possible disjoint dyadic decomposition — a partitioning of the frequency interval $[0, 1/2]$ — which may be derived from the WPT of Fig. 1.

$\tilde{g}_l = g_l/\sqrt{2}$, and let the transfer function corresponding to $\{\tilde{h}_l\}$ with $\tilde{h}_l = h_l/\sqrt{2}$, be defined similarly. At the first level of the transform $\mathbf{X}$ is circularly filtered by the low-pass filter $\{\tilde{g}_l\}$, with corresponding transfer function $\widetilde{G}(f)$, to give an N-length vector of first level coefficients $\widetilde{\mathbf{W}}_{1,0}$ and $\mathbf{X}$ is also circularly filtered by the high-pass filter $\{\tilde{h}_l\}$, with transfer function $\widetilde{H}(f)$, to give the vector $\widetilde{\mathbf{W}}_{1,1}$. For subsequent levels j of the transform we insert $2^{j-1} - 1$ zeros, $j \geq 1$, between the elements of $\{\tilde{g}_l\}$; the resulting filter has a transfer function given by $\widetilde{G}(2^{j-1}f)$. We can do likewise for $\{\tilde{h}_l\}$ to obtain $\widetilde{H}(2^{j-1}f)$. The jth level coefficients are obtained by filtering the level $j - 1$ coefficients with the circular filter having discrete Fourier transform $\{\widetilde{H}(2^{j-1}\frac{k}{N})\}$ or $\{\widetilde{G}(2^{j-1}\frac{k}{N})\}$, as appropriate. The filtering steps up to level $j = 3$ are shown in Fig. 1 and the collection of MODWPT coefficients is called a (maximal overlap) wavelet packet table (WPT). If, as here, the ordering of the filters means that at level j, band n is associated with frequencies $\mathcal{I}_{j,n} = \left[-\frac{1}{2^{j+1}}, \frac{1}{2^{j+1}} \right]$, for $n = 0$, and $\mathcal{I}_{j,n} = -\left[\frac{n+1}{2^{j+1}}, \frac{n}{2^{j+1}} \right) \cup \left(\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}} \right]$, for $n > 0$, then the transform is said to be 'sequency ordered.'

At level j and frequency band n the wavelet packet coefficients for $t = 0, \ldots, N - 1$ can be written as $\widetilde{W}_{j,n,t} = \sum_{l=0}^{L_j - 1} \tilde{u}_{j,n,l} X_{(t-l) \bmod N}$ where $\{\tilde{u}_{j,n,l}\}$ is a jth level and nth band MODWPT wavelet packet filter of length $L_j = (2^j - 1)(L - 1) + 1$ with a transfer function which may be derived from the filtering steps shown in Fig. 1; for example for $j = 3, n = 3$ this transfer function is given by $\widetilde{G}(f)\widetilde{H}(2f)\widetilde{G}(4f)$.

In matrix form we can write $\widetilde{\mathbf{W}}_{j,n} = \widetilde{\mathcal{W}}_{j,n}\mathbf{X}$ where $\widetilde{\mathbf{W}}_{j,n}$ is a N-length column vector, and $\widetilde{\mathcal{W}}_{j,n}$ is a $N \times N$ matrix composed of elements from $\{\tilde{u}_{j,n,l}\}$ and zeros. If we let $\widetilde{\mathcal{D}}_{j,n} = \widetilde{\mathcal{W}}_{j,n}^T \widetilde{\mathcal{W}}_{j,n}\mathbf{X}$ then the signal can be recovered via $\mathbf{X} = \sum_{n=0}^{2^j - 1} \widetilde{\mathcal{D}}_{j,n}$. For $n > 0$ the set $\{\widetilde{\mathcal{D}}_{j,n,t}, t = 0, \ldots, N - 1\}$ are level j, frequency band n, MODWPT detail coefficients.

However, the decomposition in (4.1) makes use of different levels of the transform, and the same can be done for the MODWPT. The doublets (j, n) that form the indices for the MODWPT coefficient vectors $\widetilde{\mathbf{W}}_{j,n}$ can be collected together to form a set $\mathcal{N} \equiv \{(j, n) : j = 0, \ldots, J_0; \ n = 0, \ldots, 2^j - 1\}$. Each $(j, n) \in \mathcal{N}$

is associated with the frequency interval $\mathcal{I}_{j,n}$. The WPT can be used to form a large collection of different transforms known as *disjoint dyadic decompositions*. By definition, each such decomposition is associated with a subset, say $\mathcal{C}$, of doublets from $\mathcal{N}$ that satisfies two properties. First, the union of all the frequency intervals $\mathcal{I}_{j,n}$ for each $(j,n) \in \mathcal{C}$ is exactly the interval $[0, 1/2]$, and, second, if (j,n) and (j',n') are any two distinct elements of $\mathcal{C}$, then $\mathcal{I}_{j,n}$ and $\mathcal{I}_{j',n'}$ have no frequencies in common. One such decomposition is shown by the grey shading in Fig. 2; here $\mathcal{C} = \{(2,1), (2,3), (3,0), (3,1), (3,4), (3,5)\})$. Fundamentally, for any such decomposition $\mathcal{C}$ we have

$$\mathbf{X} = \sum_{(j,n>0\in\mathcal{C})} \widetilde{\mathcal{D}}_{j,n} + \widetilde{\mathcal{D}}_{j',0},$$

for some $j' \leq J_0$. The last term covers low frequencies and trends.

The projection matrices are $\mathbf{P}_{\mathcal{R}_{j,n}} = \widehat{\mathcal{W}}_{j,n}^T \widehat{\mathcal{W}}_{j,n}$ where $\mathbf{P}_{\mathcal{R}_{j,n}}$ is approximately the identity matrix on time-frequency region $\mathcal{R}_{j,n}$, and zero elsewhere, where

$$\mathcal{R}_{j,n} = [0, (N-1)] \times \mathcal{I}_{j,n}.$$

This is another example of a set of projection operators.

When utilising the wavelet packet transform we must decide on some method of choosing a suitable disjoint dyadic decomposition $\mathcal{C}$ which has the effect of producing a time-frequency tiling commensurate with separating the original signal into a set of monocomponent signals.

(b) *Choosing the disjoint dyadic decomposition*

Consider an algorithm starting at level $j = J_0$. If the lowest frequency feature of interest in the signal occurs at a frequency higher than f_L say, we need to choose J_0 so that $\frac{1}{2^{J_0+1}} < f_L$. (This is because the Hilbert spectrum discards the lowest frequency components to avoid the zero frequency or infinite scale). Consequently, $\widetilde{\mathcal{D}}_{J_0,0}$ and $\widetilde{\mathcal{D}}_{J_0,1}$ are always preserved in any decomposition.

