Invited paper

On Spectral Noncircularity of Natural Signals

Scott Wisdom[1], Les Atlas[1], and James Pitton[1,2]
[1]Department of Electrical Engineering, University of Washington, Seattle, WA, USA
[2]Applied Physics Laboratory, University of Washington, Seattle, WA, USA
{swisdom,atlas}@uw.edu, pitton@apl.uw.edu

***Abstract*—Natural signals are typically nonstationary. The complex-valued frequency spectra of nonstationary signals do not have zero spectral correlation, as is assumed for wide-sense stationary processes. Instead, these spectra have non-zero second-order noncircular statistics–that is, they are not rotationally invariant–that are potentially useful for detection, classification, and enhancement. These noncircular statistics are especially significant for transient events, which are common in many natural signals. In this paper we provide practical and effective estimators for spectral noncircularity and spectral correlation. We illustrate the behavior of our spectral noncircularity estimators for synthetic signals. Then, we derive a generalized likelihood ratio test using both circular and noncircular models and show how estimates of spectral noncircularity provide performance improvements for detection of natural acoustic events.**

I. Introduction and related work

Physical signals that arise naturally, such as speech, environmental sounds, and machinery vibrations are obviously real-valued. Yet Fourier transforms of these natural physical signals are complex-valued. In this paper, we show that the complex spectra of these signals can be beneficially characterized by considering whether their second-order statistics are noncircular. That is, rotations of the complex quantities will change their probability distributions [1]. We thus propose that spectral noncircularity is a useful feature for detection, enhancement, and classification of natural signals. This paper details the case for detection, especially when the spectral noncircularity of the natural signal of interest is substantial compared to the circularity of the background noise.

We give explicit expressions and illustrations for the auto- and cross-spectral correlation of several models of real-valued natural signals, including wide-sense stationary (WSS) noise and deterministic signals of any bandwidth. We also derive practical estimators for spectral noncircularity and show how these estimators are useful for acoustic event detection in natural data.

Schreier and Scharf [2] were among the first to recognize that many real-valued nonstationary random signals exhibit second-order noncircularity (or "impropriety") in their analytic signals. Furthermore, they also examined the bifrequency Loève spectrum of such signals. Atlas [3] considered bifrequency spectral correlations and their application to separation and modification of speech from the perspective of modulation frequency. Douglas and Mandic proposed using the "panorama" [4], which they defined as the Fourier transform of a zero-mean signal's autoconvolution estimated using ensemble averaging. In concert with the conventional power spectrum–the Fourier transform of a zero-mean signal's autocorrelation–they proposed a detector for deterministic sinusoidal components in the presence of WSS noise. We go beyond previous work by considering both the auto- and cross-spectral statistics directly for a wider variety of deterministic signals. Additionally, our estimators of spectral noncircularity do not require ensemble averaging.

Millioz and Martin [5] studied the circularity of the short-time Fourier transform (STFT) coefficients of nonstationary signals in white Gaussian noise, and used noncircularity to segment synthetic and natural signals in time-frequency. Clark [6] observed that speech signals exhibit spectral noncircularity. Spectral noncircularity is also closely related to the modulation frequency content of signals, which can correspond to syllabic rates of speech [7], [8]. Wisdom et al. [9] and Okopal et al. [10], [11] empirically showed that spectral noncircularity is useful for detection of speech signals and can be used for blind separation of multichannel mixtures of speech and noise.

II. Background

A. Noncircularity and correlation of complex-valued random data

A zero-mean complex-valued random variable (RV) x requires two second order moments [12]: Hermitian covariance, defined as $R_x \triangleq E\{|x|^2\} = E\{xx^*\}$, and complementary covariance, defined as $\widetilde{R}_x \triangleq E\{x^2\} = E\{xx\}$. The RV x is called *second-order circular* if and only if $\widetilde{R}_x = 0$. The degree of noncircularity[1] of x is defined as

$$\kappa_x \triangleq \frac{|\widetilde{R}_x|}{R_x}, \tag{1}$$

and takes on values between 0 and 1.

