# ADAPTING MULTITAPER SPECTROGRAMS TO LOCAL FREQUENCY MODULATION

*James W. Pitton*

Applied Physics Laboratory, University of Washington
1013 NE 40th St.
Seattle, WA 98105
E-mail: pitton@apl.washington.edu

**ABSTRACT**

This paper presents further extensions to the multitaper time–frequency spectrum estimation method developed by the author. The method uses time–frequency (TF) concentrated basis functions which diagonalize the nonstationary spectrum generating operator over a finite region of the TF plane. Individual spectrograms computed with these eigenfunctions form direct TF spectrum estimates, and are combined to form the multitaper TF spectrum estimate. A method is presented for adapting the multitaper spectrogram to locally match frequency modulation in the signal, which can cause broadening of the spectral estimate. An F-test for detecting and removing frequency-modulated tones is also given.

## 1. INTRODUCTION

Thomson's multitaper spectral estimation approach [1] is a powerful method for nonparametric spectral estimation. This method uses a set of orthogonal data tapers that are maximally concentrated in frequency and diagonalize the spectral generating operator. These tapers are used to approximately invert the operator and estimate the spectrum. The multitaper approach was first applied to time–frequency (TF) analysis by a direct extension to the nonstationary case through a sliding-window framework [2], in which spectrograms are computed with each of the tapers and combined to form an estimate of the TF spectrum. A multitaper TF spectrum was constructed using spectrograms computed with Hermite windows [3], which had previously been shown to maximize a TF concentration measure [4]. This method was extended to include a means of reducing artifacts using a TF mask [5]. More recently, a multitaper method for TF analysis was presented by this author [6] that diagonalized the *nonstationary* spectral generating operator, formally extending Thomson's approach to TF. Subsequent work by the author gave bias and variance measures for the estimated TF spectrum, presented an adaptive procedure to reduce the bias of the individual spectrograms, and derived other properties of the eigenfunctions and the resulting TF spectral estimate [7, 8].

In this paper, a method is presented for adapting the multitaper spectrogram to locally match frequency modulation in the signal, which can cause broadening of the spectral estimate. Frequency modulation (FM) in the signal will degrade the resolution and accuracy of the multitaper spectrogram due to well-known spectral broadening effects. One common way of alleviating the effects of the spectral broadening is to match the spectrogram to the FM by frequency-modulating the window. This approach works perfectly well when there is only one FM rate in the signal, as is the case with chirped sonar and radar. However, in multicomponent signals such as speech, biological, and mechanical signals, there can be multiple FM rates present at any given time. To accurately analyze these types of signals, it is necessary to *locally* adapt the multitaper spectrogram to the FM at a given TF region. This paper presents a method for performing this local adaptation. An F-test for detecting and removing frequency-modulated tones is also given.

## 2. BACKGROUND: MULTITAPER TIME–FREQUENCY SPECTROGRAMS

This approach to TF spectral estimation is based on a straightforward extension of the spectral representation theorem for stationary processes [9], 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

This work was supported by the National Science Foundation and the Office of Naval Research.

0-7803-5988-7/00/$10.00 © 2000 IEEE

filter. The signal can then be written as:

$$s(t) = \int H(t,\omega)e^{j\omega t}dZ(\omega), \qquad (1)$$

where $H(t,\omega)$ is defined as the Fourier transform of the LTV filter $h(t,t-\tau)$ [10]. The TF spectrum is defined by:

$$P(t,\omega) = |H(t,\omega)|^2. \qquad (2)$$

This formulation for a TF spectrum is of the same general form as Priestley's evolutionary spectrum [9]; however, $H(t,\omega)$ is not constrained to be slowly–varying.

