MULTITAPER TIME-FREQUENCY REASSIGNMENT

Jun Xiao and Patrick Flandrin

Laboratoire de Physique (UMR CNRS 5672), École Normale Supérieure de Lyon
46, allée d'Italie 69364 Lyon Cedex 07, France
{Jun.Xiao,flandrin}@ens-lyon.fr

ABSTRACT

A method is proposed for estimating chirp signals embedded in nonstationary noise, with the twofold objective of a sharp localization for the chirp components and a reduced level of statistical fluctuations for the noise. The technique consists of combining time-frequency reassignment with multitapering. The principle of the method is outlined, its implementation based on Hermite functions is justified and discussed, and some examples are provided for supporting the efficiency of the approach, both qualitatively and quantitatively.

1. INTRODUCTION

In nonstationary contexts, it is well-known [7] that Fourier-based methods of (time-varying) spectrum estimation are classically faced with intrinsic limitations and different kinds of trade-offs: (*i*) from a statistical point of view, the usual bias-variance trade-off inherent to any estimation procedure is amplified when analyzing nonstationary stochastic processes by the fact that time-averaging, aimed at reducing variance, introduces some bias not only in the frequency direction but also in time; (*ii*) from a geometrical perspective, windowing — aimed at guaranteeing a form of local stationarity — ends up with a different kind of trade-off related to the time-frequency localization in the case of chirp-like signals. Such difficulties have been recognized long ago, and numerous studies have tried to address the problem. As far as localization is concerned, Wigner-based approaches have been developed and shown to outperform windowed (Fourier or wavelet-based) methods, at least in the case of noise-free single chirps [7]. In more realistic situations of multi-chirps, a dramatic improvement over both Fourier and Wigner-based methods has come from the use of the so-called *reassignment* technique [1], with an efficiency that is however limited to the cases where the signal-to-noise ratio is high enough. Turning to the estimation issue in a statistical sense, different attempts have been made to advantage of the idea of *multitapering*, pioneered by D.J. Thomson in a stationary setting [12], and thanks to which an improved statistical stability can be obtained without having recourse to a time-averaging step. As it has been extended to nonstationary situations, the "classical" method of multitapering suffers however still from the time-frequency localization trade-off mentioned above [4, 5, 9, 13]. Some attempts have been made to circumvent this limitation, in particular by identifying chirp-like components and excising them prior applying the multitaper machinery [3, 4]. The purpose of this paper is to avoid such a complication and to rather combine multitapering (for a sake of variance reduction) with reassignment (for localization).

2. NONSTATIONARY SPECTRUM ESTIMATION

Defining a time-varying "spectrum" for a nonstationary process $\{x(t), t \in \mathbb{R}\}$ is a question that has no unique answer [7]. Among the various possibilities stands first the *Wigner-Ville Spectrum* (WVS), whose definition reads:

$$\mathbf{W}_x(t,f) = \int_{-\infty}^{+\infty} \mathbb{E}\left\{x\left(t+\frac{\tau}{2}\right)x^*\left(t-\frac{\tau}{2}\right)\right\} e^{-i2\pi f\tau} d\tau,$$

where t and f refer to time and frequency, respectively, and $\mathbb{E}\{.\}$ stands for the expectation operator. This definition, though not unique, presents the advantage of extending the usual concept of *Power Spectrum Density* (PSD) and making it time-dependent in a rather natural way. Without entering into much details, it is worth recalling that the WVS reduces to the PSD for all times if the analyzed process happens to be stationary. Moreover, its marginal distributions are directly related to meaningful quantities (variance in time, Loève's distribution function in frequency) and it also satisfies the important property of preserving supports, in both time and frequency.

If we now introduce the (non random) quantity

$$W_x(t,f) = \int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right)\, x^*\left(t-\frac{\tau}{2}\right)\, e^{-i2\pi f\tau}\, d\tau,$$

which is referred to as the *Wigner-Ville Distribution* (WVD), it can be shown [7] that, under mild conditions, the WVS of a process is nothing but the ensemble average of the WVDs of all possible realizations of this process:

$$\mathbf{W}_x(t,f) = \mathbb{E}\left\{W_x(t,f)\right\}. \tag{1}$$

Given one observed realization of a nonstationary process, estimating the WVS amounts to find a substitute for the unattainable ensemble average entering eq.(1). One standard way is to assume for $x(t)$ a form of local stationarity in both time and frequency, i.e., some locally slow evolution of the actual WVS in the two directions. Such an assumption paves the road for a replacement of the ensemble average at a given time-frequency location by a local smoothing over a neighbouring domain. This idea can be formalized by introducing as a family of WVS estimators the quantity [7]:

$$\hat{\mathbf{W}}_x(t,f) = \iint_{-\infty}^{+\infty} W_x(s,\xi)\,\Pi(s-t,\xi-f)\,ds\,d\xi, \quad (2)$$

where $\Pi(t,f)$ is some time-frequency smoothing kernel.

