Transient Detection in Impulsive Noise Using Low-Variance Spectrum Estimation

Dragana Carevic

Maritime Operations Division, DSTO, Bldg A-51, HMAS Stirling, Rockingham, 6958, Australia

ABSTRACT

This paper considers the problem of detecting unknown bandpass acoustic transients in highly impulsive ambient noise such as that produced by snapping shrimp. Standard transient detection techniques apply the periodogram which is a classical direct nonparametric spectrum estimation method. The periodogram can be limited due to its poor bias properties resulting from substantial sidelobe leakage and is an inconsistent estimator in the sense that its variance does not decrease with the sample size. We propose the use of low-variance spectrum estimation techniques such as multitaper spectrum estimation whereby the power spectrum estimate is obtained as an average of a number of direct spectrum estimates computed using different windows (or tapers). In addition, methods for improving power spectrum estimates based on denoising by wavelet thresholding are considered. Detectors derived based on these approaches are tested using artificially generated bandpass transients inserted in impulsive ambient sea noise.

INTRODUCTION

Detection of passive acoustic transients in an underwater environment is a hard problem as ambient noise may be nonstationary and is usually corrupted by noises produced from biological and man-made sources. In warm shallow underwater environments noise is often dominated by extremely short, highly impulsive snaps produced by snapping shrimp (Versluis et al. 2000). The snaps are broadband with frequencies covering the range from about 600 Hz up to 250 kHz and the superposition of large number of snaps leads to sustained background crackle that is commonly heard in warm coastal waters (Cato and Bell 1992). This paper investigates the detection of bandpass transients of unknown spectral and temporal characteristics in highly impulsive noise, such as that dominated by snapping shrimp, where it is assumed that the time duration of the signal to be detected is much longer than the duration of individual snaps.

Detectors usually exploit the notion that transient signal is a localized burst in time and that its spectrum covers a contiguous frequency band. Signals may also have local oscillatory behavior thereby yielding high spectral peaks. Many current transient detectors apply an appropriate linear transform, such as wavelet (Del Marco and Weiss 1997; Liu and Fraser-Smith 2000; Plett 2007), Gabor (Friedlander and Porat 1989; Friedlander and Porat 1993), or Fourier transform (Nuttall 1996; Nuttall 1997; Streit and Willett 1999), to the received signal. The transform is expected to match the characteristics of the transient and to condense the signal into a few transform coefficients with large magnitude while in the absence of a transient the distribution of the transform coefficients is spread.

Several transient detectors that use discrete Fourier transform (DFT) and discrete wavelet transform (DWT) (S. G. Mallat 1989), and are based on the Nuttall's 'power-law' detector (Nuttall 1996), are described in Wang and Willett 2001. These detectors enable prewhitening (or self-normalisation) of unknown noise levels and allow for a constant false-alarm rate (CFAR) detection. Besides, they are easy to implement and make minimal assumptions about the structure of the transient signal. The use of the DFT in these detectors is equivalent to using the periodogram, which is a classical direct nonparametric method for spectrum estimation of a stationary random time series. It is computed as the magnitude-squared DFT of the (possibly zero-padded) windowed time series, and in Wang and Willett 2001 the window is rectangular.

Although the periodogram provides a fast method to compute an estimate of the power spectrum, it can be limited due to its poor bias properties resulting from substantial sidelobe leakage. This estimator is also inconsistent in the sense that its variance does not decrease with the sample size. In order to tackle this problem a multiple window (or multitaper) spectrum estimation technique is proposed whereby a number of direct spectrum estimates are computed, each by using different window (or taper), and these estimates are averaged (Thompson 1982). The tapers are chosen to be pairwise orthogonal and are designed so as to prevent leakage. This method gives a consistent spectrum estimator with a variance inversely proportional to the number of tapers used (Percival and Walden 1993).

A spectrum estimate obtained either as a periodogram or by using the multitaper method, can be further enhanced by applying denoising techniques based on wavelet thresholding (Percival and Walden 1993; Donoho and Johnstone 1994), (Gao 1997; Moulin 1994; Walden, Percival, and McCoy 1998; Cristan and Walden 2002). This approach utilises the notion that the logarithm of the spectrum estimate of an input time series can be represented as "signal" plus "noise," with the signal being equal to the true log spectrum (Percival and Walden 1993; Gao 1997). Then, a wavelet transform is first applied to the logarithm of the power spectrum estimate and the resulting wavelet coefficients are measured against a threshold; if they fall below this threshold they are counted as noise and set to zero. The smooth estimate of the log spectrum is next obtained by inverting the thresholded wavelet representation. Initially the periodogram is used as the input spectrum estimate (Gao 1997; Moulin 1994) and this technique is improved by replacing the log periodogram with the log of a multitaper spectrum estimator computed using L sine tapers (Walden, Percival, and McCoy 1998; Cristan and Walden 2002).

The paper is organised as follows. Section 2 explains the multitaper spectrum estimation and gives some details about refining spectrum estimates by using wavelet denoising techniques. Section 2 describes approaches to computing detection statistics for transient detection using the estimated spectra. The test results obtained using artificially generated bandpass transients with different bandwidths inserted in impulsive ambient acoustic noise are presented in Section 4. Finally, Section 5 contains some concluding remarks.