B. Spectral representation and correlations

The Cramér-Loève spectral representation of a random process $x(t)$ is [13]

$$x(t) = \int_{-\infty}^{\infty} e^{j2\pi ft} dX(f), \tag{2}$$

where $dX(f)$ is a complex-valued spectral increment process. Since $dX(f)$ is complex valued, it requires two second-order statistics. Thus, these increment processes have bifrequency Hermitian spectral correlation (HSC) and complementary spectral correlation (CSC) which are, respectively,

$$S_{xx}^{(bf)}(f_1, f_2) df_1 df_2 \triangleq E[dX(f_1) dX^*(f_2)], \tag{3}$$

$$\widetilde{S}_{xx}^{(bf)}(f_1, f_2) df_1 df_2 \triangleq E[dX(f_1) dX(f_2)]. \tag{4}$$

This work is funded by U.S. Office of Naval Research contract N00014-12-G-0078, delivery order 0013, and U.S. Army Research Office grant number W911NF-15-1-0450.

[1]Throughout the remainder of this paper, *circular* will refer to second-order circularity.

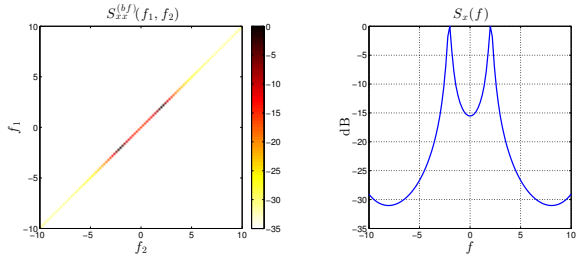

Fig. 1. Theoretical bifrequency Hermitian spectral correlation (HSC) and power spectral density (PSD) for a real-valued wide-sense stationary (WSS) random process. Left panel: bifrequency HSC; right panel: conventional PSD $S_x(f)$, equivalent to $S_{xx}^{(bf)}(f, f)$.

III. Theoretical spectral impropriety

This section presents the theoretical spectral noncircularity and spectral correlations of several models of natural signals.

A. Wide-sense stationary random processes

A WSS random process has orthogonal increments [14], which means that a WSS process exhibits no spectral correlation between different frequencies. That is,

$$S_{xx}^{(bf)}(f_1, f_2) = 0, \text{ for } f_1 \neq f_2 \tag{5}$$

Furthermore, a WSS process's spectral increments are circular [12], which means that the process has zero complementary spectral correlation everywhere. That is,

$$\widetilde{S}_{x}^{(bf)}(f_1, f_2) = 0, \text{ for all } f_1, f_2. \tag{6}$$

Thus, a WSS process is completely characterized by its power spectral density (PSD) along the "stationary manifold," $f_1 = f_2$ [2]. The PSD is thus

$$S_x(f) \triangleq S_{xx}^{(bf)}(f_1 = f, f_2 = f). \tag{7}$$

Figure 1 illustrates these concepts. The theoretical bifrequency HSC (left panel) and conventional PSD (right panel) are plotted for an example WSS second-order autoregressive process.

B. Deterministic signal in WSS noise

Consider a deterministic signal $p(t)$ in additive WSS noise:

$$x(t) = p(t - t_0) + v(t), \tag{8}$$

where t_0 is a deterministic delay. In this case, the spectral increments are

$$dX(f) = P(f)e^{-j2\pi f t_0} df + dV(f), \tag{9}$$

and the HSC and CSC along the stationary manifold are, respectively,

$$S_x(f) = S_{xx}^{(bf)}(f, f) = |P(f)|^2 + S_v(f), \tag{10}$$

$$\widetilde{S}_x(f) = \widetilde{S}_{xx}^{(bf)}(f, f) = P^2(f)\exp(-j4\pi f t_0). \tag{11}$$

Thus, along the stationary manifold, $x(t)$ exhibits spectral noncircularity at all frequencies where $P(f)$ has non-zero power, with degree of noncircularity

$$\kappa_x(f) = \frac{|P(f)|^2}{|P(f)|^2 + S_v(f)} = \frac{\mathrm{SNR}(f)}{\mathrm{SNR}(f) + 1}, \tag{12}$$

where $\mathrm{SNR}(f) \triangleq \frac{|P(f)|^2}{S_v(f)}$ is the frequency-dependent power ratio between the deterministic signal and the random noise.