Now consider a measure of the bandwidth of the details $\widetilde{\mathcal{D}}_{j,n,t}$, $t = 0, \ldots, N-1$ at level j, band $n > 0$. The discrete Fourier transform (DFT) of the details at frequencies $f_k = k/N$ is

$$\xi_{j,n,k} = \sum_{t=0}^{N-1} \widetilde{\mathcal{D}}_{j,n,t} \mathrm{e}^{-\mathrm{i}2\pi tk/N}, \quad k = 0, \ldots, N-1,$$

and the normalised modulus squared, which sums to unity, takes the form

$$\Upsilon_{j,n,k} = |\xi_{j,n,k}|^2 \Big/ \sum_{k=0}^{\lfloor N/2 \rfloor} |\xi_{j,n,k}|^2, \quad k = 0, \ldots, \lfloor N/2 \rfloor,$$

where $\lfloor x \rfloor$ denotes the greatest integer $\leq x$. We denote the bandwidth by $B_{j,n}$ with

$$B_{j,n}^2 = \sum_{k=0}^{\lfloor N/2 \rfloor} (f_k - \langle f \rangle_{j,n})^2 \, \Upsilon_{j,n,k} \quad \text{where} \quad \langle f \rangle_{j,n} = \sum_{k=0}^{\lfloor N/2 \rfloor} f_k \, \Upsilon_{j,n,k}.$$

The fundamental idea behind the decomposition algorithm is illustrated in Fig. 3. Consider whether a frequency band $\mathcal{I}_{j,n}$ should be split into two subbands $\mathcal{I}_{j+1,2n}$ and $\mathcal{I}_{j+1,2n+1}$; we note
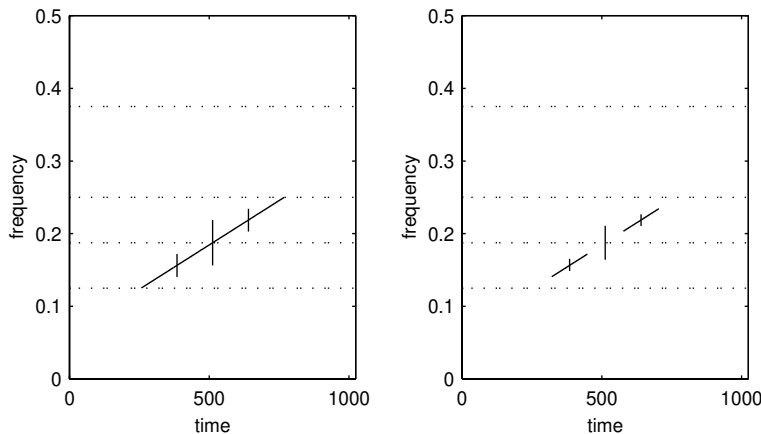
Figure 3. Illustration of points (i) (left) and (ii) (right) of the decomposition algorithm. Left: one chirp signal extending from $f = 0.125$ to $f = 0.25$ (diagonal line) The bandwidth (long vertical line) is 0.0625. If split into two, we obtain two chirps extending from $f = 0.125$ to 0.1875 and one from 0.1875 to 0.25 The separate chirps both have a bandwidth (short vertical lines) of 0.03125 which is half that of the entire signal. Right: two separate chirps, each with bandwidths of about 0.0166. If they are regarded as one signal, then this has a bandwidth of about $0.04628 > 0.0166 + 0.0166$.

(i) if $\mathcal{I}_{j,n}$ contains a monocomponent signal then if it is split, the sum of the ensuing bandwidths is equal to that of the entire signal (Fig. 3, left), while

(ii) if the subbands each contain separate monocomponent signals then the sum of their bandwidths will be less than the bandwidth obtained if they are regarded as one signal (Fig. 3, right).

The algorithm starts at level $j = J_0 - 1$ and moves up the tree. If

$$B_{j,n} > \alpha(B_{j+1,2n} + B_{j+1,2n+1}) \tag{5.1}$$

we keep $\widetilde{\mathcal{D}}_{j+1,2n}$ and $\widetilde{\mathcal{D}}_{j+1,2n+1}$ (i.e., use subbands $\mathcal{I}_{j+1,2n}$ and $\mathcal{I}_{j+1,2n+1}$) and in order to ensure that the split remains as we move up the tree we set $B_{j,n} = 0$. Here α is a constant chosen to avoid splitting the band due simply to random variation. It is more important to avoid not splitting a band when we should, rather than splitting when not required, as we seek monocomponent signals in each band — if we do not split a band when we should, then a meaningless instantaneous frequency will result. A value $\alpha = 1.2$ is suitable for this requirement.

A slight problem arises when dealing with time-frequency tiles which are empty of signal, so that the bandwidth is ill-defined. It is advisable to factor in the actual energy of the details at level j and band n and modify the bracketed term in (5.1) to read

$$\frac{B_{j+1,2n} \sum_{t=0}^{N-1} A_{j+1,2n,t}^2 + B_{j+1,2n+1} \sum_{t=0}^{N-1} A_{j+1,2n+1,t}^2}{\sum_{t=0}^{N-1} A_{j+1,2n,t}^2 + \sum_{t=0}^{N-1} A_{j+1,2n+1,t}^2} \tag{5.2}$$

where $A_{j,n,t}$, $t = 0, \ldots, N-1$ are the amplitudes of the analytic MODWPT details, defined next.

<center>(*c*) *The Hilbert spectrum via wavelet packets*</center>

We can create analytic MODWPT details for level j and band $n > 0$ using

$$\widetilde{\mathcal{D}}^a_{j,n,t} = \widetilde{\mathcal{D}}_{j,n,t} + \mathrm{i}(\widetilde{\mathcal{D}}_{j,n} * q)_t = \widetilde{\mathcal{D}}_{j,n,t} + \mathrm{i}\widetilde{\mathcal{D}}^{(H)}_{j,n,t}. \tag{5.3}$$

Given the analytic details (5.3), the corresponding amplitude, $A_{j,n,t}$, phase, $\phi_{j,n,t}$ and instantaneous frequency, $\nu_{j,n,t}$, are computed as in section 4(b), (an additional subscript 'n' being required throughout).

The Hilbert energy spectrum of $\{\widetilde{\mathcal{D}}_{j,n,t}\}$ is given by

$$S_{j,n;t,f_l} = A^2_{j,n,t}\delta_{l,\langle \nu_{j,n,t}/\Delta f\rangle}, \qquad t = 0,\dots,N-1;\ f_l = 0,\dots,1/2, \tag{5.4}$$

and the time-varying mean frequency for the entire signal follows as

$$\nu_t = \frac{\sum_{(j,n>0\in\mathcal{C})} A^2_{j,n,t}\nu_{j,n,t}}{\sum_{(j,n>0\in\mathcal{C})} A^2_{j,n,t}}. \tag{5.5}$$

<center>(*d*) *Leakage of time-frequency projections*</center>

A major problem with any projection is the unavoidable leakage caused by the fact that no projection can create a signal which is simultaneously time and frequency limited. Using finite-length filters, wavelet packets lose some of their frequency resolution at high frequencies (e.g., Hess-Nielsen & Wickerhauser, 1996). Nielsen (2001) derived classes of half-band low-pass DWT filters $\{g_l, l = 0,\dots,L-1\}$ such that the L_p norm of the difference between the power transfer function $|G(f)|^2$, (where $G(f)$ is the Fourier transform of $\{g_l\}$), and the power transfer function of the ideal half-band filter — rectangular on $[-1/4,1/4]$ — is $O(1/L)$ for $p \in [1,2]$. One class of such optimal asymptotic frequency resolution filters are the Fejér-Korovkin filters.

