NONSTATIONARY SPECTRUM ESTIMATION AND TIME–FREQUENCY CONCENTRATION

James W. Pitton

MathSoft
1700 Westlake Ave N, Suite 500
Seattle, WA USA 98109
yojimbo@statsci.com

ABSTRACT

This paper extends Thomson's multitaper spectrum estimation method [17] to nonstationary signals. The method uses a newly–derived set of basis functions which generalize the concentration properties of the prolate spheroidal waveforms [15] to the time–frequency case. We solve for the basis which diagonalizes the nonstationary spectrum generating operator over a finite region of the time–frequency plane. These eigenfunctions are maximally concentrated to and orthogonal over the specified time-frequency region, and are thus doubly orthogonal. Individual spectrograms computed with these eigenfunctions form direct time–frequency spectrum estimates. We next present a multitaper time–frequency spectrum estimation procedure using these time–frequency eigenestimates. Bias and variance expressions are derived, allowing for a statistical characterization of the accuracy of the estimate. The time–frequency concentration property of the basis functions yields an estimator with excellent bias properties, while the variance of the estimate is reduced through the use of multiple orthogonal windows.

1. TIME–FREQUENCY SPECTRAL ANALYSIS

There have generally been two approaches to time–frequency spectral analysis. The evolutionary spectrum approaches (*e.g.*, [14, 7, 8]) model the spectrum as a slowly varying envelope of a complex sinusoid. This assumption allows the averaging of short–time spectral estimates to stabilize the variance. The second approach is commonly referred to as Cohen's bilinear class [3], which provides a general formulation for joint time–frequency distributions. Computationally, the evolutionary spectrum methods fall within Cohen's class.

A subclass of time–frequency distributions are the positive time–frequency distributions (TFDs) [4]. Postive TFDs are everywhere nonnegative, and yield the correct univariate marginal distributions in time and frequency (the instantaneous energy and the energy spectral density):

$$P(t,\omega) \geq 0, \tag{1}$$

$$\int P(t,\omega)d\omega = |s(t)|^2, \tag{2}$$

$$\int P(t,\omega)dt = |S(\omega)|^2, \tag{3}$$

where $S(\omega)$ denotes the Fourier transform of the finite energy signal $s(t)$, and all integrals are from $-\infty$ to ∞.[1]

The first method for generating positive TFDs used constrained optimization, minimizing the cross–entropy to a prior distribution subject to a set of linear constraints [9]. Positive TFDs have been linked to the evolutionary spectrum and estimated via deconvolution [13]. Least-squares estimation has also been used to compute positive TFDs [11]. Approximate solutions for positive TFDs have been obtained through a nonlinear combination of spectrograms [10].

Another approach to computing time–frequency spectra has been to extend Thomson's multitaper spectral estimation method [17] to the nonstationary case through a sliding–window framework [16]. [1] developed a multitaper time–frequency spectrum, including a significance test for nonstationary tones, using Hermite windows, which have previously been shown to maximize a time–frequency concentration measure [5]. [2] extended the Hermite multiwindow method to include a means of reducing artifacts using a time–frequency mask.

While these methods all provide some representation of the time–varying frequency content of a signal, they do not relate the computed distribution to an underlying time–frequency spectrum (*e.g.*, [1] minimizes the bias between the multitaper TFD and the Wigner distribution; however, the Wigner distribution is not nonnegative for arbitrary signals, and as such is not a valid time–frequency spectrum). As a result, there is no quantitative measurement of the accuracy of the representation. For time–frequency analysis to be useful in a wide variety of real–world applications, some method of measuring the bias and variance of the estimated time–frequency spectrum is required. To meet this requirement, we present a statistical spectral estimation method for nonstationary signals. The method is based on a time–varying filter formulation for positive TFDs, as discussed in [12]. We solve for the eigenvectors which diagonalize the nonstationary spectral generating function. These eigenvectors are maximally concentrated (and doubly–orthogonal)

[1] Throughout the analysis that follows, we use integral formulations of the various operations. The extension to the discrete, finite case is straightforward and not presented here. Integrals with no limits are over the entire domain of support of the integrand. The corresponding summations in the discrete case are then over the length of the corresponding vectors.

0-7803-4428-6/98 $10.00 © 1998 IEEE

in time–frequency. We then derive a multitaper estimation procedure to solve for the time–frequency spectrum. We also present bias and variance measures for the estimated time–frequency spectrum.