Off the stationary manifold where $f_1 \neq f_2$, the HSC is only nonzero at all combinations of frequencies where $P(f_1)$ and $P(f_2)$ are nonzero. The HSC and CSC are

$$S_{xx}^{(bf)}(f_1, f_2) = P(f_1)P^*(f_2), \tag{13}$$

$$\widetilde{S}_{xx}^{(bf)}(f_1, f_2) = P(f_1)P(f_2). \tag{14}$$

C. Random signals

Certain random signals also exhibit spectral noncircularity and cross-spectral correlation. For example, WSS noise that is amplitude-modulated by a sum of sinusoids, which are good models of rhythmic signals like ship noise, are spectrally noncircular at multiples of the modulation frequencies [6], [7], [15], [16]. Wisdom showed [17] that jittered pulse trains, which can be used to model glottal excitations in voiced human speech, exhibit spectral noncircularity at harmonic multiples of the fundamental frequency.

IV. Estimating spectral noncircularity

In this section we propose methods to estimate spectral correlations and spectral noncircularity in practice. Assume that N samples, x_n, of the random process are measured. In this case, applying an analysis window $h \in \mathbb{R}^N$ to x, computing the discrete spectrum X_m using a N-length FFT, and taking the magnitude-squared corresponds to a direct spectral estimate $\hat{S}_x^{(D)}(f_m)$ [14], which is equivalent to an estimate of the Hermitian spectral correlation $S_{xx}^{(bf)}(f_m, f_m)$ along the stationary manifold at the discrete Fourier frequencies f_m.

Unfortunately, estimating the complementary spectral correlation is not as simple as just squaring X_m (i.e., multiplying X_m by itself), since this would produce an estimator that is maximally noncircular at every frequency. An effective estimator requires some amount of averaging over multiple realizations.

Since only one realization is available, ensemble averaging is impossible. Instead, multiple realizations can be acquired in one of two ways. First, x can be segmented into L shorter chunks and we can average in time over the spectra of the chunks to compute the statistics[2], which is akin to Welch's method [18]. Alternatively, by using L orthogonal multitapers h_n^ℓ such that $\langle h_n^{\ell_1}, h_n^{\ell_2} \rangle = 0$ for $\ell_1 \neq \ell_2$, we can acquire L independent observations of the spectra: $X_m^\ell \triangleq \mathrm{FFT}\left\{h^\ell \odot x\right\}$, where m is discrete frequency index and ℓ is taper index. A multitaper estimator (a method originally proposed by Thomson [19]) averages across frequency, with the bandwidth increasing as the number of tapers increases. The standard choice for the multitapers are the discrete prolate spheroidal sequences (DPSS), which maximize the spectral concentration of the windows h_ℓ for a given bandwidth W [19].

Using either Welch's or Thomson's multitaper method, the

[2] As long as the deterministic phase caused by window hops is removed by multiplying X_m^ℓ by $e^{j2\pi f \ell N_{hop}}$.