Fig. 4 gives the power transfer functions of level $j = 3$ wavelet packet filters $\{\tilde{u}_{j,n,l}\}$ generated from the $L = 18$ point Fejér-Korovkin filter $\{g_l\}$ — abbreviated to FK(18) — and $L = 20$ point Daubechies least asymmetric filter $\{g_l\}$ — abbreviated to LA(20). Even though it is shorter by two points than the frequently used Daubechies least asymmetric filter, the Fejér-Korovkin filter gives rise to the better concentrated wavelet packets, the power transfer functions being in particular more rectangular over their pass-bands $\mathcal{I}_{j,n}$. The minimally short two point Haar wavelet filter generates wavelet packet filters with much poorer concentration than the LA(20). The efficacy of the FK(18), LA(20) and Haar wavelet filters for Hilbert spectrum or time-varying mean frequency calculation is demonstrated in the examples which follow.

6. Examples

In this section we will compare plots of the Hilbert energy spectrum and the time-varying mean frequency made using the MODWT and EMD approaches. For the EMD approach the same computational scheme is used as set out in Section 4b except that the wavelet detail coefficients are replaced by the IMF sequences; in particular the same phase derivative method is used — see (4.4) and (4.5).
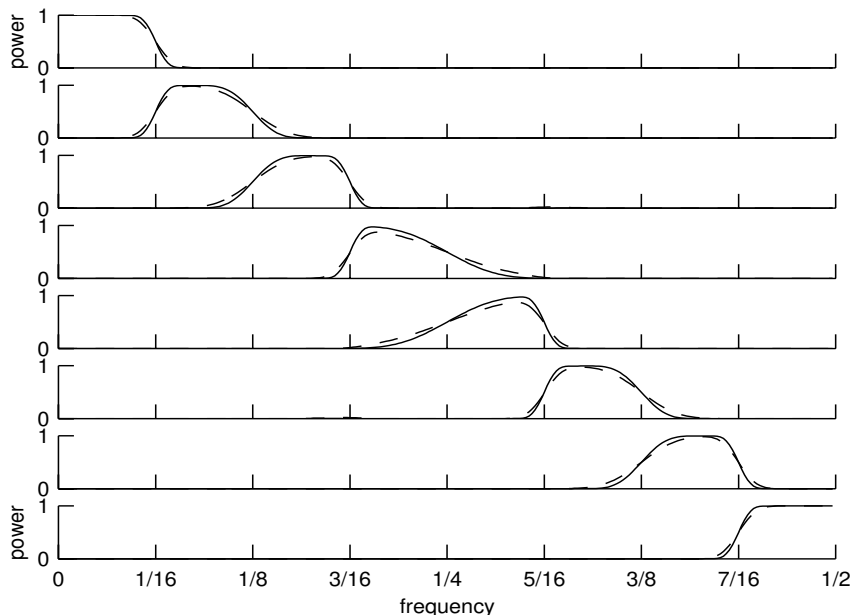
Figure 4. Power transfer functions of level $j = 3$ wavelet filters $\{\tilde{u}_{j,n,l}\}$ for (from top to bottom) frequency bands $n = 0, \ldots, 7$, generated from the 18-point Fejér-Korovkin filter (solid lines) or 20-point Daubechies least asymmetric filter (dashed lines).

(a) Sum of sinusoids

Consider a signal which is the sum of two equal-amplitude sine waves:

$$x_1(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t), \tag{6.1}$$

where $f_1 = 0.08$ and $f_2 = 0.16$; the theoretical mean frequency is 0.12.

Figs. 5(a)–(c) show the Hilbert energy spectrum and the time-varying mean frequency calculated using the MODWT with $J_0 = 5$ when this signal is sampled with a unit sample interval, $N = 512$ times starting at $t = 0$. The half-band low-pass filter $\{g_l, \, l = 0, \ldots, L-1\}$ used in the MODWT was (a) Haar, (b) Daubechies LA(20), and (c) Fejér-Korovkin FK(18). Leakage-generated oscillations in the Hilbert spectrum are present in all three cases. Leakage of the projections means that in a particular frequency band not only is the appropriate sinusoid present, but also sinusoidal contributions from neighbouring bands, with all the resulting complications such as (1.1). These oscillations decrease to almost nothing as the concentration of the filters improves; the FK(18) result is strikingly good. Fig. 6 gives the MODWT detail sequences for $j = 1, \ldots, 5$ for the case of the Fejér-Korovkin FK(18) filter. The five sequences are ordered (top to bottom) according to the top five frequency bands delineated by dotted lines in Figs. 5(c); it is clear that leakage between bands is very small. The representation of the sinusoids is excellent as predicted in the discussion at the end of §4a.

