MULTIPLE WINDOW TIME-FREQUENCY ANALYSIS

Metin Bayram and Richard G. Baraniuk *

Department of Electrical and Computer Engineering
Rice University
Houston, TX 77005–1892, USA
mebay@rice.edu, richb@rice.edu

ABSTRACT

We propose a robust method for estimating the time-varying spectrum of a non-stationary random process. Our approach extends Thomson's powerful multiple window spectrum estimation scheme to the time-frequency and time-scale planes. The method refines previous extensions of Thomson's method through optimally concentrated window and wavelet functions and a statistical test for extracting chirping line components.

1. INTRODUCTION

Many methods exist for estimating the power spectrum of stationary signals. However, these methods are insufficient for the non-stationary signals that occur in important applications such as radar, sonar, acoustics, biology, and geophysics. These applications demand time-frequency representations that indicate how the power spectrum changes over time. To date research in time-frequency analysis has focused on deterministic signals. Only recently has attention turned to non-stationary random processes [1–4].

Unlike the power spectrum for stationary random processes, there is no unique definition for the time-varying spectrum of a non-stationary random process. Because it satisfies a number of desirable properties, we choose the *Wigner-Ville spectrum* (WVS) [1] as our definition of the time-varying spectrum in this paper. The WVS $\mathbf{W}_x$ is the expected value of the Wigner distribution W_x [5] of one realization of the process $x(t)$

$$\begin{aligned} \mathbf{W}_x(t,f) &\equiv E\{W_x(t,f)\} \\ &= E\left\{\int x^*(t-\tau/2)\, x(t+\tau/2)\, e^{-j2\pi f\tau} d\tau\right\}, \end{aligned} \tag{1}$$

and reduces to the classical power spectrum for stationary signals. The problem of time-varying spectrum estimation can be stated as the estimation of $\mathbf{W}_x(t,f)$ given only one realization of the non-stationary process $x(t)$.

A number of different WVS estimates have been proposed. The simplest is the empirical Wigner distribution $W_x(t,f)$ itself. However, while it is unbiased, it has infinite variance. Smoothing reduces the variance of the empirical Wigner distribution. Two-dimensional convolution of the Wigner distribution with a signal-independent smoothing kernel yields a distribution in the time and frequency shift-covariant Cohen's class [1, 5]. Two-dimensional affine convolution yields a distribution in the time and scale covariant affine class [6]. Sayeed and Jones have developed a method for optimal kernel design for WVS estimation when the statistics of the process are known [2].

Unfortunately, the large amount of smoothing required to obtain a low variance WVS estimate can damage the resolution of line components. Line components are deterministic chirping signals of the form $e^{j2\pi\gamma(t)}$, whose ideal time-frequency representations have the form $\delta(f-\gamma'(t))$.

Realizing that random and deterministic spectral components must be dealt with separately, Thomson introduced a powerful multiple window (MW) spectrum estimator for stationary signals in [7]. The method uses a statistical significance test to detect and extract all sinusoids from the data, computes a MW spectrum estimate of the sinusoid-free data with optimal windows, and reshapes the spectrum to account for the excised sinusoids. Because of its excellent performance, several groups have applied this technique, ad hoc, to non-stationary signals in a piecewise fashion [3, 4, 8]. In this paper we refine these methods into an improved time-varying MW spectrum estimate for non-stationary signals. Our method preserves the resolution of line components, has low variance, and offers fine control over the bias-variance trade-off.

2. THOMSON'S MW METHOD

The classical spectrum estimator for stationary signals, the periodogram, is defined as simply the squared magnitude of the Fourier transform of the data. While the periodogram suffers from a large variance, this variance can be reduced by cutting the data into blocks, computing a periodogram of each block, and then averaging the periodograms. However, this procedure also smoothes and biases the resulting spectrum estimate.

Inspired by the notion of averaging but displeased with the resulting bias, Thomson suggested computing several periodogram estimates of the *entire signal* using a set of different windows and then averaging the resulting periodograms to construct a spectrum estimate [7]. For a low variance, low bias estimate, he demanded that the windows be (1) orthogonal (to minimize variance), and (2) optimally concentrated in frequency (to minimize bias). The optimal windows satisfying these requirements for signals of finite extent are the

*This work was supported by the National Science Foundation, grant no. MIP–9457438, the Office of Naval Research, grant no. N00014–95–1–0849, and the Texas Advanced Technology Program, grant no. TX–ATP 003604–002.

prolate spheroidal functions. In addition to multiple windows, Thomson also introduced into his estimate a separate procedure for deterministic sinusoidal components as mentioned above.