LOW-VARIANCE SPECTRUM ESTIMATION

Multitaper Spectrum Estimation

Given the N-dimensional observation vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, a multitaper spectrum estimate of $\mathbf{x}$ is defined by

$$\hat{S}_k(\mathbf{x}) = \frac{1}{L}\sum_{l=0}^{L-1} \hat{S}_{l,k}(\mathbf{x}) \tag{1}$$

with

$$\hat{S}_{l,k}(\mathbf{x}) = \left| \sum_{i=1}^{N} v_{l,i}\, x_i\, e^{-j2\pi(k-1)\,(i-1)/N} \right|^2 \tag{2}$$

where $\mathbf{v}_l = (v_{l,1}, v_{l,2}, \ldots, v_{l,N})$ is the lth data taper used to compute the spectral estimate $\hat{S}_{l,k}(\mathbf{x})$, $l = 0, 1, \ldots, L-1$, k, $1 \le k \le N$ denotes the spectral bin number, and $j = \sqrt{-1}$. The tapers are chosen to be pairwise orthonormal, *i.e.*, $\sum_i v_{l,i} v_{m,i} = 1$ for $m = l$ and $\sum_i v_{l,i} v_{m,i} = 0$ for $m \neq l$. In the special case $L = 1$, the estimator (1) becomes a tapered periodogram estimator, and if the tapering is also uniform (*i.e.*, $v_{1,1} = v_{1,2} = \ldots = v_{1,N} = \frac{1}{N}$) it is called the periodogram estimator. Therefore, for $L > 1$, the estimator (1) is a linear combination of L orthogonal tapered periodogram estimators. The resulting spectrum estimate is superior to the periodogram in terms of the reduced bias and variance.

Thompson 1982 used Slepian or discrete prolate spheroidal sequences (DPSS) as the tapers. DPSS are unique orthogonal sequences that maximise the spectral concentration of the window main lobe within $[-W, W]$, where W is the prescribed main lobe width expressed in units of normalised frequency, $0 < W < 1/2$, and have good leakage properties. The number of tapers L is chosen to be less then $2NW$. Additionally, Riedel and Sidorenko 1995 propose the use of sine tapers defined by

$$v_{l,i} = \sqrt{\frac{2}{N+1}} \sin\frac{\pi(l+1)i}{N+1}, \quad i = 1, 2, \ldots, N.$$

The sine tapers are easy to compute. They also produce smaller local bias than the DPSS and have roughly same spectral concentration.

Smoothing the Spectrum Estimate by Wavelet Thresholding

Assuming that the eigenspectra in (2) are uncorrelated, the ratio of the estimated multitaper spectrum $\hat{S}_k(\mathbf{x})$ to the true power spectrum $S_k(\mathbf{x})$ can be approximated by a chi-squared distribution with $2L$ degrees of freedom, *i.e.*, as (Percival and Walden 1993)

$$v_k = \frac{\hat{S}_k(\mathbf{x})}{S_k(\mathbf{x})} \sim \frac{\chi^2_{2L}}{2L}. \tag{3}$$

Then the random variable (rv) η_k defined by

$$\eta_k = \log v_k - \psi(L) + \log L \tag{4}$$

is approximately correlated zero mean Gaussian with known variance $\sigma_\eta^2 = \psi'(L)$, where $\psi(\cdot)$ and $\psi'(\cdot)$ denote respectively digamma and trigamma functions. Applying logarithm to (3) and using (4) we have

$$\log \hat{S}_k(\mathbf{x}) - \psi(L) + \log L = \log S_k(\mathbf{x}) + \eta_k \tag{5}$$

that is, the log multitaper spectrum (plus a constant $(\log L - \psi(L))$) can be written as the true log spectrum plus approximately correlated Gaussian noise with zero mean and known variance σ_η^2. The model in (5) is well suited for wavelet denoising with a goal of removing the noise term η_k and obtaining the smooth estimate of the log spectrum.

A number of methods have been proposed for signal denoising based on thresholding wavelet coefficients (Donoho and Johnstone 1994; Donoho 1995; Donoho and Johnstone 1995; Krim et al. 1999; Johnstone and Silverman 1997). We briefly describe only those that are relevant for this paper. Denote by $\{z_{j,m}\}$ a set of wavelet coefficients computed by applying a DWT to the signal representation on the left side of (5) where subscript j indicates decomposition level (or scale) and m, $m = 1, \ldots, N/2^j$ is the wavelet coefficient index associated with the scale j. By the linearity of the DWT it follows from (5) that

$$z_{j,m} = s_{j,m} + n_{j,m} \tag{6}$$

where $\{s_{j,m}\}$ are wavelet coefficients of $\log S_k(\mathbf{x})$ and $\{n_{j,m}\}$ are wavelet coefficients of η_k.

Hard and soft thresholding functions can be used for noise reduction (Donoho 1995). The hard thresholding function, for a given threshold T, is defined by

$$\delta_h = \begin{cases} z, & \text{if } |z| \ge T \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

and the soft thresholding function is defined by

$$\delta_s = \begin{cases} z - T, & \text{if } z \ge T \\ 0, & \text{if } |z| < T \\ z + T, & \text{if } z \le -T. \end{cases} \tag{8}$$

For either soft or hard thresholding it is important to determine an appropriate threshold level T. A simple approach is to use scale-independent thresholding whereby one threshold is applied to all wavelet coefficients independent of the scale j (Walden, Percival, and McCoy 1998). This method uses the assumption that noise is Gaussian distributed and uncorrelated. The threshold does not depend on the input data, but only on the noise variance σ_η^2 and is given by $T = \sigma_\eta\sqrt{2\log N}$.

However, the noise term η_k in (4) is stationary and coloured, so the variance of the corresponding noise wavelet coefficients $n_{j,m}$ depends on level (or scale) j of the wavelet decomposition. Walden, Percival, and McCoy 1998 investigated the correlation structure of η_k across frequencies and devised a method to estimate the level-dependent variances of the wavelet coefficients σ_j^2. In this level-dependent thresholding scheme the threshold T at each scale j of the DWT is computed as $T = \sigma_j\sqrt{2\log N}$.