Fig. 7 gives the five IMFs produced by EMD. The differences between the components determined by the MODWT in Fig. 6 and EMD in Fig. 7 do not seem significant, yet they are; this can be seen from the Hilbert spectrum given in Fig. 5(d)
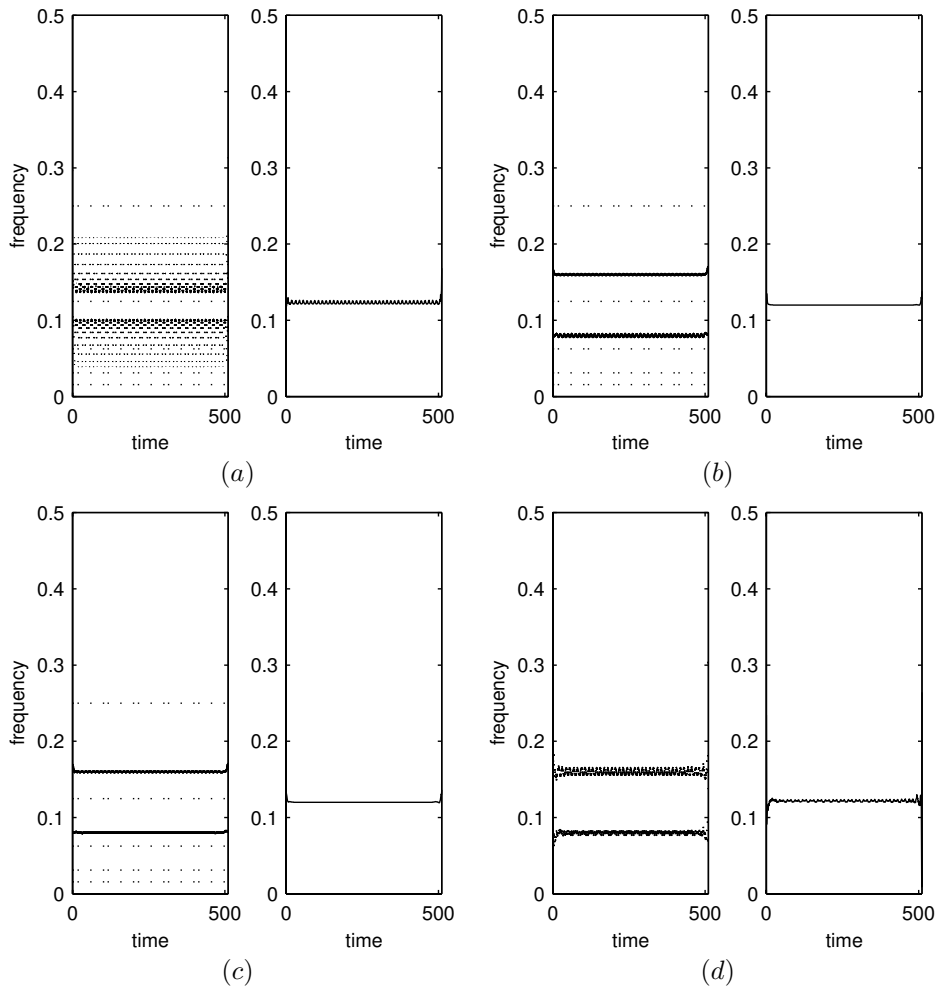
Figure 5. Sum of sines example. $f_1 = 0.08$ and $f_2 = 0.16$ and theoretical mean frequency is 0.12 (a) Left: the Hilbert spectrum using the MODWT, Haar wavelet filter, and $J_0 = 5$. Right: corresponding time-varying mean frequency. (b) Left: the Hilbert spectrum using the MODWT, Daubechies LA(20) wavelet filter, and $J_0 = 5$. Right: corresponding time-varying mean frequency. (c) Left: the Hilbert spectrum using the MODWT, Fejér-Korovkin FK(18) wavelet filter, and $J_0 = 5$. Right: corresponding time-varying mean frequency. The horizontal dotted lines delineate the MODWT octave bands. (d) Left: the Hilbert spectrum using EMD and five IMFs. Right: corresponding time-varying mean frequency.

using the five IMFs where relatively poor frequency resolution is achieved and the result is inferior to all but the Haar MODWT. This is somewhat surprising since we would expect the EMD method to produce good representations of sinusoids since theoretically sinusoidal components are IMFs.
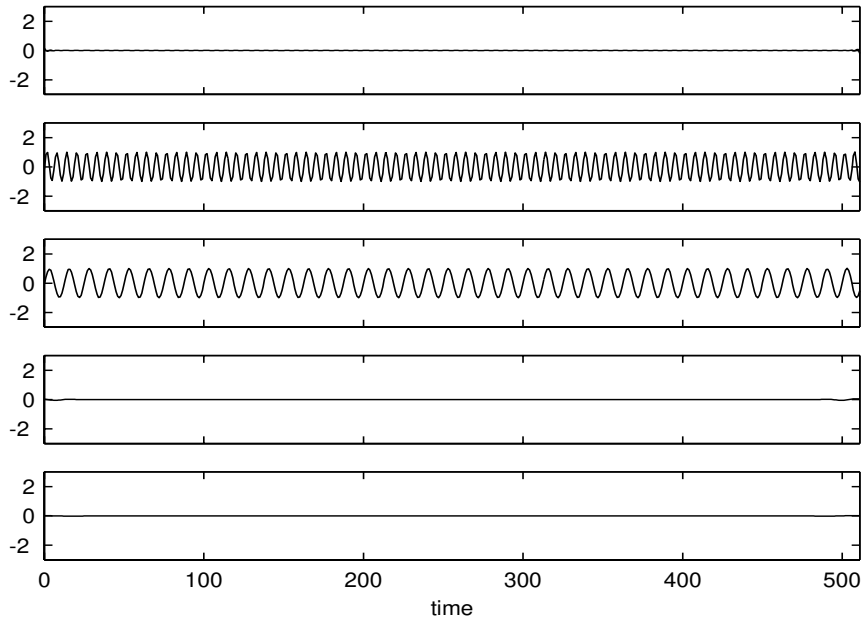
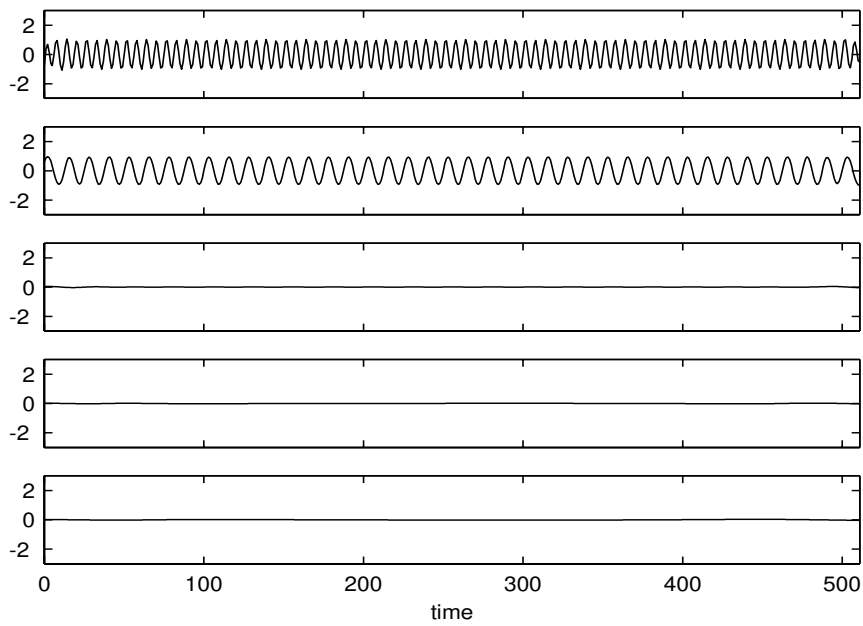Figure 6. The five MODWT detail sequences (FK(18) filter) for sum of sines.

Figure 7. Five IMFs for sum of sines.