2.1. Summary of the Method

Thomson's MW method can be summarized in three steps [7]:

1. Detect and extract all significant sinusoids in the data $x(t)$[1] (see Section 2.2) to obtain the part $y(t)$ of the data having a continuous power spectrum

$$y(t) = x(t) - \{sinusoids\}. \quad (2)$$

2. Average K "orthogonal" periodogram estimates of $y(t)$ using prolate spheroidal data windows $\{v_k(t)\}$ [7]

$$P_T(f) = \sum_{k=0}^{K-1} d_k(f) \left| \int y(t)\, v_k(t)\, e^{-j2\pi ft}\, dt \right|^2. \quad (3)$$

   The orthogonal prolate spheroidal windows are perfectly suited to stationary spectrum estimation, because they are simultaneously compactly supported in time and optimally concentrated in frequency. This concentration property results in a low bias estimate of the spectrum. The weights $d_k(f)$ can be chosen adaptively to further reduce the bias [7].

3. Reshape the spectrum $P_T(f)$ to account for the sinusoids excised in Step 1.

2.2. Thomson's F-test For Sinusoids

Before we can extract the significant sinusoids from the data $x(t)$ as in (2), we must detect their presence and estimate their parameters.

We assume the signal model

$$x(t) = y(t) + \sum_i \mu(f_i)\, e^{j2\pi f_i t} \quad (4)$$

with $y(t)$ zero mean and Gaussian. Assuming that a sinusoid is present at frequency f, we estimate its complex amplitude and subtract it from the data to obtain an estimate of the "background" continuous spectrum around f. Comparing this power in the background spectrum with the power in the assumed sinusoid results in an F variance-ratio test $F(f)$ [7]. If $F(f)$ exceeds a significance threshold, we say that a sinusoid exists at frequency f. When $F(f)$ exceeds the threshold when no sinusoid is present at frequency f, we say a *spurious peak* occurs.

Averaging orthogonal periodogram estimates reduces the variance of the MW power spectrum estimate by approximately K times compared to the variance of a single periodogram (in which $K = 1$). Furthermore, concentrated windows, adaptive weights, and sinusoid extraction keep resolution very high. These properties make Thomson's MW method the tool of choice for estimating the power spectrum of stationary random processes.

[1] In [7], Thomson sets up the spectrum estimation problem in discrete-time.

3. MW TIME-FREQUENCY ANALYSIS

The excellent performance of Thomson's MW method has led several groups to apply the method to time-varying spectrum estimation by simply sliding the estimate (3) along the signal and computing a MW spectrogram estimate about each time point [3, 4, 8]. While reasonably effective on certain classes of piecewise stationary signals, this approach suffers from two primary drawbacks. First, prolate spheroidal window functions have no inherent optimality properties in the joint time-frequency domain. Second, Thomson's F-test sinusoid extraction procedure fails on chirping line components of rapidly changing instantaneous frequency. In this section, we will extend Thomson's MW method to the time-frequency and time-scale planes by identifying sets of optimal windows/wavelets, and by developing a linear-chirp extraction algorithm that better matches non-stationary line components.

3.1. Hermite Windows

The foundation of the stationary MW method rests on the fact that the prolate spheroidal functions are optimal windows for estimating the spectrum of a time-limited signal. This optimality does not carry over into time-frequency, however, since the prolate spheroidal functions treat the time-frequency plane as two separate spaces rather than as one geometric whole [9–11].

For time-frequency signal analysis, it is natural to average over multiple orthogonal windows that are optimally concentrated in an appropriate time-frequency domain. To date, optimal orthogonal functions of this kind have been found only for a few very special domains. The Hermite functions, defined as

$$h_k(x) = \pi^{-1/4}(2^k k!)^{-1/2} \left(x - \frac{d}{dx}\right)^k e^{-x^2/2} \quad (5)$$

for $k = 0, 1, 2, \cdots$, are optimally concentrated in the circular region [9–11]

$$R = \left\{(t, f): \quad t^2 + f^2 \le C\right\} \quad (6)$$

with C a constant. Thus, they treat all time-varying spectral features in the same fashion. The first four Hermite windows and their concentration region are shown in Figure 1.