This expression, which turns out to coincide with *Cohen's class* [7] (thereafter denoted $C_x(t,f;\Pi)$) for the observed realization, offers a unified setting in which the two trade-offs mentioned previously (regarding fluctuations and localization) appear clearly. If we consider the toy example of a linear chirp embedded in broadband noise, the fluctuations of the WVD due to noise will be smoothed out provided that the kernel in eq.(2) is low-pass. However, the WVD of the linear chirp (which has the unique property of being perfectly localized along the instantaneous frequency [7]) will be smoothed out too. A way out is however possible by reconsidering the apparently contradictory issues of fluctuations reduction and localization at the light of the two refinements offered by reassignment and multitapers.

2.1. Reassignment

Since the smoothing kernel in eq.(2) is *a priori* arbitrary, a possible choice is to make use of a function that would be itself the WVD of some signal $h(t)$ supposed to be reasonably well localized in both time and frequency, a property that carries over to the time-frequency plane thanks to the structure of the WVD. Doing so, it is easy to show that we end up with $C_x(t,f;W_h) = S_x^{(h)}(t,f)$, where $S_x^{(h)}(t,f)$ is nothing but the *spectrogram* of $x(t)$ with window $h(t)$, a time-frequency distribution that is usually rather expressed as:

$$S_x^{(h)}(t,f) = \left|F_x^{(h)}(t,f)\right|^2, \quad (3)$$

where

$$F_x^{(h)}(t,f) = \int_{-\infty}^{+\infty} x(s)\,h^*(s-t)\,e^{-i2\pi fs}\,ds$$

stands for the *Short-Time Fourier Transform* (STFT).

A spectrogram appears therefore as an estimator for the WVS, with the well-known time-frequency localization trade-off attached to this type of distribution: the shorter the window $h(t)$, the better the time localization, but the poorer the frequency localization, and *vice versa*. In this respect however, the alternative interpretation of the spectrogram as a smoothed WVD (according to eq.(2)) rather than as a squared STFT (according to eq.(3)) gives the clue for improving upon its localization limitations. Indeed, if we recall that the WVD of a linear chirp perfectly localizes on a time-frequency line, the spreading of any corresponding spectrogram just comes from the fact that, when centering the analysis window at some time-frequency point that does not belong to this line, a non-zero contribution is nevertheless obtained as long as the line passes through the local time-frequency window (whose joint support cannot be made arbitrarily small). Reasoning by a mechanical analogy identifying energy with mass, the situation is as if a whole distribution of mass within a domain (here, the time-frequency window) would be replaced by one single number (the total mass) assigned to the *geometrical* center of the domain. Such an assignment is clearly not well adapted to situations where the distribution is not uniform over the domain. In such cases, a much more meaningful assignment is the *center of mass* within the domain. This is precisely the essence of the reassignment technique, which consists in evaluating for each time-frequency location, not only the integrated signal WVD within the domain defined by the window WVD (in other words, the spectrogram value at this point), but also the center of mass of the signal WVD, a location that is possibly different from the window WVD center and to which the spectrogram value is reassigned. In the idealized case where one single linear chirp is "seen" through the time-frequency window, it is clear that the center of mass necessarily belongs to the line along which the WVD is localized, thus guaranteeing a perfect localization of the spectrogram after its reassignment.

Previous studies [1, 8] have shown that an efficient evaluation of the local centers of mass $(\hat{t}_{t,f}, \hat{f}_{t,f})$ can be made implicitly, according to

$$\begin{cases} \hat{t}_{t,f} &= t + \mathrm{Re}\{F_x^{(\mathcal{T}h)}(t,f)/F_x^{(h)}(t,f)\}; \\ \hat{f}_{t,f} &= f - \mathrm{Im}\{F_x^{(\mathcal{D}h)}(t,f)/F_x^{(h)}(t,f)\}, \end{cases}$$

where the two additional windows needed in the computation are defined from the mother window $h(t)$ as $(\mathcal{T}h)(t) = t\,h(t)$ and $(\mathcal{D}h)(t) = (dh/dt)(t)$. Given the field of all above centroïds, the reassigned spectrogram $RS_x^{(h)}(t,f)$ attached to the conventional spectrogram $S_x^{(h)}(t,f)$ follows as:

$$RS_x^{(h)}(t,f) = \iint_{-\infty}^{+\infty} S_x^{(h)}(s,\xi)\,\delta(t-\hat{t}_{s,\xi})\delta(f-\hat{f}_{s,\xi})dsd\xi.$$