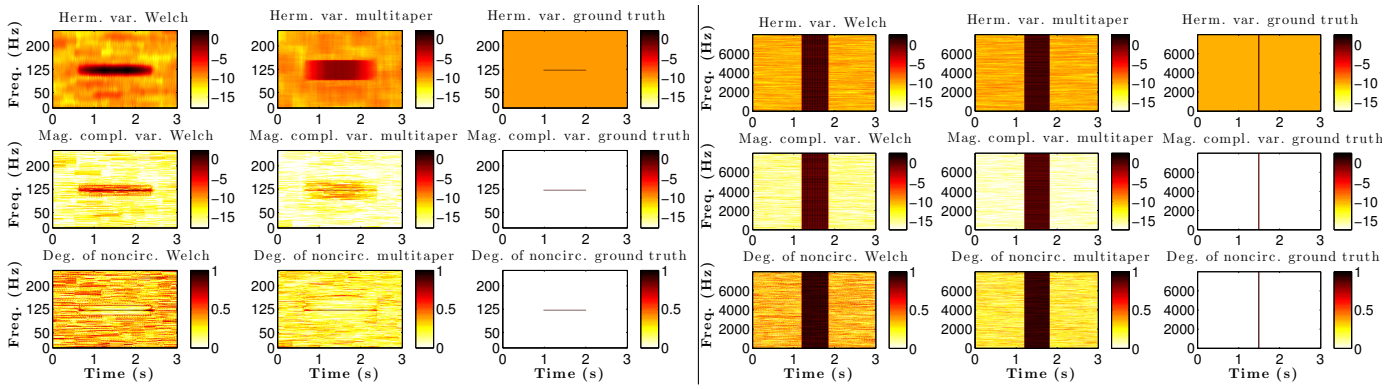

Fig. 2. Outputs of Welch and multitaper estimators of spectral Hermitian variance (top row), complementary variance (middle row), and degree of noncircularity (bottom row) for two different synthetic test signals in 10dB SNR white noise: a sinusoid (left panel) and a Kronecker delta function (right panel). The ideal ground truth is given in the right columns of the panels. Both estimators use the same bandwidth of $2/(64\text{ms}) = 31.25\text{Hz}$.

estimators for HSC and CSC at frequencies f_{m_1}, f_{m_2} are

$$\hat{S}_{xx}^{(bf)}(f_{m_1}, f_{m_2}) = \frac{1}{L}\sum_{\ell=0}^{L-1} c_\ell X_{m_1}^\ell (X_{m_2}^\ell)^*, \tag{15}$$

$$\hat{\tilde{S}}_{xx}^{(bf)}(f_{m_1}, f_{m_2}) = \frac{1}{L}\sum_{\ell=0}^{L-1} c_\ell X_{m_1}^\ell X_{m_2}^\ell \tag{16}$$

where $c_\ell = 1$ for all ℓ when using Welch's method and when using the multitaper method, c_ℓ are weights corresponding to the eigenvalues of the tapers. When $m_1 = m_2 = m$, by the triangle inequality these estimators satisfy the property $\left|\hat{\tilde{S}}_{xx}^{(bf)}(f_m, f_m)\right| \leq \hat{S}_{xx}^{(bf)}(f_m, f_m)$ for all m. A version of the multitaper CSC estimator along the stationary manifold was originally proposed by Clark et al. [20].

Using (15) and (16), the spectral noncircularity estimator is

$$\hat{\kappa}_x(f_m) = \frac{\left|\hat{\tilde{S}}_{xx}^{(bf)}(f_m, f_m)\right|}{\hat{S}_{xx}^{(bf)}(f_m, f_m)}. \tag{17}$$

If the samples $X_m^\ell/\sqrt{c^\ell}$ are assumed to be distributed as i.i.d. complex-valued Gaussians, Delmas et al. [21, Remark 5] provide an approximate finite-sample distribution for the estimator of the degree of noncircularity $\hat{\kappa}$ in (17). According to this approximation, the distribution of $\hat{\kappa}$ converges as $L \to \infty$ to a Rayleigh distribution with scale $L^{-1/2}$ when $\kappa = 0$ and to a normal distribution with mean κ and variance $\sigma_\kappa^2 = L^{-1/2}(1-\kappa^2)^2$ when $\kappa > 0$. In practice, this approximation seems to hold for L as low as 10 samples [6, Appendix C]. Using this approximate finite-sample distribution, we can define a threshold T for deciding whether a frequency bin is noncircular for a given probability of false alarm P_{FA}.