Another thresholding scheme using the principle of minimising Stien's unbiased estimate of risk (SURE) is proposed by Donoho and Johnstone 1995 and Johnstone and Silverman 1997. For a specified threshold T and the signal $\mathbf{z} = \{z_m\}_{m=1}^N$, the Stien's unbiased estimate of risk using the soft thresholding function is given by

$$\hat{R}(T;\mathbf{z}) = \sigma^2 N + \sum_{i=1}^{N} \left\{ \min(z_i^2, T^2) - 2\sigma^2 I(|z_m| \le T) \right\} \tag{9}$$

where σ^2 is noise variance and I is an indicator function, $I(\cdot) = 1$ if $|z_m| \le T$ and $I(\cdot) = 0$ if $|z_m| > T$. A SURE threshold is then selected as

$$T_{SURE} = \arg\min_{0 \le T \le \sigma\sqrt{2\log N}} \hat{R}(T;\mathbf{z}). \tag{10}$$

In the cases where the data is dominated by noise, the SURE threshold is found to be too small. A hybrid thresholding scheme is proposed as a combination of the SURE thresholding and a fixed thresholding method (Donoho and Johnstone 1995). This scheme decides whether to use the SURE threshold or the fixed threshold $\sigma\sqrt{2\log N}$ based on the test of the significance of the

presence of the signal in the data. In particular, the threshold is computed by

$$T = \begin{cases} \sigma\sqrt{2\log N}, & s_d^2 \le \eta_d \\ T_{SURE}, & s_d^2 > \eta_d. \end{cases} \quad (11)$$

where $s_d^2 = N^{-1}\sum_{i=0}^{N-1} x_i^2 - \sigma^2$ and $\eta_d = \sigma(\log_2 N)^{1.5}/\sqrt{N}$.

A level-dependent hybrid thresholding that assigns a threshold to each dyadic level j of the DWT can be devised by using level-dependent variances σ_j in (10) and (11). These variances are computed as

$$\sigma_j = \frac{\mathrm{MAD}(z_{j,m})}{0.6745} \quad (12)$$

where $\mathrm{MAD}(z_{j,m})$ is the median absolute deviation of the wavelet coefficients at the decomposition level j, and 0.6745 is a normalising factor (Johnstone and Silverman 1997).

Denote the spectrum refined by wavelet denoising as $\mathscr{W}(\hat{S}_k(\mathbf{x}))$. Given the spectrum estimate $\hat{S}_k(\mathbf{x})$ the procedure for computing $\mathscr{W}(\hat{S}_k(\mathbf{x}))$ is summarised by the following steps:

1. Obtain $Z_k(\mathbf{x}) = \log\hat{S}_k(\mathbf{x}) - \psi(L) + \log L$ where L is the number of tapers used to compute the multitaper spectrum estimate $\hat{S}_k(\mathbf{x})$.
2. Apply a standard periodic DWT (S. G. Mallat 1989) to $Z_k(\mathbf{x})$ to obtain the wavelet coefficients $z_{j,k}$ at each level j.
3. Apply thresholding to the wavelet coefficients up to the decomposition level q_0 keeping the wavelet coefficients at the levels above q_0 and the scaling coefficient unchanged. The thresholding method is either scale-independent or scale-dependent as decribed above.
4. Apply the inverse DWT to the thresholded coefficients to obtain the refined log spectrum.
5. Compute the estimate $\mathscr{W}(\hat{S}_k(\mathbf{x}))$ by applying the exponential function to the estimated log spectrum.

NON-GAUSSIAN TRANSIENT DETECTOR STRUCTURE

The transient detection process is formalised in terms of binary hypothesis testing. Consider an N-dimensional complex-valued observation vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ where the ith complex sample x_i is given by

$$x_i = x_i^I + jx_i^Q. \quad (13)$$

and where the superscripts I and Q indicate in-phase and quadrature components of the data respectively. We test a hypothesis H_1 that the observation vector $\mathbf{x}$ contains a transient signal embedded in additive noise against a null hypothesis H_0 that the data contains only a realisation of a wide sense stationary noise random process. The transient is taken to be entirely contained within the window of N data samples, and noise is assumed to be non-Gaussian and impulsive. Also, it is assumed that bandwith and centre frequency, duration and time of arrival of the transient are not known.

Usually, the detection of signals in impulsive noise involves initially passing the data through a non-linearity with a goal of suppressing samples with large instantaneous envelope. Locally optimum detectors use parametric non-linearities that require detailed knowledge of the noise statistics (Chitre, Potter, and Ong 2006). As an approximation to parametric non-linearities Lu and Eisenstein 1983 propose the use of non-parametric, suboptimal, non-linearity that does not require the knowledge of the noise statistics. They define this non-linearity by

$$g(y_i) = \begin{cases} \frac{1}{y_i}, & y_i > 0 \\ 0, & y_i = 0 \end{cases} \quad (14)$$

where y_i is the instantaneous envelope of the complex sample x_i

$$y_i = \sqrt{(x_i^I)^2 + (x_i^Q)^2}. \quad (15)$$

and use it in the detection of band-pass and narrow-band signals in highly impulsive noises.

Using the non-linearity in (14) we define the modified (preprocessed) observation vector $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N)$ as

$$\tilde{x}_i = \left(\sum_{i=1}^{N} y_i^{\eta}\right) g(y_i) x_i, \quad i = 1, 2, \ldots, N \quad (16)$$

where $\sum_{i=1}^{N} y_i^{\eta}$ is a weighting factor and η is an appropriately chosen constant. The use of the weighting factor in (16) improves the detection performance for transients with larger bandwidths at high SNR's. Since, again, we wish to compress the samples with large instantaneous envelope that are due to snapping shrimp noise, we set $\eta < 1$.

The transient detection process involves a comparison of the observation vector $\mathbf{x}$ to a complex-valued noise-only vector $\mathbf{n}$ (see (Plett 2007) and (Wang and Willett 2001)). The non-overlapping data block of the same length as the input signal $\mathbf{x}$ that immediately precedes $\mathbf{x}$ in time is assumed to be transient-free and is used as an estimate of the noise vector $\mathbf{n}$. The following 3 functions are used to define the transient detectors tested in this paper

$$f_{1,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}) = \frac{\hat{S}_k(\tilde{\mathbf{x}}) - \hat{S}_k(\tilde{\mathbf{n}})}{\hat{S}_k(\tilde{\mathbf{n}})} = \frac{\hat{S}_k(\tilde{\mathbf{x}})}{\hat{S}_k(\tilde{\mathbf{n}})} - 1 \quad (17)$$

$$f_{2,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}) = \frac{\mathscr{W}(\hat{S}_k(\tilde{\mathbf{x}}))}{\mathscr{W}(\hat{S}_k(\tilde{\mathbf{n}}))} - 1 \quad (18)$$

$$f_{3,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}) = \mathscr{W}\left(\frac{\hat{S}_k(\tilde{\mathbf{x}})}{\hat{S}_k(\tilde{\mathbf{n}})}\right) - 1 \quad (19)$$

where $\hat{S}_k(\cdot)$ is a power spectrum estimate, $\mathscr{W}(\hat{S}_k(\cdot))$ is the spectrum estimate refined by denoising based on wavelet thresholding, and $\tilde{\mathbf{n}}$ represents the result of processing the noise vector $\mathbf{n}$ using (16).