2.2. Multitapers

In the case of a stationary process, the spectral characterization is fully described by means of the PSD $\mathbf{S}_x(f)$, which could be thought of as:

$$\mathbf{S}_x(f) = \lim_{T\to\infty} \mathbb{E}\left\{\frac{1}{T}\left|\int_{-T/2}^{+T/2} x(t)\,e^{-i2\pi ft}\,dt\right|^2\right\}.$$

In practice, the above quantity is unattainable when only one realization of finite duration is given. The Squared Fourier Transform (SFT) of a finite duration observation is a crude, non consistent, estimator, whose variance is of the order of the squared PSD [11]. Since an improvement can only come from an averaging of (almost) uncorrelated estimations, an ergodic argument suggests to first chop the observation into (almost) disjoint blocks and then average their

SFTs. This procedure is usually referred to as the Welch method of averaged periodograms [11]. Adopting the notation of spectrograms, it turns out that the corresponding (Welch) estimator can be written as:

$$\hat{\mathbf{S}}_{x,K}^{(W)}(f) = \frac{1}{K}\sum_{k=1}^{K} S_x^{(h)}(t_k, f),$$

where the spacing $\Delta = t_{k+1} - t_k$ between adjacent t_k's is of the order of the window length. Assuming that this spacing will ensure an approximate decorrelation between blocks, one can expect that the variance will be inversely proportional to the number K of blocks (i.e., roughly T/Δ for an observation of duration T). Whereas variance can be decreased this way by increasing K (to some extent), the finite duration constraint necessarily leads to shorten Δ, increasing in turn the bias in frequency since a window of duration Δ has a frequency resolution of the order of $1/\Delta$.

In order to circumvent this trade-off, D.J. Thomson suggested [12] to still average SFTs stemming from (almost) uncorrelated sequences in order to reduce variance, but to construct such sequences by using for each of them the whole data set so as to not sacrify bias. The way to achieve this program consists in projecting the observation on a family of basis functions $\{h_k(t), k \in \mathbb{Z}\}$ that are orthonormal over the observation interval. This results in a (Thomson) estimator that can be written as:

$$\hat{\mathbf{S}}_{x,K}^{(T)}(f) = \frac{1}{K}\sum_{k=1}^{K} S_x^{(h_k)}(0, f). \tag{4}$$

Assuming that the spectrum can be considered as flat over a given bandwidth B associated with the expected frequency resolution, the basis can be obtained as the family of orthonormal functions (on the given time interval) that maximize their energy in the given frequency band. The solution to this problem is given by the family of functions known as the *Prolate Spheroidal Wave Functions* or, in a discrete-time setting, as the *Discrete Prolate Spheroidal Sequences* (DPSS) [11].

Extending the approach to nonstationary situations is appealing [9, 13]. The main reason is that the inconsistency (and large variance) of a PSD estimator based on a crude SFT directly carries over to spectrograms considered as WVS estimators. The traditional way out would be to smooth over time and frequency, but at the expense of further increasing bias. In this respect, resorting to multitapers allows for a variance reduction with a bias that only sticks to the common length of the different windows. This is certainly an improvement as compared to spectrograms and smoothed spectrograms with respect to statistical efficiency, but the question of time-frequency resolution still remains not improved. Wedding multitapering with reassignment is therefore proposed as the key for such an improvement.

3. MULTITAPER TIME-FREQUENCY REASSIGNMENT

3.1. Principle and implementation

The direct application of multitapering to nonstationary processes consists in making the estimator (4) time-dependent according to [9, 13]:

$$\hat{\mathbf{S}}_{x,K}^{(T)}(f) \to S_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K} S_x^{(h_k)}(t, f).$$

What we propose here is to adopt the same strategy, but applied to reassigned spectrograms, i.e., to consider as a WVS estimator the quantity:

$$RS_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K} RS_x^{(h_k)}(t, f). \tag{5}$$

The rationale for this approach can be justified in a twofold way: (*i*) as far as chirp components are concerned, reassignment increases localization in a way that can be made independent of the window, thus permitting (5) to act as a *coherent* averaging; (*ii*) in noise regions on the contrary, the same windows lead to uncorrelated surrogate data whose time-frequency distributions are different, (5) acting in this case as a form of *incoherent* averaging tending to smooth the estimate.