2. INTEGRAL EQUATION FOR A TIME–FREQUENCY SPECTRUM

As is the case in stationary spectral estimation, a rigorous approach to time–frequency spectral estimation should be based upon the integral formulation underlying the generation of nonstationary signals. The formulation used here is a straightforward extension of the spectral representation theorem for stationary processes [14], and is equivalent to a linear time–varying (LTV) filter model. Define the signal $s(t)$ as the output of a white-noise-driven LTV filter:

$$s(t) = \int h(t,\tau)e(\tau)d\tau. \tag{4}$$

$e(t)$ is bandlimited Gaussian white noise with bandwidth much greater than that of the filter $h(t,\tau)$:

$$e(t) = \int e^{j\omega t}dZ(\omega). \tag{5}$$

$dZ(\omega)$ is an orthogonal process with unit variance. The signal can then be written as:

$$s(t) = \int H(t,\omega)e^{j\omega t}dZ(\omega), \tag{6}$$

where $H(t,\omega)$ is defined as the Fourier transform of $h(t,t-\tau)$ [12]. The time-frequency spectrum is defined by:

$$P(t,\omega) = |H(t,\omega)|^2. \tag{7}$$

This formulation for a time–frequency spectrum is of the same general form as Priestley's evolutionary spectrum [14]. However, we do not require that $H(t,\omega)$ be slowly–varying. This form for $P(t,\omega)$ also satisfies the stochastic equivalent of the time and frequency marginals (equations 2–3); the relationship between the above time–varying spectrum and positive TFDs is discussed in [12].

Given a signal $s(t)$, we want to estimate $P(t,\omega)$; however, direct inversion of equation 6 is impossible. We can gain some idea of the time–varying frequency content of $s(t)$ by computing the short-time Fourier transform (STFT):

$$S_s(t,\omega) = \int s(\tau)g(t-\tau)e^{-j\omega\tau}d\tau, \tag{8}$$

where $g(t)$ is a rectangular window of length T. The relationship between the STFT and $H(t,\omega)$ is obtained by replacing $s(t)$ by its time–frequency spectral formulation:

$$S_s(t,\omega) = \int\int H(\tau,\theta)g(t-\tau)e^{-j(\omega-\theta)\tau}dZ(\theta)d\tau. \tag{9}$$

To solve for the time-varying spectrum $H(\tau,\theta)$, we need to invert the STFT operator $g(t-\tau)e^{-j\omega\tau}$. This inversion is an inherently ill-posed problem. Instead, we approximate the inverse solution by regularizing it to some region $R(t,\omega)$ in the time–frequency plane, much as Thomson regularized the spectral inversion to a bandwidth W in his multitaper approach [17]. For simplicity throughout, we will define $R(t,\omega)$ to be a square region of time–frequency of dimension $\Delta T \times \Delta W$; however, the results readily generalize to arbitrary regions $R(t,\omega)$.

In the case of spectral estimation, the operator is square and Toeplitz; its regularized inverse is found through an eigenvector decomposition. Such is not the case in the time–frequency problem; the STFT operator is neither full rank nor square. To diagonalize it, we apply a Singular Value Decomposition, finding the left and right eigenvectors $u(\tau)$ and $V(t,\omega)$ and the associated eigen (singular) values λ:

$$g(t-\tau)e^{-j\omega\tau} = \sum_k \lambda_k u_k(\tau)V_k^*(t,\omega). \tag{10}$$

The eigenvectors $u(\tau)$ and $V(t,\omega)$ form an STFT pair:

$$V(t,\omega) = \int u(\tau)g(t-\tau)e^{-j\omega\tau}d\tau \tag{11}$$

The SVD relationship between $u(\tau)$ and $V(t,\omega)$ is obtained by applying the STFT operator to $V(t,\omega)$, computing the integrals only over $\Delta T \times \Delta W$:

$$\lambda u(\tau) = \int_{\Delta T}\int_{\Delta W} V(t,\omega)g(t-\tau)e^{j\omega\tau}d\omega dt \tag{12}$$

The inverse STFT computed over all (t,ω) also holds. This equation can be reduced to a standard eigenvector equation by substituting for $V(t,\omega)$:

$$\lambda u(\tau) = \int_{\Delta T}\int_{\Delta W}\int u(s)g(t-s)g(t-\tau)e^{j\omega(\tau-s)}ds d\omega dt \tag{13}$$