Given a signal $s(t)$, an estimate $P(t,\omega)$ is desired; however, direct inversion of equation (1) is impossible. A rough estimate of the time–varying frequency content of $s(t)$ may be obtained by computing its short-time Fourier transform (STFT):

$$S_s(t,\omega) = \int s(\tau)g(t-\tau)e^{-j\omega\tau}d\tau, \qquad (3)$$

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

$$S_s(t,\omega) = \int\int H(\tau,\theta)g(t-\tau)e^{-j(\omega-\theta)\tau}dZ(\theta)d\tau. \qquad (4)$$

To solve for the time–varying spectrum $H(\tau,\theta)$, the STFT operator $g(t-\tau)e^{-j\omega\tau}$ must be inverted. This inversion is an inherently ill-posed problem. Instead, the inverse solution is approximated by regularizing it to some region $R(t,\omega)$ in the TF plane, much as Thomson regularized the spectral inversion to a bandwidth $W$ in his multitaper approach [1]. For simplicity throughout, $R(t,\omega)$ is defined to be a square TF region of dimension $\Delta T \times \Delta W$; however, the results readily generalize to arbitrary regions.

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 TF problem; the STFT operator is neither full rank nor square. This operator is diagonalized using a Singular Value Decomposition, giving left and right eigenvectors $u(\tau)$ and $V(t,\omega)$ and the associated eigen (singular) values $\lambda$:

$$g(t-\tau)e^{-j\omega\tau} = \sum_k \lambda_k u_k(\tau)V_k^*(t,\omega). \qquad (5)$$

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

$$V(t,\omega) = \int u(\tau)g(t-\tau)e^{-j\omega\tau}d\tau. \qquad (6)$$

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 $\Delta T \times \Delta W$:

$$\lambda u(\tau) = \int_{\Delta T}\int_{\Delta W} V(t,\omega)g(t-\tau)e^{j\omega\tau}d\omega dt. \qquad (7)$$

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

$$\lambda u(\tau) = \int 2\Delta W \mathrm{sinc}(\Delta W(\tau - s))f(\tau,s)u(s)ds, \qquad (8)$$

where

$$f(\tau,s) = \int_{\Delta T} g(t-s)g(t-\tau)dt. \qquad (9)$$

$u(\tau)$ can be computed using standard eigenvalue solution methods. As has been discussed elsewhere, the eigenvectors are concentrated in TF and doubly orthogonal, both over the entire TF plane and over $\Delta T \times \Delta W$. These properties are critical for the estimation method.

Next, $H(t,\omega)$ is estimated regularized to the rectangular region $\Delta T \times \Delta W$ by projecting it onto $\Delta T \times \Delta W$ in the vicinity of $(t,\omega)$ using the $k^{th}$ left eigenvector $u_k(t)$:

$$H_k(t,\omega) \doteq \lambda_k^{-\frac{1}{2}} \int_{\Delta T}\int_{\Delta W} H(\tau,\theta)u_k(t-\tau)e^{j(\theta-\omega)\tau}dZ(\theta)d\tau. \qquad (10)$$

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

These expansion coefficients are then estimated using the STFT of $s(t)$ computed using $u_k(t)$:

$$S_k(t,\omega) = \int\int H(\tau,\theta)u_k(t-\tau)e^{-j(\omega-\theta)\tau}dZ(\theta)d\tau, \qquad (11)$$

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

$$E[|S_k(t,\omega)|^2] = |H(t,\omega)|^2 = P(t,\omega). \qquad (12)$$

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

Next, $H(t,\omega)$ is estimated over $\Delta T \times \Delta W$ using the right eigenvectors $V_k(t,\omega)$ weighted by the projections of $H(t,\omega)$ onto $u_k(t)$, *i.e.*, the $k^{th}$ spectrogram:

$$\hat{H}(\bar{t},\bar{\omega};t,\omega) = \sum_{k=1}^{K} V_k(\bar{t}-t,\bar{\omega}-\omega)S_k(t,\omega), \qquad (13)$$

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 $\hat{H}(\bar{t}, \bar{\omega}; t, \omega)$ is an estimate of $P(t, \omega)$ over $\Delta T \times \Delta W$. This estimate is a $\chi^2$ random variable with two degrees of freedom (except for DC and Nyquist) with variance $P^2(t, \omega)$. The variance of this estimate can be reduced by averaging over $\Delta T \times \Delta W$ and invoking the orthogonality of $V_k(t, \omega)$:

$$\begin{aligned} \hat{P}(t,\omega) &= \frac{1}{\Delta T \Delta W} \int_{\Delta T} \int_{\Delta W} \left| \hat{H}(\bar{t}, \bar{\omega}; t, \omega) \right|^2 d\bar{t} d\bar{\omega} \\ &= \frac{1}{\Delta T \Delta W} \sum_{k=1}^{K} \lambda_k \left| S_k(t,\omega) \right|^2. \end{aligned} \quad (14)$$