In (19) denoising by wavelet thresholding is applied to the spectrum estimates related to signal-plus-noise and noise-only time series $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{n}}$ separately, and the ratio of the resulting refined spectra is then obtained. By contrast in (20) wavelet denoising is applied to the ratio of the two estimated spectra $\hat{S}_k(\tilde{\mathbf{x}})/\hat{S}_k(\tilde{\mathbf{n}})$. Hu and Loizou 2004 showed that the logarithm of the ratio of the estimated signal-plus-noise and noise-only spectra $\hat{S}_k(\tilde{\mathbf{x}})/\hat{S}_k(\tilde{\mathbf{n}})$ can be modelled as the logarithm of the ratio of the true spectra $S_k(\tilde{\mathbf{x}})/S_k(\tilde{\mathbf{n}})$ plus a Gaussian distributed noise with the variance $\sigma_\eta^2 = 2\psi'(L)$. Therefore the procedure for spectrum enhancement by wavelet thresholding given by steps 1-5 in Section 2.2 can also be applied to the spectral ratio $\hat{S}_k(\tilde{\mathbf{x}})/\hat{S}_k(\tilde{\mathbf{n}})$.

The functions $f_{r,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}})$, $r = \{1, 2, 3\}$, in (18)-(20) can have negative values, in particular in the noise-only spectral bins related to the input signal $\mathbf{x}$. Therefore, these functions are further processed as

$$f_{r,k}^{\{P\}}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}) = \begin{cases} f_{r,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}), & f_{r,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}) > 0 \\ 0, & f_{r,k}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}}) \le 0. \end{cases} \quad (20)$$

Finally, using $f_{r,k}^{\{P\}}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}})$ in (20) the detection statistics $D(\mathbf{x})$ is then defined by

$$D(\mathbf{x}) = \sum_{k=K_{min}}^{K_{max}} \left(f_{r,k}^{\{P\}}(\tilde{\mathbf{x}}, \tilde{\mathbf{n}})\right)^{\nu} \quad (21)$$

where $r \in \{1,2,3\}$, K_{min} and K_{max}, $K_{max} > K_{min}$, $K_{min} \geq 1$, $K_{max} \leq N$, are the spectral bin numbers that correspond to the lowest and the highest frequency of interest in the power spectrum, and the exponent ν is real and its value is chosen based on the characteristics of the signal to be detected (Nuttall 1996; Wang and Willett 2001). For the band-pass transients considered in this paper $1.5 \geq \nu \geq 3$.

DETECTOR PERFORMANCE

Based on the functions $f_{r,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$, $r = \{1,2,3\}$ in (18)-(20), we formulate 8 detection statistics that use different preprocessing, spectrum estimation, and wavelet denoising techniques. Four detectors denoted by PMT, PPER, MT and PER are based on the function $f_{1,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ in (18). The PMT detector is computed using multitaper spectrum estimation with L tapers and the input signals $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{n}}$ are preprocessed using (16). The PPER detector also applies preprocessed input signals, and computes the spectra using a standard tapered periodogram. The detectors denoted by MT and PER apply multitaper spectrum estimates and the periodograms, respectively, but utilise unprocessed input signals $\mathbf{x}$ and $\mathbf{n}$. The remaining 4 detectors are based on the functions $f_{2,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ and $f_{3,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$. They use exclusively multitaper spectrum estimates of the preprocessed input signals and apply denoising by wavelet thresholding. The detectors based on the function $f_{2,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ are designated by W1-PMT and W2-PMT. The detector W1-PMT applies soft scale-independent wavelet thresholding whereas W2-PMT is computed using scale-dependent thresholding based on the SURE principle. The detectors based on the function $f_{3,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ are denoted by W3-PMT and W4-PMT; they also use, respectively, the same two denoising techniques.

To test the performance of the detectors we use a 27 minutes long set of ambient noise data collected off the cost of Western Australia. The sampling rate is 192 kHz, the water depth is 10 meters and the seabed is sandy/muddy. The intense sustained cracking noise characteristic of snapping shrimp can be heard when listening to the data on audio.

The spectral pass-band chosen for the analysis is centered at 6100 Hz and has bandwith 10200 Hz. The raw data received by the sensor is initially complex demodulated by shifting the centre of the pass-band to zero frequency, low-pass filtering and then decimating appropriately. The resulting sampling rate after the decimation is 12 kHz. A ten seconds long segment of in-phase complex demodulated time-series is shown in Fig. 1(a). The values of the kurtosis computed using in-phase and quadrature time-series within consecutive 2 seconds long segments of the data set, where in-phase and quadrature components are taken to be independent and identically distributed, are shown in Fig. 1(b).