The integral in ω reduces to a sinc function, or for the discrete case, Dirichlet's kernel. The integral in t can be easily computed over the two rectangular windows, the result of which for convenience we define by:

$$f(\tau,s) = \int_{\Delta T} g(t-s)g(t-\tau)dt \tag{14}$$

The eigenvalue equation for $u(\tau)$ is then:

$$\lambda u(\tau) = \int 2\Delta W \mathrm{sinc}(\Delta W(\tau-s))f(\tau,s)u(s)ds \tag{15}$$

We can then solve for $u(\tau)$ using standard eigenvector solution methods. Figure 1 plots the first 20 eigenvalues for this equation with $T = 256$, $\Delta T = 128$, and $\Delta W = 6/256$, so that $\Delta T \Delta W = 3$. Figure 2 illustrates the first 4 eigenvectors for the same values of ΔT and ΔW. Note that the $u_k(\tau)$ are of length $T + \Delta T$, as determined by the region of support of $f(\tau,s)$ in equation 15. This increase in length beyond T is a result of the convolution in time inherent in the STFT.

There are two extremely important properties of these eigenvectors for the time–frequency spectral estimation problem. These properties are obvious results of diagonalizing the STFT operator over a finite region of the time–frequency plane. First, the first left eigenvector u_1 maximizes a time–frequency energy concentration measure

$$u_1(\tau) = \arg_u \max \int_{\Delta T}\int_{\Delta W} |S_u(t,\omega)|^2 d\omega dt \tag{16}$$

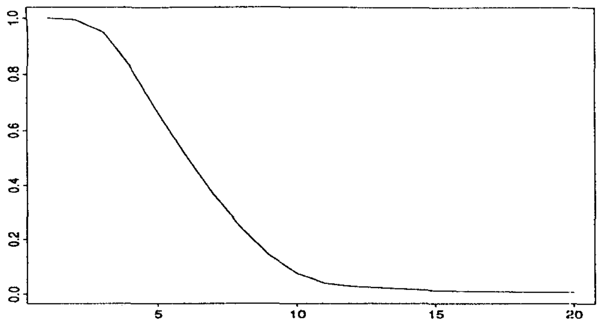

Figure 1: First 20 eigenvalues of the STFT kernel for $\Delta T \Delta W = 3$.

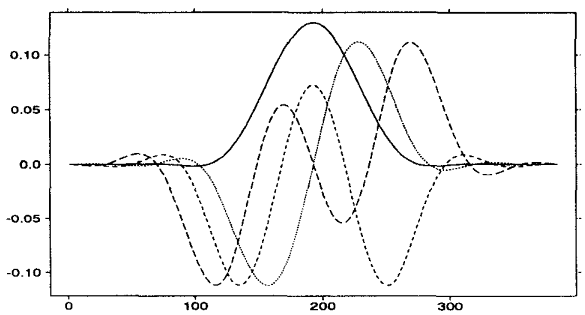

Figure 2: First 4 eigenvectors of the STFT kernel for $\Delta T \Delta W = 3$.

over the space of unit–energy functions. In other words, the first right eigenvector $V_1(t,\omega)$, the STFT of the first left eigenvector $u_1(\tau)$, has the greatest proportion of energy in $\Delta T \times \Delta W$ (more generally, in R) for any STFT. The second eigenvector maximizes the energy in a subspace orthogonal to the first, and so on. It is straightforward to show that maximization of this integral reduces to the eigenequation for $u(\tau)$. This time–frequency concentration property is akin to the frequency concentration property of the prolate spheroidal waveforms (or the prolate spheroidal sequences, in the discrete case) [15]. The second important property, again similar to the Slepian functions, is a double–orthogonality property. Since the $u_k(\tau)$ are orthonormal functions on the real line, the $V_k(t,\omega)$ are orthonormal on the entire time–frequency plane. In addition, the $V_k(t,\omega)$ are also orthogonal over $\Delta T \times \Delta W$. This orthogonality is to be expected, given the maximization of the time–frequency concentration measure. The Slepian functions, by contrast, are orthonormal on the real line and orthogonal on the interval W. Looking ahead, it is this orthogonality on $\Delta T \times \Delta W$ that will let us construct a regularized estimate of the time–frequency spectrum, just as the orthogonality of the Slepian functions on W provide the basis for Thomson's multitaper spectral estimation method.