In stationary spectrum estimation, multitapers are chosen as DPSSs because the data is of finite duration and estimation concerns frequency only. In the nonstationary case, there is no *a priori* reason to dissymmetrize time and frequency by choosing tapers that would be perfectly localized in the time domain rather than in frequency. Indeed, it makes much more sense to fully exploit the two degrees of freedom offered by the time-frequency plane and, as suggested in [3, 4], to rather pick up those functions that maximally concentrate in time-frequency domains with elliptic symmetry. As shown in [6], those functions are the *Hermite* functions (HF), whose definition is given by

$$h_k(t) = (-1)^k \frac{1}{\sqrt{\pi^{1/2}2^k k!}}\, g(t)\, (\mathcal{D}^k \gamma)(t),$$

with $g(t) := \exp\{-t^2/2\}$ and $\gamma(t) := g(it\sqrt{2}) = \exp\{t^2\}$. From a practical point of view, HFs can be computed recursively, according to $h_k(t) = H_k(t)\, g(t)/\sqrt{\pi^{1/2}2^k k!}$, where the $\{H_k(t), k \in \mathbb{N}\}$ stand for the Hermite polynomials that obey the recursion:

$$H_k(t) = 2\,t\,H_{k-1}(t) - 2\,(k-2)\,H_{k-2}(t), k \geq 2,$$

with the initialization $H_0(t) = 1$ and $H_1(t) = 2\,t$.

Not only the HFs are orthonormal, but they also guarantee a perfect localization of the corresponding reassigned spectrograms in the case of a linear chirp, for any k. This can be easily understood by noting that the WVD of a HF (which is basically a 2D Laguerre function) has elliptic symmetry [4, 6]. Recalling that the WVD is covariant with respect to dilations and rotations, it is enough to check that reassignment ends up with a perfect localization in the case of a pure tone, what can be done by an elementary calculation.

In the context of reassignment, HFs offer one further advantage, as compared to DPSSs. In the standard implementation of spectrogram reassignment, only the mother window $h(t)$ has to be given and the two additional windows $(\mathcal{T}h)(t)$ and $(\mathcal{D}h)(t)$ that are needed are evaluated numerically [2, 8]. This may cause difficulties, especially when differentiating tapers whose order k is large, since they are highly oscillating. This problem can be easily avoided when using HFs since their successive derivatives obey a recursion that can be explicitly plugged in the algorithm, namely

$$(\mathcal{D}h_k)(t) = (\mathcal{T}h_k)(t) - \sqrt{2(k+1)}\, h_{k+1}(t).$$

3.2. Examples and performance evaluation

Since the objective of multitaper time-frequency reassignment is to decrease fluctuations while preserving localization, our first example is concerned with the idealized situation of a bandpass filtered Gaussian white noise within a time-limited support. Although not strictly attainable (because of the uncertainty relation), the model $M(t,f)$ for the WVS of such an observation is the indicator function of a rectangle within the time-frequency plane. Figure 1 illustrates what happens in such a case by comparing the WVD and a sample (reassigned) spectrogram with the corresponding multitaper estimates based on $K = 10$ Hermite functions. The two effects of reduced fluctuations and support preservation are clearly visible, and ensemble averages (based on 10 independent estimates) are also provided for supporting the effectiveness of the approach and its improved convergence rate as compared to an empirical estimate of the WVS.

Figure 1 gives a qualitative account of the behaviour of the method, that can be supplemented by the more quantitative performance measure

$$E(K) = \frac{1}{\|M\|_1} \iint_{-\infty}^{+\infty} \left| \hat{\mathbf{W}}_x(t,f) - M(t,f) \right| \, dt\, df, \tag{6}$$