The simulated transients used for testing are generated as random band-pass Gaussian signals of length 0.4 s for a number of signal bandwidths. The transient center frequencies are selected randomly in such a way that the signals fall entirely within the pass-band $[1100, 10000]$ Hz. The ambient noise data is divided in two halves and in-phase and quadrature time-series from both halves are further divided into non-overlapping segments of length 0.7 s. The test transient signals are added to the segments from the second half of the data set at a range of SNR's. The signal variance σ_t^2 is chosen so that

$$SNR = 10\log_{10}\frac{N_t\,\sigma_t^2}{N\,\sigma_n^2} \tag{22}$$

where N_t denotes transient length, N is segment length, and σ_n^2 is noise variance estimated using data samples from the current segment. The transient starting times are selected randomly so that the signals are entirely contained within the corresponding

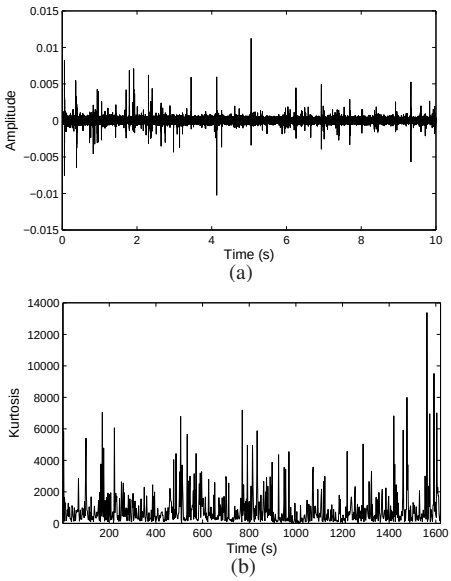


Figure 1: (a) A ten seconds long segment of in-phase time series for the complex demodulated data; (b) kurtosis computed using in-phase and quadrature time-series within consecutive 2 second long segments.

segments. The segments from the first half of the data are taken to be noise-only (or transient-free).

The detection statistics described above are computed for each data segment and SNR, where the segments are taken to be the observation vectors $\mathbf{x}$. The noise-only (transient-free) segment that immediately precedes the current segment in the data set is used as the estimate of the noise vector $\mathbf{n}$. The detection statistics computed using the segments from the first half of the data are utilised to determine probabilities of false alarm P_{fa} for a number of detection thresholds. These thresholds are also applied to the detection statistics computed using the segments from the second half of the data to obtain the probabilities of detection P_d for the required values of P_{fa} at different SNR's.

Since for denoising by wavelet thresholding the estimated power spectra need to be processed using a standard periodic DWT, it is required that the length of the spectrum be dyadic. In a general case the length of the underlying signals $\mathbf{x}$ and $\mathbf{n}$ is not dyadic, so these vectors need to be appropriately zero-padded to attain the dyadic length of $N_D = 2^p$, where p is chosen such that $2^{p-1} < N \leq 2^p$. The multitaper power spectra are computed using the zero-padded time series (Walden, Percival, and McCoy 1998) and the periodograms are computed using the input signals of the original length and the Hanning window. For each spectrum the spectral components from two neigbouring spectral bins are summed thus exploiting frequency contiguity of transient signals (Plett 2007; Wang and Willett 2001). The resulting spectra are further processed as described in Section 3 and used to compute detection statistics. The thresholds for wavelet thresholding are computed as $T = \sigma\sqrt{2\log N_D}$ where the noise variance σ^2 depends on the particular thresholding method used (see Section 2.2). The DWT is computed using Symmlet wavelet of order 4 (symmlet4) (S. Mallat 2001) and the decomposition level to which the thresholding is applied is $q_0 = 4$. The summation in (21) is done over the spectral bin numbers between K_{min} and K_{max}, where K_{min} and K_{max} are chosen so as to correspond to the frequencies 1100 Hz and 10000 Hz, respectively. We also set $\eta = 0.5$ in (16).

Figs. 2(a)-(d) show examples of the functions $f_{r,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$, $r = \{1,2,3\}$ in (18)-(20), where the observation vector contains a

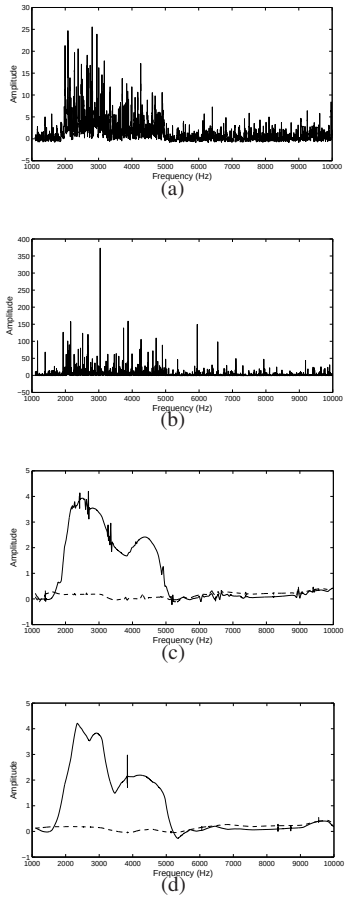


Figure 2: Examples of functions $f_{r,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$, $r=\{1,2,3\}$, for an observation vector $\mathbf{x}$ that contains a transient signal with bandwith 3000 Hz and cetre frequency 3457 Hz at SNR -6 dB (a) $f_{1,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ computed using multitaper spectrum estimates with $L=6$ tapers (b) $f_{1,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ computed using periodograms with Hanning taper (c) $f_{2,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ computed using multitaper spectrum estimates with $L=6$ tapers and spectrum denoising with scale-independent wavelet thresholding for signal-plus-noise vector $\mathbf{x}$ (solid line) and noise-only vector $\mathbf{x}$ (dashed line) (d) $f_{3,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ computed using multitaper spectrum estimates with $L=6$ tapers and spectrum denoising with scale-dependent wavelet thresholding for signal-plus-noise vector $\mathbf{x}$ (solid line) and noise-only vector $\mathbf{x}$ (dashed line).

transient signal with bandwith $B_T = 3000$ Hz and centre frequency $CF_T = 3457$ Hz inserted into a segment of ambient noise at the SNR of -6 dB. In all examples both observation and noise vectors are preprocessed using (16). Figs. 2(a) and (b) show function $f_{1,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ computed, respectively, using multitaper spectra estimated with $L=6$ tapers and the periodograms computed with Hanning window. Figs. 2(c) and (d) show the results of applying denoising by wavelet thresholding to the multitaper spectra estimated using $L=6$ tapers. Fig 2(c) shows function $f_{2,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ computed using scale-independent wavelet thresholding and Fig 2(d) shows function $f_{3,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ processed by level dependent thresholding based on the SURE principle. Also, in Figs. 2(c) and (d) dashed curves represent functions $f_{2,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$ and $f_{3,k}(\tilde{\mathbf{x}},\tilde{\mathbf{n}})$, respectively, computed using noise-only (transient-free) observation vector $\mathbf{x}$.