The average of $K$ direct estimates is a $\chi^2$ random variable with $2K$ degrees of freedom; hence, the variance of this estimate is $P^2(t, \omega)/K$. If $\Delta T$ is chosen to be a fixed proportion of the window length $T$, then this estimator is consistent for fixed $\Delta W$. Note that the form of this estimator differs slightly from that presented previously [6, 7, 8] in the weighting by the eigenvalues.

## 3. LOCALLY STATIONARY PROCESSES

The estimate for $P(t, \omega)$ given in equation (14) is unbiased for white noise. For the estimate to be unbiased for signals other than white noise, it is only necessary that $P(t, \omega)$ be *locally* white in TF, since the estimate is regularized to $\Delta T \times \Delta W$. A similar requirement is seen in the stationary case [1], wherein the spectrum is assumed to be smoothly varying so that it is approximately white over $\Delta W$. A class of stochastic processes known as *locally stationary* processes [12] satisfy the requirement of being smoothly varying in TF, and can be used to describe a wide variety of nonstationary signals. Locally stationary processes are stochastic processes with covariance functions of the form

$$R(t_1, t_2) = E\left[s(t_1)s^*(t_2)\right] = g\left(\frac{t_1 + t_2}{2}\right) f(t_1 - t_2), \quad (15)$$

where $g(\cdot)$ is a nonnegative function and $f(\cdot)$ is a valid covariance function; that is, $f(t)$ possesses a nonnegative Fourier transform $F(\omega)$. Through a change of variables, the symmetric form of the covariance function is seen to be:

$$R_s(t, \tau) = E\left[s(t + \tau/2)s^*(t - \tau/2)\right] = g(t)f(\tau), \quad (16)$$

The TF spectrum is thus given by [11]:

$$P_s(t, \omega) = g(t)F(\omega). \quad (17)$$

For locally stationary $s(t)$, $P_s(t, \omega)$ will be approximately constant over $\Delta T \times \Delta W$, and equation (12) will still hold.

The class of processes with such nonnegative TF spectra is easily extended to include a wider range of nonstationary processes [13]. Let $s(t)$ be a locally stationary process with covariance function $R_s(t, \tau)$ and corresponding TF spectrum $P_s(t, \omega)$. Then the linearly frequency modulated signal $s(t)e^{j\beta t^2/2}$ will have covariance $R_s(t, \tau)e^{j\beta t\tau}$ and corresponding nonnegative TF spectrum $P_s(t, \omega - \beta t)$. More generally, let $x(t) = s(t)e^{j\phi(t)}$, where $s(t)$ is locally stationary with symmetric covariance function $R_s(t, \tau)$ from equation (16). Then the covariance of $x(t)$ is

$$R_x(t, \tau) = g(t)f(\tau)e^{j(\phi(t+\tau/2) - \phi(t-\tau/2))}. \quad (18)$$

By making use of the principle of stationary phase [14], it can be shown [13] that the TF spectrum of $x(t)$ is given by:

$$P_x(t, \omega) = g(t)F(\omega - \phi'(t)) = P_s(t, \omega - \phi'(t)). \quad (19)$$

Thus, a frequency modulated locally stationary (FMLS) process will have a TF spectrum equal to that of the locally stationary process centered around the instantaneous frequency of the FM. The generalization can be taken one step further to define a *composite FMLS process*, consisting of a sum of statistically independent FMLS processes. The composite signal will also have a nonnegative TF spectrum equal to the sum of the spectra of the individual processes.

However, when $s(t)$ is an FMLS process, $P(t, \omega)$ will most certainly *not* be constant over $\Delta T \times \Delta W$, and equation (12) will fail to be valid. In this case, the smoothing region $\Delta T \times \Delta W$ must be oriented to match the FM of the signal. This reorientation is equivalent to matching the spectrogram window to the FM of the signal. This matching can be accomplished by using a frequency modulated window in the original STFT computation. However, in signals with multiple FM rates, as in a composite FMLS signal, this adaptation must be performed locally in TF, as discussed next.