where $\hat{\mathbf{W}}_x(t,f)$ stands for the WVS estimate ($S_{x,K}(t,f)$ or $RS_{x,K}(t,f)$), the L_1-norm being here chosen so as to put emphasis on localization in the estimates. Figure 2 presents results with different domains, all rectangular and centered in the analyzed time-frequency region, but with different areas D. In the pure white noise case where the model support identifies with the whole plane (in this case, $D = 256$), we observe for both spectrograms and reassigned spectrograms that the error measure behaves asymptotically as $E(K) \propto K^{-1/2}$ when using K tapers. In the spectrogram case, this can be justified by the fact that, for each taper, the values are known to have a χ^2 distribution with 2 degrees of freedom [11]. It then follows from the orthogonality of the tapers that the sum of the K first (Hermite) spectrograms is also χ^2 distributed, but with $2K$ degrees of freedom. Such a distribution can be shown to have for absolute deviation $4K^K \exp(-K)/\Gamma(K)$, a quantity which varies as $K^{1/2}$ for large K's (from Stirling's formula), leading to the $K^{-1/2}$ behavior for the mean $S_{x,K}(t,f)$. Although no proof is available yet, the experiments reported in Figure 2 evidence a similar behaviour for multitaper reassigned spectrograms $RS_{x,K}(t,f)$, but with a higher level of fluctuations. However, when the area of the domain D is reduced, the situation evolves quite differently for the two estimates: on the one hand, $S_{x,K}$ is smoother that $RS_{x,K}$; on the other hand, $RS_{x,K}$ is essentially confined to the non-zero support of the model, whereas $S_{x,K}$ spreads outside this domain. The criterion (6) is a measure of this bias-fluctuations trade-off that is illustrated in Figure 2 for smaller and smaller domains, evidencing eventually crossings indicating that conventional multitaper spectrograms may be outperformed by their reassigned counterparts when localized components are to be analyzed.

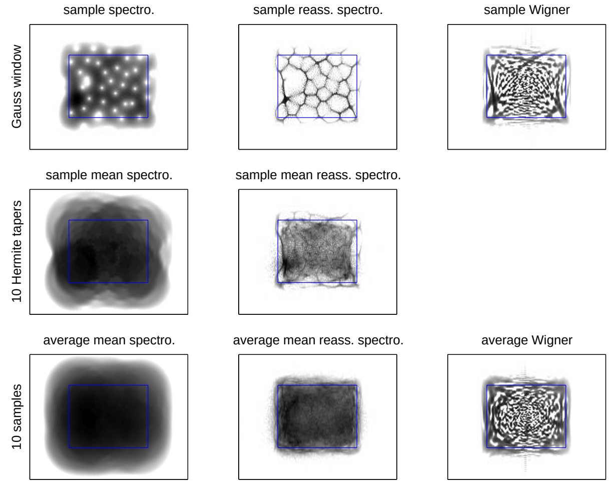

Fig. 1. Comparison of noise WVS estimates — *Each diagram represents a WVS estimate in the case of a white noise process limited in time and frequency within the superimposed rectangular domain (ideally, the estimate should be constant over this rectangle and zero outside). The first row consists of a spectrogram, its reassigned version and the WVD, based on one realization. The corresponding multitaper estimates (*10 *Hermite functions) are given in the middle row, whereas the bottom row displays ensemble averages of such estimates (*10 *independent realizations), together with the empirical WVS estimate on the same data set. In each diagram, time is horizontal, frequency vertical, and the energy is coded with gray levels on a logarithmic scale with a dynamic range of* 30 *dB.*

In this respect, we consider in Figure 3 the case already discussed in [4] and [5], with both a (nonlinear) chirp component and a (bandpass) time-varying noise. The effectiveness of the approach is clearly supported by this example which evidences the good trade-off achieved between time-frequency localization along the chirp and smoothness within the (time-varying) frequency band of the noise.

4. CONCLUSION AND PERSPECTIVES

A novel approach, combining reassignment and multitapering, has been proposed for better estimating time-varying spectra with possibly localized components. Due to space

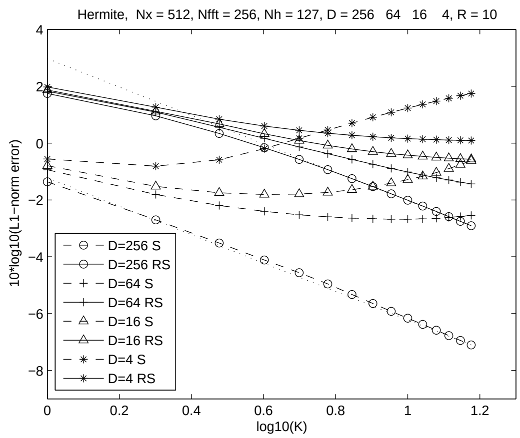

Fig. 2. Error measures in WVS multitaper estimates — *The figure plots, as a function of the number K of tapers, the error measure (6) attached to multitaper (reassigned) spectrograms when the model is a Gaussian white noise process limited in time and frequency over a rectangular domain of area D. The simulations have been conducted (with up to $K = 30$ Hermite tapers, each of length $Nh = 127$) on the basis of $R = 10$ independent realizations of $N = 512$ data points each, with $Nfft = 256$ frequency bins over the whole frequency range* $[0, 1/2)$. *In the pure white noise situation (which corresponds in the present case to the area $D = 256$), asymptotic decays in $K^{-1/2}$ (see text) have been superimposed as dotted lines.*