The performance of the detectors depends on a number of parameters, such is the number of tapers used to compute multitaper spectrum estimates and the parameter ν in (21), and, also, on the bandwidth of transient signals. In the following we present the results that show how detector performance depends on these parameters.

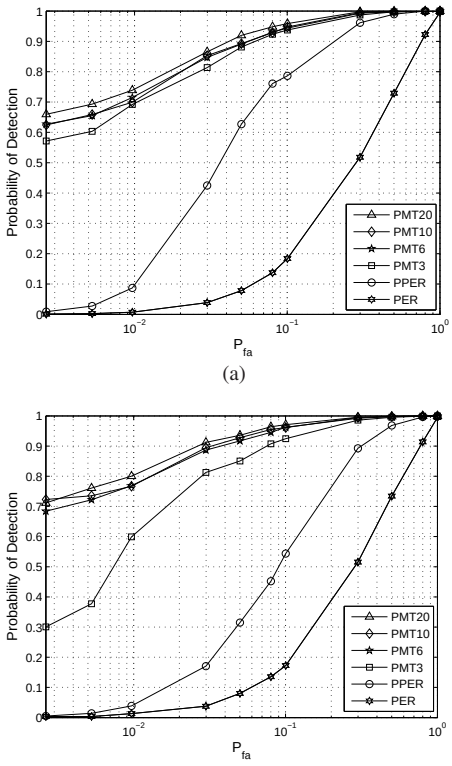
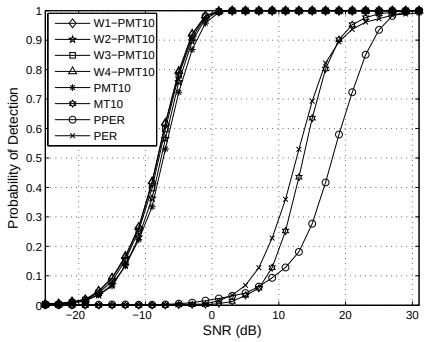


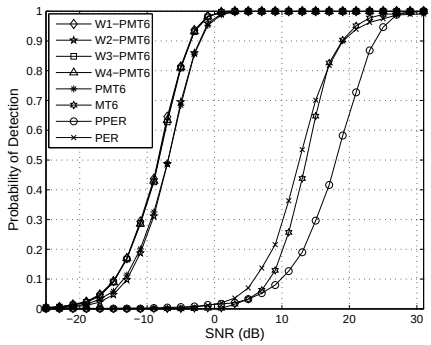
Figure 3: Probability of detection as a function of P_{fa} for SNR of -5 dB for PMT detectors where multitaper spectrum estimates are obtained using $L=\{20,10,6,3\}$ tapers. The transients used in the tests have bandwith $B_T = 3000$ Hz. Detection statistics are computed by setting ν in (21) to (a) $\nu = 1.6$ and (b) $\nu = 2.5$.

Figs. 3(a) and (b) show the probability of detection P_d as a function of P_{fa} for SNR of -5 dB for PMT detectors where the parameter ν takes, respectively, values $\nu = 1.6$ and $\nu = 2.5$. Multitaper spectrum estimates are obtained using $L=\{20,10,6,3\}$ tapers and the corresponding detectors are denoted by PMT20, PMT10, PMT6, and PMT3. The transients used in the tests have bandwith $B_T = 3000$ Hz and centre frequency is randomly selected. For comparison Figs. 3(a) and (b) also show test results for PPER detector computed using periodograms of the preprocessed vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{n}}$ and for PER detector computed using periodograms of the unprocessed signals. It can be seen that PMT detectors perform much better than PER and PPER detectors in particular at lower SNR's. Also, the performance of PMT detectors for $L \geq 6$ is somewhat better for $\nu = 2.5$ then for $\nu = 1.6$, whereas the performance of other detectors deteriorates for the larger value of this parameter.

Figs. 4 and 5 show the probability of detection P_d as a function of SNR for all 8 detectors where the detection statistics that depend on multitaper spectrum estimation are computed using different number of tapers and the parameter $\nu = 2.5$. The probability of false alarm is $P_{fa} = 0.003$ and the test transient signals are the same as in Fig. 3. In Figs. 4(a) and (b), where $L=\{10,6\}$ respectively, it can be seen that the performance of the detectors based on wavelet thresholding is comparable to that obtained using PMT detectors. In Figs. 5(a) and (b), where $L=\{3,1\}$, wavelet-based detectors give much better results than the corresponding PMT detectors. However, the performance of all wavelet-based detectors for different L in Fig. 4 and 5 is very similar even in the case where $L=1$ (note that the

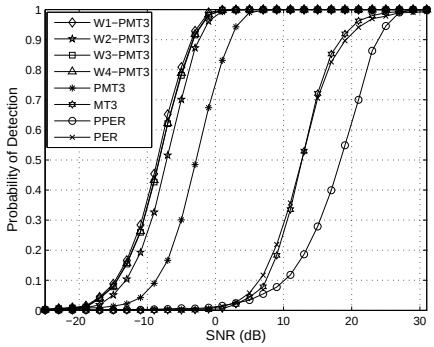


(a)

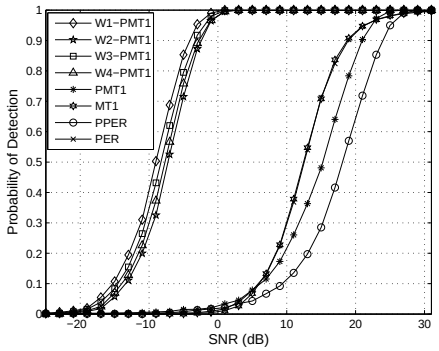


(b)

Figure 4: Probability of detection as function of SNR for $P_{fa} = 0.003$ for all detectors. Bandwidth of transients used in the tests is $B_T = 3000$ Hz and $\nu = 2.5$. Detectors based on multitaper spectrum estimation use different number of tapers: (a) $L = 10$ tapers (b) $L = 6$ tapers.