3. EIGENESTIMATES

We now want to estimate $H(t,\omega)$ regularized to $\Delta T \times \Delta W$. Following Thomson's approach, we form a Fourier–Bessel expansion of $H(t,\omega)$ over the rectangular region $\Delta T \times \Delta W$ around (t,ω):

$$H_k(t,\omega) \doteq \frac{1}{\sqrt{\lambda_k}} \int_{\Delta T}\int_{\Delta W} H(\tau,\theta) V_k(t-\tau,\theta-\omega) dZ(\theta) d\tau. \tag{17}$$

H_k is thus a direct, but unobservable, projection of $H(t,\omega)$ onto $\Delta T \times \Delta W$.

We next form an estimate of these expansion coefficients using the STFT. We have from above that the STFT of $s(t)$ is given by:

$$S_s(t,\omega) = \int\int H(\tau,\theta) g(t-\tau) e^{-j(\omega-\theta)\tau} dZ(\theta) d\tau. \tag{18}$$

Inserting the STFT $S_g(t,\omega)$ for the rectangular window $g(t)$ gives:

$$\begin{aligned} S_s(t,\omega) \quad = \quad & \int\int\int H(\tau,\theta) S_g(t-\tau,\phi) \\ & e^{-j(\omega-\theta+\phi)\tau+j\phi t} d\phi dZ(\theta) d\tau. \end{aligned} \tag{19}$$

In this formulation, $H(\tau,\theta)$ is the time–varying envelope of a complex exponential $e^{j\theta\tau}$, and as such is relatively smooth. Since $g(t)$ and $S_g(t,\omega)$ are also smooth, the integral in τ will be approximately zero when the argument of the complex exponential is nonzero. This condition holds when $\phi = \theta - \omega$, and equation 19 reduces to:

$$S_s(t,\omega) = \int\int H(\tau,\theta) S_g(t-\tau,\theta-\omega) e^{j(\theta-\omega)t} dZ(\theta) d\tau. \tag{20}$$

It follows (with a change in notation) that the STFT of $s(t)$ computed with $u_k(\tau)$ is

$$S_k(t,\omega) = \int\int H(\tau,\theta) V_k(t-\tau,\theta-\omega) e^{j(\theta-\omega)t} dZ(\theta) d\tau, \tag{21}$$

i.e., the k^{th} eigenspectrum $S_k(t,\omega)$ is a projection of $H(t,\omega)$ onto the k^{th} right eigenvector $V_k(t,\omega)$, estimating $H_k(t,\omega)$ over $\Delta T \times \Delta W$. When $s(t)$ is a stationary white noise process, it follows that

$$E[|S_k(t,\omega)|^2] = |H(t,\omega)|^2 = P(t,\omega). \tag{22}$$

Thus, the individual eigenspectra are direct estimates of $P(t,\omega)$, and are unbiased when the spectrum is white.

Next, we form an estimate of $H(t,\omega)$ over $\Delta T \times \Delta W$ using a Fourier–Bessel series:

$$\hat{H}(\bar{t},\bar{\omega};t,\omega) = \sum_{k=1}^{K} V_k(\bar{t}-t,\bar{\omega}-\omega) \frac{H_k(t,\omega)}{\lambda_k}, \tag{23}$$

where $K \approx \Delta T \Delta W$. Choosing $\Delta T \Delta W$ too small will result in estimates with poor bias and variance properties. The magnitude–square of $\hat{H}(\bar{t},\bar{\omega};t,\omega)$ yields our estimate of $P(t,\omega)$ over $\Delta T \times \Delta W$. This estimate is a χ^2 random variable with two degrees of freedom for frequencies other than DC or Nyquist. As such, its variance is $P^2(t,\omega)$. To reduce the variance, we average the estimate over $\Delta T \times \Delta W$, making use of the orthogonality of $V_k(t,\omega)$ on this region:

$$\hat{P}(t,\omega) = \frac{1}{\Delta T \Delta W} \int_{\Delta T} \int_{\Delta W} \left| \hat{H}(\bar{t},\bar{\omega};t,\omega) \right|^2 d\bar{t} d\bar{\omega}$$
$$= \frac{1}{\Delta T \Delta W} \sum_{k=1}^{K} \frac{|S_k(t,\omega)|^2}{\lambda_k}. \quad (24)$$

Averaging K "eigenspectrograms" results in a χ^2 random variable with $2K$ degrees of freedom; the variance of this estimate is then $P^2(t,\omega)/K$. If we choose ΔT to be a fixed proportion of the window length T, then this estimator is consistent for fixed ΔW.