(*b*) *Intermittent and complete sinusoids*

Here we consider a signal made up from four sinusoids:

$$x_2(t) = 1.2\sin(2\pi f_1 t)\mathcal{I}_{T_1}(t) + \sin(2\pi f_2 t)\mathcal{I}_{T_2}(t) + 1.2\sin(2\pi f_3 t) + \sin(2\pi f_4 t), \tag{6.2}$$
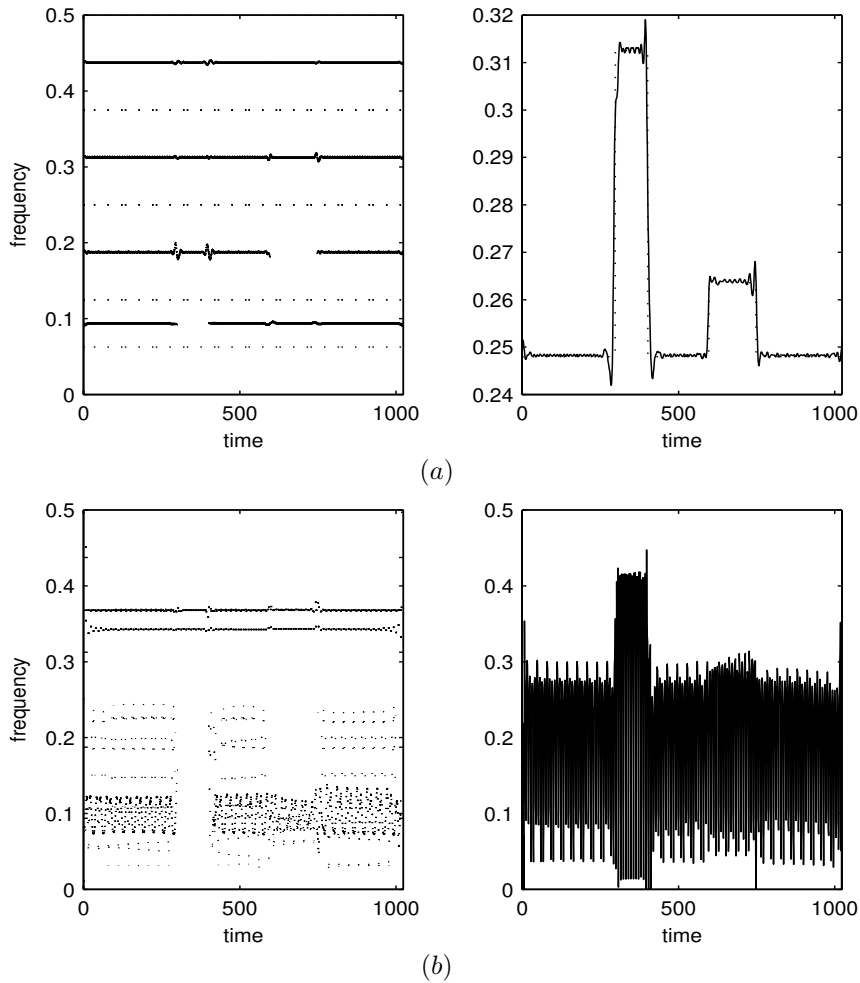
Figure 8. Intermittent and complete sines example. $f_1 = 0.0938$, $f_2 = 0.1875$, $f_3 = 0.3125$ and $f_4 = 0.4375$. (a) Left: the Hilbert spectrum using the MODWPT, Fejér-Korovkin FK(18) wavelet filter, and $J_0 = 4$. The horizontal dotted lines delineate the MODWPT frequency bands. Additional short tick marks indicating the frequencies $f_1, \ldots, f_4$ are not visible under the estimated frequencies. Right: corresponding time-varying mean frequency on a magnified scale, with the theoretical mean frequency (dotted) overlain. (b) Left: the Hilbert spectrum using EMD and nine IMFs. The additional short tick marks indicating the frequencies $f_1, \ldots, f_4$ are now visible. Right: corresponding time-varying mean frequency.

where $f_1 = 0.0938$, $f_2 = 0.1875$, $f_3 = 0.3125$, $f_4 = 0.4375$, $\mathcal{I}_T(t) = 1$ if $t \in T$ and zero otherwise, $T_1 = [0, 0.297] \cup [0.402, \infty)$ and $T_2 = [0, 0.597] \cup [0.748, \infty)$. The first two sinusoids are intermittent. This signal is sampled with a unit sample interval, $N = 1024$ times starting at $t = 0$. The two higher frequency sinusoids have frequencies both falling in the band $(1/4, 1/2)$ and so cannot be separated by the MODWT. The more general MODWPT is thus required, and we can allow the algorithm in section 5(b) to select an appropriate disjoint dyadic decomposition —

the results are shown in Fig. 8(a). The MODWPT method gives a very good and accurate description of the signals, with a few oscillations near the discontinuities in the signal. The corresponding results using EMD are given in Fig. 8(b). This used the nine IMFs found by the algorithm. (Use of only the the first four IMFs — there are four monocomponent signals present — produced an indistinguishable result.) Huge oscillations are found with the EMD algorithm, and there are also problems with mode-mixing.

(*c*) *Application to real data*

We now analyse the digitized echolocation pulse — shown in Fig. 9(a) — emitted by the Large Brown Bat, *Eptesicus Fuscus*. The sample interval is 7μs. Fig. 9(b) shows the MODWPT results, while the corresponding results for EMD are shown in Fig. 9(c). For EMD the seven IMFs found by the algorithm are used, but taking only the first four caused no appreciable difference. The MODWPT clearly outperforms the EMD approach in resolving the hyperbolic chirps present in this signal. Notice that the algorithm of section 5b for choosing the disjoint dyadic decomposition for the MODWPT has worked successfully: at any time the different frequency chirps have been separated into different frequency subbands.

7. Denoising

(*a*) *Theory*

Since we are in general dealing with signals contaminated by noise, it makes sense to denoise the signal where possible. We assume $x(t) = b(t) + \epsilon(t)$, where $b(t)$ is deterministic signal and $\epsilon(t)$ is real-valued noise such that discretization at unit sample interval gives the N-length vector

$$\mathbf{X} = \mathbf{b} + \boldsymbol{\epsilon}, \tag{7.1}$$

where $\{\epsilon_t, t = 0, \dots, N-1\}$ is a white Gaussian noise sequence with mean zero and variance σ_ϵ^2. Coifman and Donoho (1995) introduced the technique of 'cycle spinning' for denoising of $\mathbf{X}$ in a translation invariant way; this is equivalent (Percival and Walden, 2000, p. 429) to applying standard thresholding to the MODWT wavelet coefficients. Here our denoising scheme will determine which analytic wavelet detail magnitudes $A_{j,t}$ should be used in the Hilbert energy spectrum of (4.6) and time-varying mean frequency given by (4.7). The MODWT details vector is given by $\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{W}}_j^T \widetilde{\mathcal{W}}_j \mathbf{X}$ so that the additive noise vector is transformed to the noise details vector $\widetilde{\mathcal{E}}_j = \widetilde{\mathcal{W}}_j^T \widetilde{\mathcal{W}}_j \boldsymbol{\epsilon}$. But the covariance of the (Gaussian) elements of $\widetilde{\mathcal{E}}_j$ is given by (Appendix A)

$$\text{cov}\{\widetilde{\mathcal{E}}_{j,t}, \widetilde{\mathcal{E}}_{j,t'}\} \approx \begin{cases} \sigma_\epsilon^2/2^j, & t = t', \\ \sigma_\epsilon^2 \frac{\{\sin[\pi(t-t')/2^{j-1}] - \sin[\pi(t-t')/2^j]\}}{[\pi(t-t')]}, & t \neq t'. \end{cases} \tag{7.2}$$

Now suppose $\{\eta_t\}$ is a stationary process with $\text{cov}\{\eta_t, \eta_{t'}\} = s_{t-t'}$, say, and define the analytic process by