(a)



(b)

Figure 5: Probability of detection as function of SNR for $P_{fa} = 0.003$ for all detectors. Bandwidth of transients used in the tests is $B_T = 3000$ Hz and $\nu = 2.5$. Detectors based on multitaper spectrum estimation use different number of tapers: (a) $L = 3$ tapers (b) $L = 1$ tapers.

spectral estimates obtained for $L = 1$ are periodograms). So, it can be concluded that the performance of detectors that use wavelet thresholding does not depend on the number of tapers L used. The performance of MT, PPER and PER detectors is much worse as compared to other detectors.

We also tested the proposed detectors using transient signals with different bandwidths. Figs. 5(a) and (b) show the probability of detection P_d as a function of P_{fa} for fixed SNR of -11 dB for all 8 detectors where the bandwidths of the test transients are $B_T = \{200, 1000\}$ Hz. Figs. 6(a) and (b) show the probability of detection P_d as a function of P_{fa} for fixed SNR of -5 dB for all 8 detectors where the bandwidths of the test transients are $B_T = \{3000, 5000\}$ Hz. In both Figs. 5 and 6 the number of tapers used to compute detection statistics based on multitaper spectrum estimation is $L = 6$ and the parameter $\nu = 2.5$. It can be seen that the detectors that use multitaper spectrum estimation and the input signal preprocessed using (16) perform well for all signals bandwidths. The performance of these detectors is also much better than the performance of other detectors. For the transient signals with a smaller bandwidth of $B_T = 200$ Hz in Fig. 5(a) PMT6 detector performs as well as W1-PMT6 detector and better then other wavelet-based detectors. For the transient signals with larger bandwidths $B_T = \{1000, 3000, 5000\}$ Hz in Fig. 5(b) and Figs. 6(a) and (b) the performance of detectors based on wavelet thresholding is better than that of PMT6 detector. Generally, good performance of the detectors for the transients with larger bandwidths and at high SNR's is achieved by using the weighting factor in (17).

CONCLUSION

This paper considered the problem of detecting unknown bandpass transients buried in impulsive non-Gaussian noise. The proposed approaches apply low-variance spectrum estimation techniques. One of these techniques is Thomson's multitaper spectrum estimation whereby the power spectrum is estimated as an average of a number of direct spectrum estimates computed using different tapers. This is a consistent spectrum estimator with a variance inversely proportional to the used number of tapers. The other technique involves spectrum enhancement that applies spectrum denoising by wavelet thresholding. Several detectors have been derived based on these two methods where the input signals are preprocessed using a memoryless non-linear filter. The detectors are tested using artificially generated bandpass transient signals inserted in impulsive ambient sea noise.

The detectors that use multitaper spectrum estimation computed with the preprocessed input signals are found to perform better than the detectors based on periodograms and those that use multitaper spactrum estimation with unprocessed (raw) input signals when the number of tapers used to compute spectrum estimates is $L \geq 6$. However, the performance of multitaper detectors deteriorates as the number of tapers decreases. The detectors based on wavelet thresholding are found to perform comparably to the detectors that utilise multitaper spectrum estimates computed using $L \geq 6$ tapers. These detectors also give good results when the underlying spectrum estimates are obtained using only one taper, which is equivalent to a periodogram. It is therefore concluded that the performance of the detectors that apply denoising by wavelet thresholding does not depend on the number of tapers used to obtain the spectrum estimates, and can be applied directly to the periodograms.

Additionally, the proposed detectors that utilise low-variance spectrum estimators of the preprocessed input signals are tested using the transients with different bandwidths. It is concluded that they perform well over a wide range of signal bandwidths

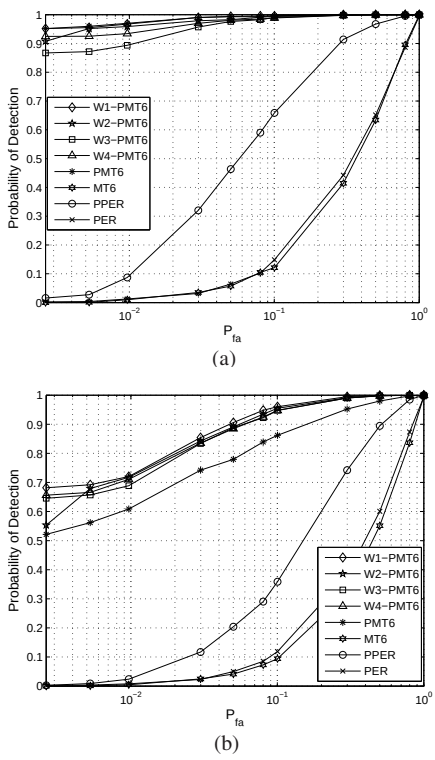


Figure 6: Probability of detection as function of P_{fa} computed for all detectors for a fixed SNR of -11 dB. Detectors based on multitaper spectrum estimation use $L = 6$ tapers and $\nu = 2.5$. Bandwidths of transients used in the tests and fixed SNR's are: (a) $B_T = 200$ Hz (b) $B_T = 1000$ Hz.