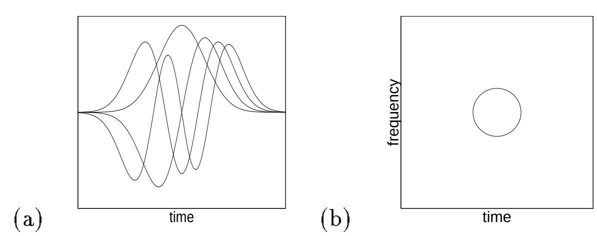

Figure 1: *(a) The first four Hermite functions in the time domain, and (b) their concentration region in time-frequency.*

3.2. Multiple Window WVS Estimate

Thomson's MW spectrum average (3) estimates the energy content of the signal at frequency f by projecting onto the windowed sinusoids $v_k(t)\, e^{j2\pi ft}$. By analogy, we estimate the energy content of a non-stationary signal at time

t and frequency f by projecting onto the sliding windowed sinusoids $h_k(\tau - t)\, e^{j2\pi f \tau}$. The result can be written as the average of K Hermite-windowed spectrograms of the data

$$\widehat{\mathbf{W}}_x(t,f) = \sum_{k=0}^{K-1} d_k(t,f) \left| \int y(\tau)\, h_k(\tau - t)\, e^{-j2\pi f \tau} d\tau \right|^2. \tag{7}$$

This WVS estimator has low variance thanks to the averaging and minimized bias due to the optimally concentrated Hermite windows. The bias-variance trade-off can easily be controlled and optimized by changing the number of windows K and by tuning the adaptive weighting functions $d_k(t,f)$ as in [7].

3.3. Cohen's Class Interpretation

The MW WVS estimate (7) is in Cohen's class of time-frequency distributions [5], each of which can be written as

$$W_x(t,f) \; ** \; \phi(t,f) \tag{8}$$

with $\phi(t,f)$ a kernel function. The kernel generating the MW WVS has a simple closed form expression

$$\phi(t,f) = \sum_{k=0}^{K-1} d_k(t,f)\, L_k(t,f) \tag{9}$$

where L_k, the k-th order Laguerre function, is the Wigner distribution of the k-th order Hermite function [12]. Since the weight functions $d_k(t,f)$ are tuned for each signal, the MW WVS estimate employs a signal-dependent kernel.

3.4. Extraction Of Line Components

As in Thomson's method for stationary signals, the averaging inherent in (7) will degrade the resolution of line components. Thus, we will first detect and extract all line components in the data before performing (7), and then reshape the estimate. We assume the signal model

$$x(t) = y(t) + \sum_i \mu_i(t)\, e^{j2\pi\gamma_i(t)} \tag{10}$$

with $y(t)$ zero mean and Gaussian.

A straightforward application of Thomson's sinusoid extraction algorithm to $x(t)$ in blocks as in [4] relies on an assumption that the chirp functions $e^{j2\pi\gamma_i(t)}$ can be closely approximated locally as sinusoids. Unfortunately, this is not the case for most chirping components; in these cases, the approach fails. In order to detect and extract highly non-stationary chirps, we have developed a statistical significance test for linear chirps of the form $e^{j2\pi(ft+ct^2)}$. Piecewise linear chirps can closely approximate all but the most rapidly changing chirp functions.

The test for linear chirp components flows as in Section 2.2, except that the F variance-ratio test must be performed at each time t, frequency f, and chirp rate c. This results in a three dimensional F-test statistic $F(t,f,c)$.

Due to the repeated application of the F variance-ratio test, the number of spurious peaks in F increases far beyond that seen in Thomson's method for stationary signals. These peaks must be suppressed in order to create a readable time-frequency image.

To suppress spurious peaks that peek above the significance threshold, we employ the following algorithm:

1. Threshold the data volume $F(t,f,c)$ and slice it along the chirp-rate dimension c.
2. For each fixed c_i, apply a nonlinear filter to $F(t,f,c_i)$ to remove peaks that have not coalesced into a region larger than the Heisenberg uncertainty principle mandates. (Intuition: spurious peaks are isolated in $F(t,f,c_i)$; true peaks lie along curves in $F(t,f,c_i)$.)
3. Combine the results from each c_i to obtain the final test statistic.

Although the above algorithm is computationally expensive, it is readily parallelizable.