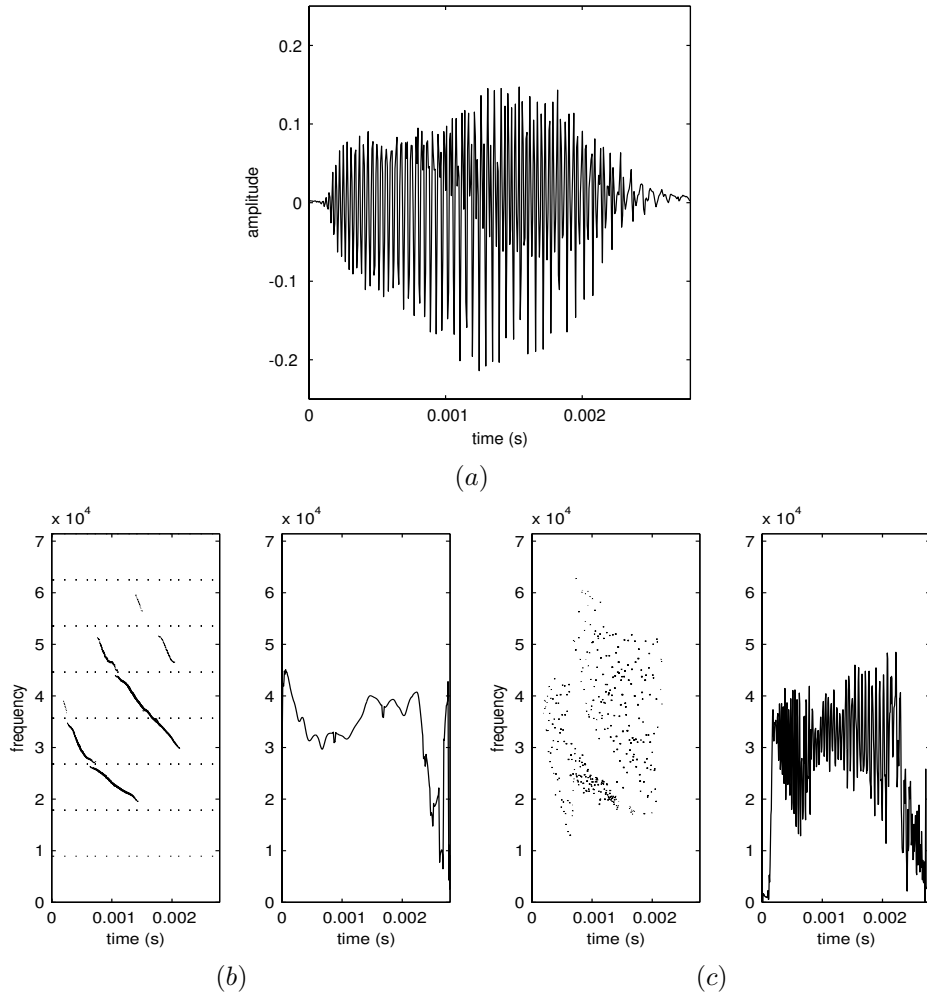$$\eta_t^{(a)} = \eta_t + \mathrm{i}(\eta * q)_t = \eta_t + \mathrm{i}\eta_t^{(H)},$$

Figure 9. (a) Bat signal. (b) Left: the Hilbert spectrum using the MODWPT, Fejér-Korovkin FK(18) wavelet, and $J_0 = 3$. Right: corresponding time-varying mean frequency. (c) Left: the Hilbert spectrum using EMD and seven IMFs. Right: corresponding time-varying mean frequency.

where, $\{q_l\}$ is the Hilbert transform filter. Then, since the Hilbert transform is an all-pass filter, it does not affect the covariance structure of $\{\eta_t\}$ so that $\mathrm{cov}\{\eta_t^{(H)}, \eta_{t'}^{(H)}\} = s_{t-t'}$, also. Moreover,

$$E\{\eta_t \eta_{t'}^{(H)}\} = E\{\eta_t \sum_l q_l \eta_{t'-l}\} = \sum_\tau s_{t-t'+l}\, q_l,$$

so that if $t = t'$ we get $\mathrm{cov}\{\eta_t, \eta_t^{(H)}\} = \sum_l s_l q_l = 0$, since $\{s_l\}$ is even and $\{q_l\}$ is odd. Hence at any time point t, the random variables η_t and $\eta_t^{(H)}$ are uncorrelated; however this does not mean the *processes* $\{\eta_t\}$ and $\{\eta_t^{(H)}\}$ are uncorrelated, let alone independent. However, if $\{\eta_t\}$ is Gaussian, then the processes $\{\eta_t\}$ and $\{\eta_t^{(H)}\}$ are
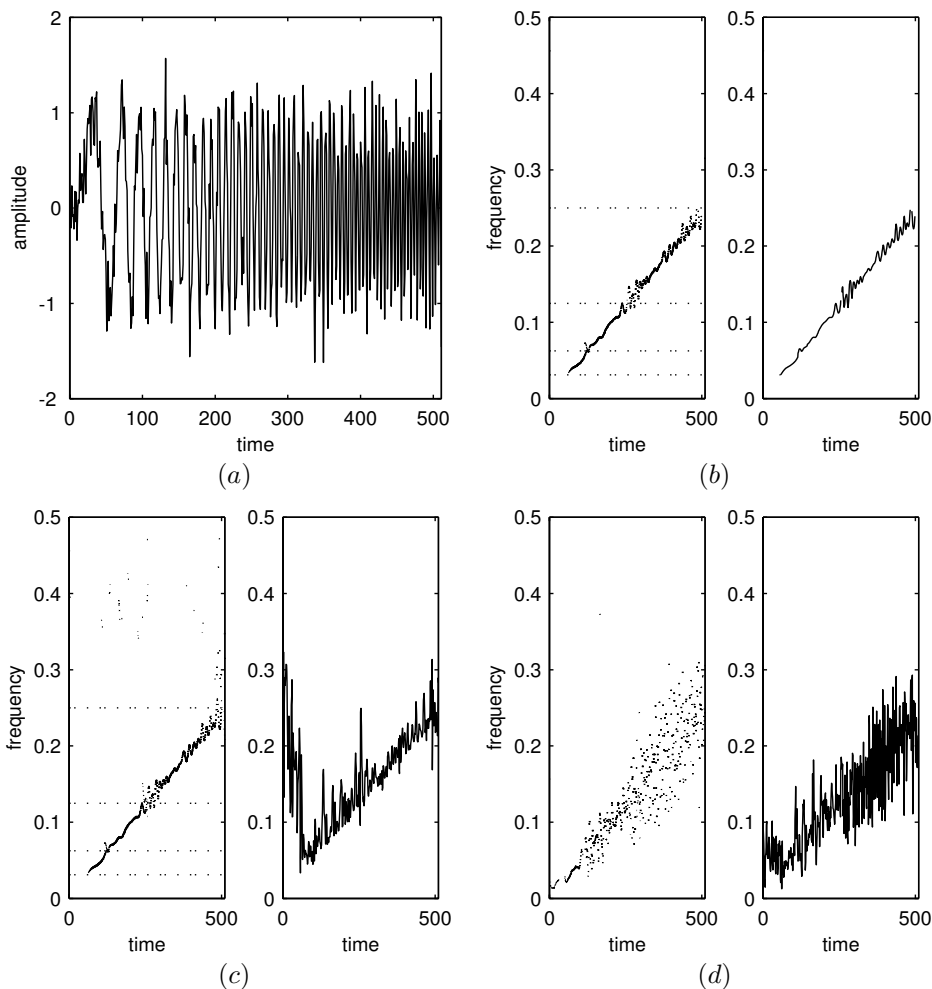
Figure 10. (a) Noisy data. (b) Left: the Hilbert spectrum using the MODWT, Fejér-Korovkin FK(18) wavelet, and $J_0 = 4$, with thresholding of wavelet details. Right: corresponding time-varying mean frequency. (c) As in (b) but with no thresholding. (d) Left: the Hilbert spectrum using EMD and six IMFs. Right: corresponding time-varying mean frequency.