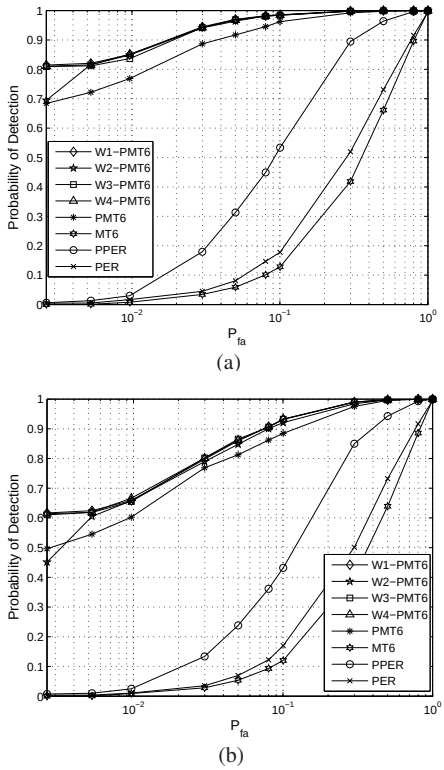


Figure 7: Probability of detection as function of P_{fa} computed for all detectors for a fixed SNR of -5 dB. Detectors based on multitaper spectrum estimation use $L = 6$ tapers and $\nu = 2.5$. Bandwidths of transients used in the tests and fixed SNR's are: (a) $B_T = 3000$ Hz (b) $B_T = 5000$ Hz.

and that their performance is better as compared to the detectors that utilise only periodograms and unprocessed input signals.

REFERENCES

Cato, D. H. and M. J. Bell (1992). *Ultrasonic ambient noise in Australian shallow waters at frequencies up to 2000 kHz*. Tech. rep. MRL-TR-91-23. Materials Research Laboratory, Australia.

Chitre, M. A., J. R. Potter, and S. H. Ong (2006). "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise". *IEEE J. Oceanic Eng.* 31.2, pp. 497–503.

Cristan, A. C. and A. T. Walden (2002). "Multitaper Power spectrum estimation and thresholding: Wavelet packets versus wavelets". *IEEE Trans. Signal Process.* 50.12, pp. 2976–2986.

Del Marco, S. and J. Weiss (1997). "Improved transient signal detection using a wavepacket-based detector with an extended translation-invariant wavelet transform". *IEEE Trans. Signal Process.* 45.4, pp. 841–850.

Donoho, D. L. (1995). "De-noising by soft thresholding". *IEEE Trans. Inform. Theory* 41.5, pp. 613–627.

Donoho, D. L. and I. M. Johnstone (1994). "Ideal spatial adaptation by wavelet shrinkage". *Biometrika* 81, pp. 425–255.

Donoho, D. L. and M. Johnstone (1995). "Adapting to unknown smoothness via wavelet srinkage". *J. Amer. Stat. Assoc.* 90, pp. 1200–1224.

Friedlander, B. and B. Porat (1989). "Detection of transient signals by the Gabor representation". *IEEE Trans. Acoustic Speech Sig. Process.* 37.2, pp. 169–180.

— (1993). "Performance analysis of transient detectors based on a class of linear transforms". *IEEE Trans. Infrom. Theory* 40.2, pp. 136–144.

Gao, H.-Y (1997). "Choice of thresholds for wavelet shrinkage estimate of the spectrum". *J. Time Series Anal.* 18, pp. 231–251.

Hu, Y. and P. C. Loizou (2004). "Speech enhancment based on wavelet thresholding the multitaper spectrum". *IEEE Trans. Speech Audio Processing* 12.1, pp. 59–67.

Johnstone, I. M. and B. W. Silverman (1997). "Wavelet threshold estimators for data with correlated noise". *J. R. Stat. Soc* B-59, pp. 319–51.

Krim, H. et al. (1999). "On denoising and best signal representation". *IEEE Trans. Info. Theory* 45.7, pp. 2225–2238.

Liu, T. T. and A. C. Fraser-Smith (2000). "Detection of transients in 1/f noise with the undecimated discrete wavelet transform". *IEEE Trans. Signal Process.* 48.5, pp. 1458–1462.

Lu, N. H. and B. A. Eisenstein (1983). "Suboptimum detection of weak signals in non-Gaussian noise". *IEEE Trans. Information Theory.* 29.3, pp. 462–466.

Mallat, S. G. (1989). "A theory of multiresolution signal decomposition: the wavelet representation". *IEEE Trans. Pattern Anal. Machine Intell.* 11, pp. 674–693.

Mallat, S. (2001). *A wavelet tour of signal processing*. Washingtn, D.C.,Chestnut Hill, MA 02176, USA: Academic Press.

Moulin, P. (1994). "Wavelet thresholding techniques for power spectrum estimation". *IEEE Trans. Signal Process.* 42.11, pp. 3126–3136.

Nuttall, A. (1996). *Near-optimum detection performance of power-low processors for random signals of unknown location, structure, extend and arbutrary setrength*. Tech. rep. 11123. Naval Undersea Warfare Center Division, Newport, Rhode Island.

— (1997). *Performance of power-low processors with normaliztion for random signals of unknown structure*. Tech. rep.

10760. Naval Undersea Warfare Center, Newport, Rhode Island.

Percival, D. B. and A. T. Walden (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. The Edinburgh Building, Cambridge CB2 2RU, UK: Cambridge Univ. Press.

Plett, M. I. (2007). "Transient detection with cross wavelet transforms and wavelet coherence". *IEEE Trans. Signal Process.* 558.5, pp. 1605–1611.

Riedel, K. S. and A. Sidorenko (1995). "Minimum bias multiple taper spectral estimation". *IEEE Trans. Signal Process.* 43.1, pp. 188–195.

Streit, R. and P. Willett (1999). "Detection of random transient signals via hyperparameter estimation". *IEEE Trans. Signal Process.* 47.7, pp. 1823–1834.

Thompson, D. J. (1982). "Spectrum Estimation and and harmonic analysis". *Proc. IEEE* 70.9, pp. 1055–1095.

Versluis, M. et al. (2000). "How snapping shrimp snap: through cavitating bubbles". *Science* 289, pp. 2114–2117.

Walden, A. T., D. B. Percival, and E. J. McCoy (1998). "Spectrum estimation by wavelet thresholding multitaper estimators". *IEEE Trans. Signal Process.* 46.12, pp. 3153–3165.

Wang, Z. and P. Willett (2001). "All-purpose and plug-in power-law detectors for transient signals". *IEEE Trans. Signal Process.* 49.11, pp. 2454–2466.