## 4. LOCALLY MATCHED MULTITAPER SPECTROGRAMS

To locally demodulate the spectrograms, it is first necessary to construct a reliable estimate of the local FM, which is denoted by $\beta(t, \omega)$. Letting the TF dependence be implicit, $\beta$ can be estimated by computing a local covariance of the multitaper spectrogram normalized by the time spread: $\langle (t-\bar{t})(\omega-\bar{\omega}) \rangle / \langle (t-\bar{t})^2 \rangle$, where $\bar{t}$ and $\bar{\omega}$ are the local average time and frequency, respectively; their dependence on $t$ and $\omega$ is implied. The covariance is computed by integrating over a finite region of the TF plane $\Delta T \times \Delta W$ as a two–dimensional sliding window to provide an estimate of $\beta$ as a function of $t$ and $\omega$:

$$\beta(t,\omega) = \frac{\int_{\Delta T}\int_{\Delta W}(t-\hat{t}-\bar{t})(\omega-\hat{\omega}-\bar{\omega})P(\hat{t},\hat{\omega})d\hat{t}d\hat{\omega}}{\int_{\Delta T}\int_{\Delta W}(t-\hat{t}-\bar{t})^2P(\hat{t},\hat{\omega})d\hat{t}d\hat{\omega}}; \tag{20}$$

$\bar{t}$ and $\bar{\omega}$ are computed similarly. Integrating over a larger region will provide better variance properties at the expense of possible bias due to multiple signal components with differing FM rates lying within the area of integration.

Once $\beta(t,\omega)$ has been estimated, each STFT $S_k(t,\omega)$ is dechirped by locally convolving it with the Fourier transform of $e^{j\beta(t,\omega)\tau^2/2}$:

$$S_k^\beta(t,\omega) = \int S_k(t,\omega-\theta)e^{-j\theta^2/2\beta(t,\omega)}d\theta. \tag{21}$$

This convolution is shift–variant; at each frequency, a new $\beta$ must be used. This convolution is equivalent to matching the STFT to the local chirp rate. While this convoluation at first would appear to be an $O(N^2)$ operation, it can actually be implemented much more efficiently. The equivalent chirp in the time domain is of length $T$, the length of the STFT window. The Fourier transform of this finite–length chirp will then have bandwidth $\beta T$. Thus, if the average bandwidth of the various FM components is $M = \beta T$ bins, an STFT with N frequency samples can be dechirped with only $NM$ multiplies per time slice, comparable to the computational complexity of the STFT itself. Once all of the $S_k(t,\omega)$ are dechirped, the multitaper estimate is constructed as usual.

## 5. F-TEST FOR FREQUENCY–MODULATED TONES

The validity of the multitaper estimate rests on the assumption that the TF spectrum is smoothly varying over $\Delta T \times \Delta W$. This assumption is violated when spectral lines (FM or otherwise) are present in the signal. In this case, it is necessary to estimate the tones and remove them from the signal. Ordinarily, estimating a tone with unknown FM would be extremely difficult. This task is made easier, however, by the local matching described above. Once the individual STFT's $S_k(t,\omega)$ have been adapted to local FM, any frequency modulated tones in the signal will behave exactly as a stationary tone would behave in a non–adapted STFT. As a result, an F-test for the existence of any FM tones in the TF spectrum can be defined by directly extending Thomson's approach in the stationary case. The expected value of the $k^{th}$ dechirped STFT for an FM tone $\mu e^{j\phi(t)}$ with instantaneous frequency $\omega = \phi'(t)$ is:

$$E[S_k(t,\omega)] = \mu U_k(0). \tag{22}$$

The mean can then be estimated via regression:

$$\hat{\mu}(t,\omega) = \frac{\sum_{k=1}^K U_k(0)S_k(t,\omega)}{\sum_{k=1}^K U_k^2(0)}. \tag{23}$$

The variance of this estimate is equal to the background TF spectrum minus the spectral line, which is:

$$P(t,\omega) = \frac{1}{K-1}\sum_{k=1}^K |S_k(t,\omega) - \hat{\mu}(t,\omega)U_k(0)|^2. \tag{24}$$