3.5. Example

In Figure 2, we illustrate the performance of the MW WVS estimate using a test signal composed of a chirp with sinusoidal instantaneous frequency in an additive bandpass Gaussian noise of linearly rising center frequency. It is not possible to identify the components of the test signal from the empirical Wigner distribution due to its large variance. The spectrogram smoothes the Wigner distribution. Unfortunately the amount of smoothing needed to reduce the variance smears the line components excessively. A sliding version of Thomson's method as proposed in [3, 4] does not perform well for this non-stationary data, since a local sine approximation to the chirping line component is inadequate. The time-frequency MW estimate of Figure 2(d) on the other hand enjoys both high resolution and low variance. The variance of the MW WVS estimate is approximately $\frac{1}{4}$ that of the spectrogram, in agreement with the fact that four windows were used in the computation of the MW WVS estimate.

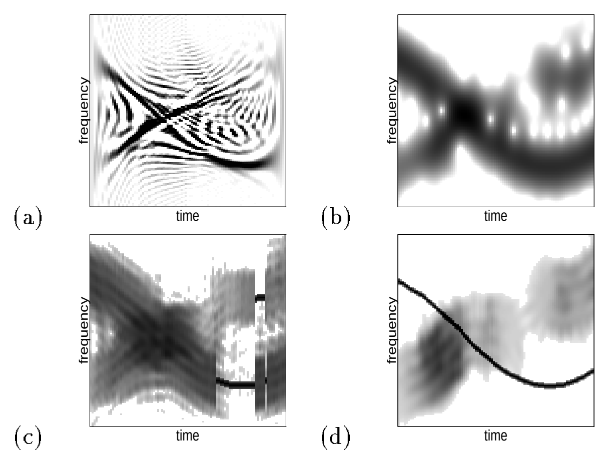

Figure 2: *Four time-varying spectrum estimates of a test signal. (a) Empirical Wigner distribution. (b) Spectrogram. (c) Sliding window Thomson's method [3, 4]. (d) Multiple window method.*

4. MW TIME-SCALE ANALYSIS

Random processes containing high frequency components of short duration and low frequency components of long duration are better matched by time-scale representations from the affine class [6]. The smoothing kernels of affine class representations change with frequency to accommodate such processes (see Figure 3).

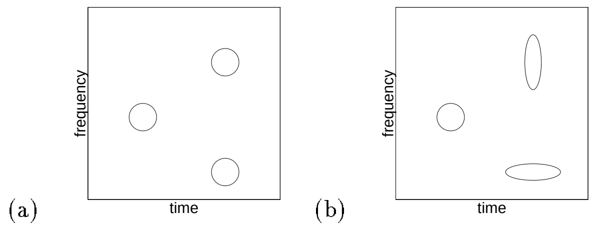

Figure 3: *Smoothing regions in the time-frequency plane in (a) Cohen's class, (b) affine class representations.*

The Morse wavelets [13,14] play a rôle in time-scale analogous to that of the Hermite windows in time-frequency. They are defined in the frequency domain as [13]

$$\Psi_k(f) = f^{\beta/2}\, e^{-f/2}\, \frac{d^\beta}{df^\beta} \left[e^f \, \frac{d^{\beta+k}}{df^{\beta+k}} \left(f^{\beta+k}\, e^{-f} \right) \right] \qquad (11)$$

with $k = 0, 1, 2, \cdots$, and β the degree of flatness at $f = 0$. The Morse functions are mutually orthogonal and maximally concentrated in the tear-drop shaped time-frequency region [14]

$$R = \left\{ (t, f) : \quad t^2 + \frac{C_1}{f^2} + 1 \leq \frac{C_2}{|f|} \right\} \qquad (12)$$

with C_1 and C_2 constants. A more general formulation of the Morse wavelets is given in [14]. Figure 4 shows the first four Morse wavelets and their concentration region R.