V. Experiments

A. Synthetic data

Two simple test signals to verify these estimators are a sinusoid at frequency $f_0 = 125$ Hz of duration 1 second and a Kronecker delta function. Both signals are embedded in white noise at a SNR of 10dB. The sampling frequency is 16kHz and frames of length 64ms are used for both the Welch's and multitaper estimators. Statistics are computed in 8ms increments using $L = 16$ frames at a time. The Welch's estimator uses a Hamming window with a hop of 32ms, and the multitaper estimator uses 16 DPSS tapers with duration equal to the total duration of the Welch's averaging region, which is $L \cdot (\text{hop}) = 16(32\text{ms}) = 512\text{ms}$. Figure 2 shows the results. For reference, we give ground truth time-frequency plots that are the output of an ideal estimator using our results from §III.B. Aside from smoothing, both estimators appear to work relatively well.

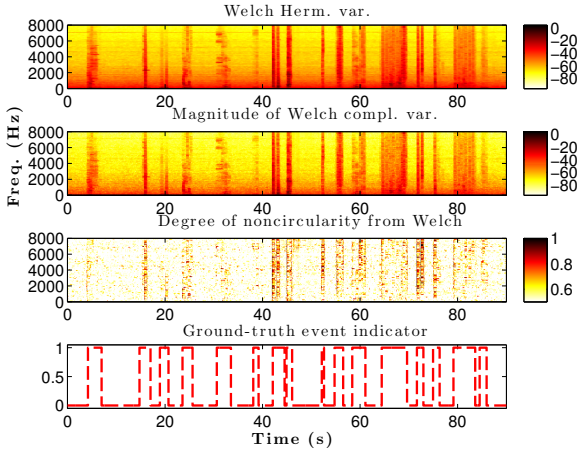

Fig. 3. Hermitian and complementary variance estimates and estimated degree of spectral noncircularity for a DCASE audio file.

B. Detection of acoustic events

To demonstrate the usefulness of spectral noncircularity for natural data, we consider the task of acoustic event detection using the publicly-available Detection and Classification of Acoustic Scenes and Events (DCASE) dataset [22]. In this dataset, 16 different types of acoustic events characteristic of a typical office environment–including beeps, throat clearing, coughs, door slams, keyboard clicks, keys dropped on a table, knocks, laughter, mouse clicks, page turns, pen drops, phones, printers, speech, and switches–are embedded in stationary background noise at SNRs of -6, 0, and 6 dB. Given a recording y, our goal is to detect when one or more acoustic events are active, in 8ms increments.

Figure 3 shows time-frequency plots for an example DCASE audio file at 0dB SNR. The sampling frequency is 16kHz, the Welch's method window is a 64ms Hamming window with a hop of 32ms, and statistics are computed in 8ms increments using $L = 16$ STFT frames at a time. The total duration of the Welch's sliding window and of the multitaper windows is thus $L \cdot (\text{hop}) = (16)(32\text{ms}) = 512\text{ms}$. Notice that the degree of spectral noncircularity is highest when acoustic events occur, especially transient events.

We test four methods of detection on this data: a baseline energy detector and three detectors based on a generalized likelihood ratio test (GLRT) using either a circular or noncircular model of the unknown signal. In general, if θ_1 are unknown parameters under $\mathcal{H}_1$ and θ_0 are unknown parameters under $\mathcal{H}_0$, a GLRT is defined as the ratio of the likelihoods under the two hypotheses, using the maximum-likelihood estimates of the unknown parameters θ_1 and θ_0:

$$G(y) \triangleq \left[\max_{\theta_1} p_{\mathcal{H}_1}(y|\theta_1)\right] / \left[\max_{\theta_0} p_{\mathcal{H}_0}(y|\theta_0)\right]. \tag{18}$$

Here, the detection problem tests between the two hypotheses

$$\begin{aligned} \mathcal{H}_0 &: Y_t = V_t, \\ \mathcal{H}_1 &: Y_t = X_t + V_t. \end{aligned} \tag{19}$$