jointly (dependent) Gaussian (Papoulis, 1972, p. 475) and so the random variables η_t and $\eta_t^{(H)}$ are independent and identically distributed for any specified t.

Since from (7.2) $\text{cov}\{\widetilde{\mathcal{E}}_{j,t}, \widetilde{\mathcal{E}}_{j,t'}\}$ depends only on $|t - t'|$ we let $\widetilde{\mathcal{E}}_{j,t}$ play the role of the stationary process η_t and conclude that, up to inaccuracies introduced by filtering end-effects, the random variables $\widetilde{\mathcal{E}}_{j,t}$ and $\widetilde{\mathcal{E}}_{j,t}^{(H)}$ can be treated as independent and identically distributed for each $t = 0, \ldots, N - 1$. In particular $\widetilde{\mathcal{E}}_{j,t}$ and $\widetilde{\mathcal{E}}_{j,t}^{(H)}$ have the same variance given by $\sigma_{\widetilde{\mathcal{E}},j}^2 = \sigma_\epsilon^2/2^j$; see (7.2). To estimate the noise standard deviation σ_ϵ we use the MODWT-based MAD estimator (Percival and

Walden, 2000, p. 429),

$$\hat{\sigma}_\epsilon = \frac{2^{1/2}\text{median}\{|\widetilde{W}_{1,0}|, |\widetilde{W}_{1,1}|, \ldots, |\widetilde{W}_{1,N-1}|\}}{0.6745}, \tag{7.3}$$

which gives an estimate of $\hat{\sigma}^2_{\widetilde{\mathcal{E}},j} = \hat{\sigma}^2_\epsilon/2^j$.

The independence and Gaussianity of $\widetilde{\mathcal{E}}_{j,t}$ and $\widetilde{\mathcal{E}}^{(H)}_{j,t}$ means that

$$\left(\frac{\widetilde{\mathcal{E}}_{j,t}}{\sigma_{\widetilde{\mathcal{E}},j}}\right)^2 + \left(\frac{\widetilde{\mathcal{E}}^{(H)}_{j,t}}{\sigma_{\widetilde{\mathcal{E}},j}}\right)^2 = \left(\frac{|\widetilde{\mathcal{E}}^{(a)}_{j,t}|}{\sigma_{\widetilde{\mathcal{E}},j}}\right)^2 \sim \chi^2_2$$

where '$\sim \chi^2_2$' means 'is distributed as a chi-square random variable with two degrees of freedom,' and $|\widetilde{\mathcal{E}}^{(a)}_{j,t}| = \sqrt{\{\widetilde{\mathcal{E}}^2_{j,t} + [\widetilde{\mathcal{E}}^{(H)}_{j,t}]^2\}}$. We thus have a sequence of dependent chi-square random variables $(|\widetilde{\mathcal{E}}^{(a)}_{j,t}|/\sigma_{\widetilde{\mathcal{E}},j})^2, t = 0, \ldots, N-1$.

Thresholding proceeds under the hypothesis of no deterministic signal being present; in this case $\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{W}}^T_j \widetilde{\mathcal{W}}_j \mathbf{X} = \widetilde{\mathcal{E}}_j = \widetilde{\mathcal{W}}^T_j \widetilde{\mathcal{W}}_j \boldsymbol{\epsilon}$ and hence $A_{j,t} \equiv |\widetilde{\mathcal{E}}^{(a)}_{j,t}|$.

A suitable threshold level for N *independent* chi-square random variables was developed in Sardy (2000), and for our case of N *dependent* chi-square random variables we can show (Olhede and Walden, 2003) that the threshold level of Sardy (2000) provides a conservative threshold level. Under the hypothesis of no deterministic signal present we use the hard-thresholding rule, combined with Sardy's threshold level of $\sqrt{[2\log(N\log N)]}$,

$$A^{(\text{ht})}_{j,t} = \begin{cases} A_{j,t}, & \text{if } A_{j,t} > \hat{\sigma}_{\widetilde{\mathcal{E}},j}\sqrt{[2\log(N\log N)]} \\ 0, & \text{otherwise.} \end{cases}$$

for levels $j = 1, \ldots, J_0$. The Hilbert energy spectrum and time-varying mean frequency calculations then proceed as in (4.6) and (4.7) with $A^{(\text{ht})}_{j,t}$ replacing $A_{j,t}$

Notice that our strategy is to threshold the amplitudes only, leaving the phases unchanged, since only when the amplitude is low do we wish to eliminate the contribution to $\mathbf{X}$. Thresholding the real and imaginary parts of $\widetilde{\mathcal{D}}_{j,t}$ separately would change the phase of the analytic signal and hence is avoided.

(b) Example

Here we consider a linear chirp, $b(t) = \sin(\frac{\pi t^2}{4N})$, plus noise. This signal has instantaneous frequency given by $t/(4N)$, i.e., a straight line with frequency changing from 0 to 0.25 as time goes from 0 to N. We use a discretization of the signal at 512 unit sample intervals, plus a white Gaussian noise $\{\epsilon_t, t = 0, \ldots, 511\}$ with $\sigma_\epsilon = 0.25$. A realization is shown in Fig. 10(a). The MODWT-based Hilbert spectrum and corresponding time-varying mean frequency following the thresholding scheme are given in Fig. 10(b), while the MODWT results omitting thresholding are shown in Fig. 10(c). The EMD results are shown in Fig. 10(d) where the number of IMFs used was six, as found by the algorithm; a reduction to the first three IMFs produced worse results.

Our denoising method, with thresholding of the amplitudes only, works very well. With regard to the Hilbert spectrum, the standard MODWT-based approach again performs better than the EMD, but the omission of thresholding significantly worsens the results.