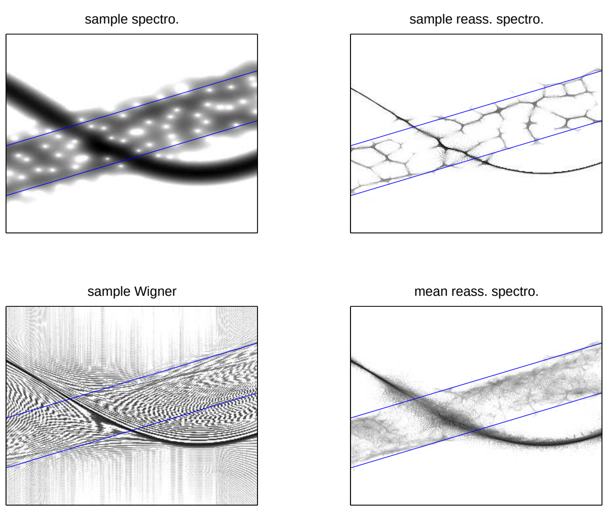

Fig. 3. Comparison of signal+noise WVS estimates — *Each diagram represents a WVS estimate in the case of a nonlinear chirp signal embedded in a bandpass time-varying noise limited within the superimposed frequency band (ideally, the estimate should be constant over this band, zero outside and perfectly localized along the chirp instantaneous frequency). The first row consists of a spectrogram and its reassigned version, based on one realization. The corresponding multitaper estimate (*6 *Hermite functions) is given in the bottom row (right), with the WVD (left) for comparison. In each diagram, time is horizontal, frequency vertical, and the energy is coded with gray levels on a logarithmic scale with a dynamic range of* 30 *dB.*

limitation, only the basic principles of the method have been outlined, with no reference to many issues and variations that can be envisioned. For instance, only the simplest way (arithmetic averaging) of combining estimates with different tapers has been considered here, but other types of averaging are possible (see, e.g., [10]) as well as refinements such as jacknifing the estimates [13], that might improve upon the performance. Such developments are under current investigation and will be reported elsewhere.

5. REFERENCES

[1] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Proc.*, vol. SP-43(5), pp. 1068–1089, 1995.

[2] F. Auger, P. Flandrin, P. Gonçalvès and O. Lemoine, *Time-Frequency Toolbox for* MATLAB. Freeware available at the URL `http://tftb.nongnu.org`.

[3] M. Bayram and R.G. Baraniuk, "Multiple window time-frequency analysis," in *Proc. IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, (Paris, France), pp. 511–514, 1996.

[4] M. Bayram and R.G. Baraniuk, "Multiple window time-varying spectrum estimation," in *Nonlinear and Nonstationary Signal Processing* (W.J. Fitzgerald et al., eds.), pp. 292–316, Cambridge Univ. Press, 2000.

[5] F. Çakrak and P.J. Loughlin, "Multiple window time-varying spectral analysis," *IEEE Trans. Signal Proc.*, vol. SP-49(2), pp. 448–453, 2001.

[6] P. Flandrin, "Maximum signal energy concentration in a time-frequency domain," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc. — ICASSP'88*, (New York, NY), pp. 2176–2179, 1988.

[7] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.

[8] P. Flandrin, F. Auger and E. Chassande-Mottin, "Time-Frequency Reassignment — From Principles to Algorithms," in *Applications in Time-Frequency Signal Processing* (A. Papandreou-Suppappola, ed.), Chap. 5, pp. 179-203, CRC Press, 2003.

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

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

[11] D.B. Percival and A.T. Walden, *Spectral Analysis for the Physical Sciences*, Cambridge Univ. Press, 1993.

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

[13] D.J. Thomson, "Multitaper analysis of nonstationary and nonlinear time series data," in *Nonlinear and Nonstationary Signal Processing* (W.J. Fitzgerald et al., eds.), pp. 317–394, Cambridge Univ. Press, 2000.