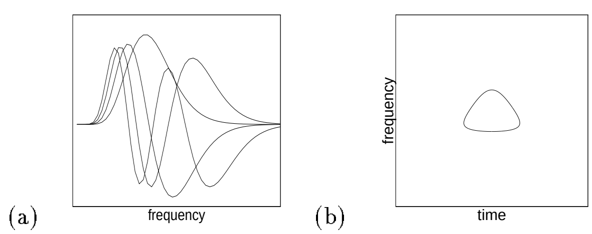

Figure 4: *(a) The first four Morse wavelets in the frequency domain, and (b) their concentration region in time-frequency.*

We form a time-scale MW WVS estimate of the data $x(t)$ as the weighted average of the squares of K wavelet transforms using the Morse wavelets

$$\widehat{\mathbf{W}}_x(t, a) \;=\; \sum_{k=0}^{K-1} d_k(t, a) \left| a^{-1/2} \int x(\tau)\, \psi_k\!\left(\frac{\tau - t}{a} \right) d\tau \right|^2 . \qquad (13)$$

where ψ_k is the k-th order Morse wavelet in the time domain. The estimate belongs to the affine class of time-scale covariant distributions [6]. Its kernel is the weighted sum of the Wigner distributions of the K Morse wavelets.

As in the time-frequency case, averaging degrades the resolution of chirping line components. Using the algorithm of Section 3.4, we can detect and extract the line components from the data before computing the estimate (13).

Lilly and Park have also considered multi-wavelet spectrum estimation [15]. In their work, they employed different wavelets and did not consider line component extraction.

5. CONCLUSIONS

In this paper, we have motivated and developed MW time-frequency and time-scale analysis for time-varying signals by fully extending Thomson's work [7] on MW spectrum estimation for stationary signals. Our contribution builds on the previous successful work of [3, 4, 8], yet differs from these approaches in two major ways:

1. We have identified two sets of optimal windows for time-varying spectrum estimation. They are the Hermite functions [9–11] for time-frequency analysis, and the Morse wavelets [13, 14] for time-scale analysis. Since these window sets are optimally concentrated in the time-frequency and time-scale planes, they result in low bias spectrum estimates.
2. We have developed an algorithm to detect and extract non-stationary line components from the data by approximating them as piece-wise linear chirps. We then form the MW WVS estimate of the chirp-free data and reshape the spectrum to account for the excised line components. This preserves the resolution of the line components.

ACKNOWLEDGMENTS

We thank Patrick Flandrin for encouraging us to publish these results, and David Thomson, Paulo Gonçalvès, and Rob Nowak for stimulating discussions.

REFERENCES

[1] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary random processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 1461–1470, Dec. 1985.

[2] A. M. Sayeed and D. L. Jones, "Optimal kernels for nonstationary spectral estimation," *IEEE Trans. Signal Processing*, vol. 43, pp. 478–491, Feb. 1995.

[3] G. Frazer and B. Boashash, "Multiple window spectrogram and time-frequency distributions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '94*, vol. IV, pp. 293–296, 1994.

[4] K. A. Farry, *Issues in Myoelectric Teleoperation of Complex Artificial Hands*. Ph.D. dissertation, Dep. Elec. Comput. Eng., Rice University, 1994.

[5] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] O. Rioul and P. Flandrin, "Time-scale energy distributions: A general class extending wavelet transforms," *IEEE Trans. Signal Processing*, vol. 40, pp. 1746–1757, July 1992.

[7] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, pp. 1055–1096, Sept. 1982.

[8] D. J. Thomson. Personal Communication.

[9] I. Daubechies, "Time-frequency localization operators: A geometric phase space approach," *IEEE Trans. Inform. Theory*, vol. 34, pp. 605–612, July 1988.

[10] P. Flandrin, "Maximum signal energy concentration in a time-frequency domain," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '88*, pp. 2176–2179, 1988.

[11] T. W. Parks and R. G. Shenoy, "Time-frequency concentrated basis functions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '90*, vol. 5, pp. 2459–2462, 1990.

[12] G. B. Folland, *Harmonic Analysis in Phase Space*. Princeton, NJ: Princeton University Press, 1989.

[13] P. Morse, "Diatomic molecules according to the wave mechanics II. Vibrational levels," *Physical Review*, vol. 34, pp. 57–64, July 1929.

[14] I. Daubechies and T. Paul, "Time-frequency localization operators — a geometric phase space approach: II. The use of dilations," *Inverse Problems*, no. 4, pp. 661–680, 1988.

[15] J. M. Lilly and J. Park, "Multiwavelet spectral and polarization analyses of seismic records," *Geophys. J. Int.*, no. 122, pp. 1001–1021, 1995.

Presented at the IEEE SP International Symposium on Time-Frequency
and Time-Scale Analysis, Paris, France, June 1996.