The F-test at time $t$ is then given by the ratio of the power of the spectral line and that of the background spectrum:

$$F(t,\omega) = \frac{(K-1)|\hat{\mu}(t,\omega)|^2\sum_{k=1}^K U_k^2(0)}{\sum_{k=1}^K |S_k(t,\omega)-\hat{\mu}(t,\omega)U_k(0)|^2}. \tag{25}$$

Under the null hypothesis, the test quantity at a single time is the ratio of two $\chi^2$ random variables with 2 and $2(K-1)$ degrees of freedom. For a signal of length $T$ and an STFT of order $N$, there will be $T/N$ independent blocks of data. Thus, the final F-test will be a ratio of $\chi^2$ random variables with $2T/N$ and $2(K-1)T/N$ degrees of freedom, integrated along the contour specified by $\omega = \phi'(t)$:

$$F(\phi'(t)) = \frac{(K-1)\sum_{t=1}^T|\hat{\mu}(t,\phi'(t))|^2\sum_{k=1}^K U_k^2(0)}{\sum_{t=1}^T\sum_{k=1}^K |S_k(t,\phi'(t))-\hat{\mu}(t,\phi'(t))U_k(0)|^2}. \tag{26}$$

If the F-test achieves the specified confidence level, the tone should be removed by subtracting from the STFT's prior to forming the TF spectrum, then added into the representation as an impulse:

$$P(t,\omega) = \hat{\mu}(t,\omega)\delta(\omega-\phi'(t)) + \frac{1}{K}\sum_{k=1}^K |S_k(t,\omega)-\hat{\mu}(t,\omega)U_k(\omega-\phi'(t))|^2. \tag{27}$$

Matching the STFTs to the local FM greatly simplifies the F-test. With no matching, the STFT of an FM tone will be spread according to the sweep rate, and will thus have a functional form dependent on $\beta$. After matching, the FM tone will have the same response as a stationary tone in an unmatched STFT. Thus, the expression for $\mu$ in equation (23) can be used for all FM rates. The procedure for testing for an FM tone is then a four-step process: compute the test statistic $F(t,\omega)$ over time and frequency; find candidate contours $\omega(t) = \phi'(t)$ in $F(t,\omega)$; compute $F(\phi'(t))$; and test its significance.

## REFERENCES


[1] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, pp. 1055–1096, 1982.

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

[3] M. Bayram and R. G. Baraniuk, "Multiple window time-frequency analysis," in *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, (Paris, France), pp. 173–176, June 1996.

[4] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach," *IEEE Transactions on Information Theory*, vol. 34, no. 4, pp. 605–612, 1992.

[5] F. Çakrak and P. Loughlin, "Multiple window non–linear time–varying spectral analysis," *IEEE Transactions on Signal Processing*, 1999. submitted.

[6] J. Pitton, "Nonstationary spectrum estimation and time-frequency concentration," in *IEEE Conference on Acoustics, Speech, and Signal Processing*, vol. IV, pp. 2425–2428, IEEE, 1998.

[7] J. Pitton, "Time-frequency spectrum estimation: an adaptive multitaper method," in *IEEE Int. Sym. Time-Frequency and Time-Scale Analysis*, (Pittsburgh, PA), pp. 665–668, 1998.

[8] J. Pitton, "Adaptive multitaper time-frequency spectrum estimation," in *SPIE Advanced Sig. Proc. Algs. Archs., Impl. VII*, 1999.

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

[10] J. Pitton, "Linear and quadratic methods for positive time-frequency distributions," in *IEEE Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 3649–3652, IEEE, 1997.

[11] P. Flandrin, "On the positivity of the wigner–Ville spectrum," *Signal Processing*, vol. 11, no. 2, 1985.

[12] R. Silverman, "Locally stationary random processes," *IRE Transactions on Information Theory*, vol. IT-3, September 1957.

[13] J. Pitton, "The statistics of time-frequency analysis," *Journal of the Franklin Institute*, 2000. To appear.

[14] E. Key, E. Fowle, and R. Haggarty, "A method of designing signals of large time-bandwidth product," *IRE Intern. Conv. Record*, vol. IV, 1961.