8. Conclusions

The analysis of a possibly multicomponent non-stationary signal via projections onto the time-frequency plane followed by computation of the Hilbert spectrum is an appealing way to determine the time-variation of frequency components. However, the EMD method seems a problematic step, and we have shown that superior results can be obtained with wavelet-based (MODWT or MODWPT) projections utilising the Fejér-Korovkin class of wavelet filters. These transforms produce decompositions which are conducive to statistical analysis, with the important result that a denoising step can easily be incorporated into this approach.

In the bat data analysis of Fig. 9, we note that although, as required in Hilbert spectrum multicomponent analysis, *at any time* 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. A paper covering work on increasing the flexibility of our approach to deal with such drawbacks and cope better with highly non-linear chirps, via suitable pre- and post-processing, is nearing completion.

Appendix A.

Here we shall prove result (7.2). From Percival and Walden (2000, p. 172),

$$\widetilde{\mathcal{E}}_{j,t} = \frac{1}{N} \sum_{k=0}^{N-1} \left| \widetilde{H}_j\left(\tfrac{k}{N}\right) \right|^2 \Xi_k \mathrm{e}^{\mathrm{i}2\pi tk/N},$$

where $\widetilde{H}_j\left(\tfrac{k}{N}\right) = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l}\mathrm{e}^{-\mathrm{i}2\pi lk/N}$ and $\Xi_k = \sum_{t=0}^{N-1} \epsilon_t \mathrm{e}^{-\mathrm{i}2\pi tk/N}$. Then,

$$E\{\widetilde{\mathcal{E}}_{j,t}\widetilde{\mathcal{E}}_{j,t'}^*\} = E\left\{ \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \left| \widetilde{H}_j\left(\tfrac{k}{N}\right) \right|^2 \left| \widetilde{H}_j\left(\tfrac{l}{N}\right) \right|^2 \Xi_k \Xi_l^* \mathrm{e}^{\mathrm{i}2\pi(tk-t'l)/N} \right\}.$$

But

$$
\begin{aligned}
E\{\Xi_k \Xi_l^*\} &= \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} E\left\{\epsilon_u \epsilon_v^*\right\} \mathrm{e}^{-\mathrm{i}2\pi(uk-vl)/N} \\
&= \sigma_\epsilon^2 \sum_{u=0}^{N-1} \mathrm{e}^{-\mathrm{i}2\pi u(k-l)/N} = \begin{cases} N\sigma_\epsilon^2, & \text{if } k = l, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
E\{\widetilde{\mathcal{E}}_{j,t}\widetilde{\mathcal{E}}_{j,t'}^*\} &= \frac{\sigma_\epsilon^2}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \left|\widetilde{H}_j\left(\tfrac{k}{N}\right)\right|^2 \left|\widetilde{H}_j\left(\tfrac{l}{N}\right)\right|^2 \delta_{k,l} \mathrm{e}^{\mathrm{i}2\pi(tk-t'l)/N} \\
&= \frac{\sigma_\epsilon^2}{N} \sum_{k=0}^{N-1} \left|\widetilde{H}_j\left(\tfrac{k}{N}\right)\right|^4 \mathrm{e}^{\mathrm{i}2\pi k(t-t')/N} \approx \sigma_\epsilon^2 \int_{-1/2}^{1/2} |\widetilde{H}_j(f)|^4 \mathrm{e}^{\mathrm{i}2\pi f(t-t')} df.
\end{aligned}
$$

However,

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

so that

$$
\begin{aligned}
\sigma_\epsilon^2 \int_{-1/2}^{1/2} |\widetilde{H}_j(f)|^4 \mathrm{e}^{\mathrm{i}2\pi f(t-t')} df &\approx \frac{\sigma_\epsilon^2}{\mathrm{i}2\pi(t-t')} \left[\mathrm{e}^{\mathrm{i}2\pi f(t-t')} \Big|_{-1/2^j}^{-1/2^{j+1}} + \mathrm{e}^{\mathrm{i}2\pi f(t-t')} \Big|_{1/2^{j+1}}^{1/2^j} \right] \\
&= \begin{cases} \sigma_\epsilon^2/2^j, & t = t', \\ \sigma_\epsilon^2 \frac{\{\sin[\pi(t-t')/2^{j-1}] - \sin[\pi(t-t')/2^j]\}}{[\pi(t-t')]}, & t \neq t'. \end{cases}
\end{aligned}
$$

and since $E\{\widetilde{\mathcal{E}}_{j,t}\} = 0$, we have $\mathrm{cov}\{\widetilde{\mathcal{E}}_{j,t}, \widetilde{\mathcal{E}}_{j,t'}\} = E\{\widetilde{\mathcal{E}}_{j,t}\widetilde{\mathcal{E}}_{j,t'}^*\}$, and the result is proven.

References

Boashash, B. 1992a Estimating and interpreting the instantaneous frequency of a signal-Part 1: Fundamentals, *Proc. IEEE*, **80**, 520–538.

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

Burden & Faires. 1993 *Numerical analysis, 5th Edition.* Boston MA: PWS Publishing Company.

Coifman, R. R. and Donoho, D. L. 1995 Translation-invariant denoising. In *Wavelets and Statistics (Lecture Notes in Statistics, Volume 103)*, (A. Antoniadis and G. Oppenheim, *eds.*), 125–50, New York: Springer-Verlag.

Cohen, L. 1995 *Time-frequency analysis.* Englewood Cliffs NJ: Prentice-Hall.

Hess-Nielsen, N. & Wickerhauser, M. V. 1996 Wavelets and time-frequency analysis, *Proc. IEEE*, **84**, 523–540.

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, N. E., Shen, Z. & Long, S. R. 1999 A new view of nonlinear water waves: the Hilbert spectrum, *Annu. Rev. Fluid Mech.*, **31**, 417–457.

Liang, J. & Parks, T. W. 1996 A translation-invariant wavelet representation algorithm with applications, *IEEE Trans. Signal Processing*, **44**, 225–232.

Loughlin, P. J. & Tacer, B. 1997 Comments on the interpretation of instantaneous frequency, *IEEE Signal Processing Letters*, **4**, 123–5.

Magrin-Chagnolleau, I. & Baraniuk, R. G. 1999 Empirical mode decomposition based time-frequency attributes. In *Proc. 69th SEG Meeting, Houston.* Soc. Expl. Geophys.

Nielsen, M. 2001 On the construction and frequency localization of finite orthogonal quadrature filters, *J. Approx. Theory*, **108**, 36–52.

Olhede, S. & Walden, A. T. 2003 'Analytic' wavelet thresholding. *Preprint.*

Papoulis, A. 1972 *Probability, random variables, and stochastic processes.* McGraw-Hill.

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

Sardy, S. 2000 Minimax threshold for denoising complex signals with Waveshrink, *IEEE Trans. Signal Processing*, **48**, 1023–8.

Rilling, G., Flandrin, P. & Gonçalvès, P. (2003) On empirical mode decomposition and its algorithms. Presented at *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03*, Grado, Italy, June 8–11.

Shensa, M. J. 1992 The discrete wavelet transform: wedding the à trous and Mallat algorithms, *IEEE Transactions on Signal Processing*, **40**, 2464–82.

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.