While the individual eigenspectrograms, and hence, their sum, are unbiased for white noise, there will be bias in the final estimate due to leakage from the time–frequency smoothing window. The expected value of spectrograms for estimating a time–frequency spectrum has been derived previously [13]. Using those results, it follows that the expected value for the multitaper time–frequency spectrum is:

$$E[\hat{P}(t,\omega)] = \int\int P(t-\tau,\omega-\theta) \sum_{k=1}^{K} \frac{W_k(\tau,\theta)}{\lambda_k} d\theta d\tau, \quad (25)$$

where $W_k(\tau,\theta)$ is the Wigner distribution of the k^{th} left eigenvector $u_k(\tau)$. Since the eigenvectors are localized to $\Delta T \times \Delta W$, their Wigner distributions will likewise be localized. Note, however, that concentration of the Wigner distribution of the eigenvectors was not a factor in the derivation; see [6] for a discussion of that topic. Using such Wigner–concentrated eigenvectors would produce a multitaper estimate with minimum broadband bias.

4. CONCLUSIONS

We have presented a multitaper method of estimating time–varying spectra. The methodology follows that used by Thomson in his seminal 1982 paper [17], extended throughout to the time–frequency case. Specifically, we used a family of orthonormal windows whose corresponding short–time Fourier transforms are doubly orthogonal and maximally concentrated in time–frequency. The multitaper estimate approximately solves the integral equation for time–frequency spectra, providing a local least squares solution. Expressions for computing the bias and variance are provided; the multitaper estimate is also consistent for fixed bandwidth resolution. This solution to the problem of estimating time–varying spectra is computationally efficient, easily automated, and, most importantly, provides a means of quantifying the accuracy and stability of the estimate.

5. REFERENCES

[1] M. Bayram and R. G. Baraniuk. Multiple window time-frequency analysis. In *Proc. IEEE-SP Int. Sym. Time-Freq. Time-Scale Anal.*, pp. 173–176, 1996.

[2] F. Çakrak and P. Loughlin. Multiple window non–linear time–varying spectral analysis. *IEEE Transactions on Signal Processing*, 1997. submitted.

[3] L. Cohen. *Time-Frequency Analysis*. Prentice-Hall, 1995.

[4] L. Cohen and T. Posch. Positive time-frequency distributions. *IEEE Transactions on Acoustics Speech and Signal Processing*, 33(1):31–38, 1985.

[5] I. Daubechies. Time-frequency localization operators: a geometric phase space approach. *IEEE Transactions on Information Theory*, 34(4):605–612, 1992.

[6] P. Flandrin. Maximum signal energy concentration in a time–frequency domain. In *ICASSP-88*, pages 2176–2179. IEEE Signal Processing Society, 1988.

[7] A. Kayhan, A. El-Jaroudi, and L. Chaparro. The evolutionary periodogram for non-stationary signals. *IEEE Transactions on Signal Processing*, 42(6), 1994.

[8] A. Kayhan, A. El-Jaroudi, and L. Chaparro. Data–adaptive evolutionary spectral estimation. *IEEE Transactions on Signal Processing*, 43(1), 1995.

[9] P. Loughlin, J. Pitton, and L. Atlas. Construction of positive time-frequency distributions. *IEEE Transactions on Signal Processing*, 42(10):2697–2705, 1994.

[10] P. Loughlin, J. Pitton, and B. Hannaford. Approximating time-frequency density functions via optimal combinations of spectrograms. *IEEE Signal Processing Letters*, 1(12):199–202, 1994.

[11] J. Pitton. An algorithm for weighted least squares positive time–frequency distributions. In *SPIE Advanced Sig. Proc. Algs. Archs., Impl. VII*, volume 3162, 1997.

[12] J. Pitton. Linear and quadratic methods for positive time-frequency distributions. In *IEEE ICASSP'97*, vol. V, pp. 3649–3652, 1997.

[13] J. Pitton, L. Atlas, and P. Loughlin. Applications of positive time-frequency distributions to speech processing. *IEEE Trans. Sp. Audio Proc.*, 2(4):554–566, 1994.

[14] M. Priestley. *Spectral Analysis of Time Series*. Academic Press - London, 1981.

[15] D. Slepian and H. Pollack. Prolate spheroidal wave functions, Fourier analysis and uncertainty. *Bell Sys. Tech. J.*, 40(1), 1961.

[16] D. Thomson and A. Chave. Jackknifed error estimates for spectra, coherences, and transfer functions. S. Haykin, (ed.), *Advances in Spectrum Analysis and Array Processing, Vol. I*, 58–113. Prentice–Hall, 1991.

[17] D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70:1055–1096, 1982.