Under the null hypothesis $\mathcal{H}_0$, a complex-valued STFT frame Y_t consists of only zero-mean circular Gaussian noise V_t with known diagonal covariance R_v. Under the alternative hypothesis $\mathcal{H}_1$, Y_t is V_t plus some zero-mean Gaussian signal X_t with unknown diagonal Hermitian covariance R_x. The signal's diagonal complementary covariance $\widetilde{R}_x$ is known and equal to 0 under the circular model, and is unknown under the noncircular model. We use the probability density function for L samples of a zero-mean noncircular Gaussian, which is

$$p(y; R_y, \widetilde{R}_y) = \frac{\exp\left(\frac{-\sum_{\ell=0}^{L-1} |y_\ell|^2 + \mathrm{Re}\left\{\frac{\widetilde{R}_y}{R_y} \sum_{\ell=0}^{L-1} y_\ell^2\right\}}{R_y(1-\kappa_y^2)}\right)}{\left(\pi R_y \sqrt{1-\kappa_y^2}\right)^L}. \tag{20}$$

The estimate of the noise spectrum $R_v(f)$ is computed from the first 4 seconds of data, which we know *a priori* to only contain noise. Using the noise estimate $\hat{R}_v(f)$, the log detection statistic for the circular GLRT is proportional to

$$\log G_C(Y_t) = \sum_f \left(\frac{\hat{R}_y(f,t)}{\hat{R}_v(f)} - \log \frac{\hat{R}_y(f,t)}{\hat{R}_v(f)} \right), \tag{21}$$

where $\hat{R}_y(f,t) \geq \hat{R}_v(f)$ is the estimated Hermitian variance of the observed data Y_t, and constants have been absorbed into the detection threshold. The first term in equation (21) defines an energy detector, which will be used as a baseline. For the noncircular GLRT, the log detection statistic is

$$\log G_{NC}(Y_t) = \log G_C(Y_t) - \frac{1}{2} \sum_f \log\left(1 - |\hat{\kappa}_y(f,t)|^2\right), \tag{22}$$

where $\hat{\kappa}_y(f,t)$ is the ratio of the estimated complementary variance magnitude $|\hat{\widetilde{R}}_y(f,t)|$ to $\hat{R}_y(f,t)$.

If the noise Hermitian variance $R_v(f)$ is not known, under a circular signal model no detector exists, since the problem has become ill-defined. However, if we continue to assume that the WSS noise is circular and that the nonstationary signal is noncircular at least some frequencies, we can design a blind noncircular detector that does not require an estimate of the Hermitian variance of the noise. Under these assumptions, the maximum likelihood estimate of the Hermitian variance under both $\mathcal{H}_0$ and $\mathcal{H}_1$ is $\hat{R}_y(f,t)$. Because of this, in (22), $\log G_C(Y_t)$ becomes 1, and the log GLRT is

$$\log G_{NC}^{(\text{blind})}(Y_t) = -\sum_f \log\left(1 - |\hat{\kappa}_y(f,t)|^2\right). \tag{23}$$

Using Delmas et al.'s approximate finite-sample distribution for κ [21], we can determine a threshold with a desired probability of false alarm.

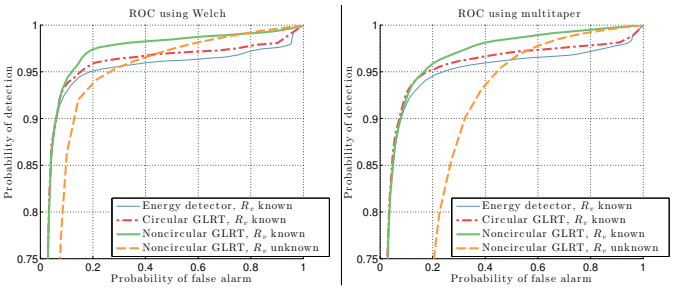

Fig. 4. Receiver operating curves for DCASE acoustic event detection using Welch's estimator (left panel) and multitaper estimator (right panel). Note that the Welch blind noncircular GLRT assuming unknown noise variance R_v (left panel: thick dashed line) performs nearly as well as the non-blind detectors that do use an estimate of the noise variance R_v (left panel: thin solid, dot-dash, and thick solid lines).

Receiver operating curves (ROCs) are shown in figure 4 for the four detectors using either the Welch or multitaper estimators. Notice that the noncircular non-blind detector improves over the circular non-blind detector for lower SNRs (i.e., when the detection threshold is lower). Also, interestingly, despite not using an estimate of the noise variance, the blind noncircular detector using the Welch estimator comes close to achieving the performance of the non-blind detectors. The area under the ROC curve (AUC) for the energy detector, non-blind circular, non-blind noncircular, and blind noncircular detectors is 0.94, 0.95, 0.96, and 0.92, respectively, using the Welch estimator, and 0.94, 0.95, 0.96, and 0.83, respectively, using the multitaper estimator.

VI. Conclusion

We have given closed-form expressions for the Hermitian and complementary correlation of the spectral increments of models of nonstationary natural signals. We proposed practical estimators for these statistics, and we demonstrated that spectral noncircularity can improve detection of realistic acoustic events. Furthermore, spectral noncircularity can be used to design a blind detector of such events that does not require an estimate of the stationary background noise power spectrum, yet does not suffer much performance degradation compared to non-blind detectors.

Future opportunities include exploitation of cross-frequency spectral correlations and the spectral noncircularity of multichannel data for natural signals.

REFERENCES

[1] B. Picinbono, "On circularity," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[2] P. Schreier and L. Scharf, "Stochastic time-frequency analysis using the analytic signal: why the complementary distribution matters," *IEEE Transactions on Signal Processing*, vol. 51, no. 12, pp. 3071–3079, 2003.

[3] L. Atlas, "Modulation Spectral Transforms–Application to Speech Separation and Modification," in *Speech Dynamics by Ear, Eye, Mouth, and Machine, An Interdisciplinary Workshop, Technical Report of IEICE*, Kyoto, Japan, Jun. 2003. [Online]. Available: http://isdl.ee.washington.edu/papers/atlas-2003-ieice.pdf

[4] S. Douglas and D. Mandic, "Autoconvolution and panorama: Augmenting second-order signal analysis," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 384–388.

[5] F. Millioz and N. Martin, "Circularity of the STFT and Spectral Kurtosis for Time-Frequency Segmentation in Gaussian Environment," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 515–524, 2011.

[6] P. Clark, "Coherent Demodulation of Nonstationary Random Processes," Ph.D. dissertation, University of Washington, 2012.

[7] P. Clark, I. Kirsteins, and L. Atlas, "Existence and estimation of impropriety in real rhythmic signals," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 3713–3716.

[8] S. Wisdom, J. Pitton, and L. Atlas, "Extending Coherence for Optimal Detection of Nonstationary Harmonic Signals," in *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2014.

[9] S. Wisdom, G. Okopal, L. Atlas, and J. Pitton, "Voice Activity Detection Using Subband Noncircularity," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Austrailia, Apr. 2015.

[10] G. Okopal, S. Wisdom, and L. Atlas, "Estimating the Noncircularity of Latent Components within Complex-Valued Subband Mixtures with Applications to Speech Processing," in *Proceedings of Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2014.

[11] ——, "Speech Analysis With the Strong Uncorrelating Transform," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 11, pp. 1858–1868, Nov. 2015.

[12] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge University Press, Feb. 2010.

[13] M. Loève, *Probability Theory*. New York: Springer-Verlag, 1977.

[14] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge University Press, Jun. 1993.

[15] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Processing*, vol. 86, no. 4, pp. 639–697, Apr. 2006.

[16] A. Napolitano, *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*. John Wiley & Sons, Aug. 2012.

[17] S. Wisdom, "Improved Statistical Signal Processing of Nonstationary Random Processes Using Time-Warping," Master's Thesis, University of Washington, Seattle, WA, Mar. 2014.

[18] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[19] D. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055–1096, Sep. 1982.

[20] P. Clark, I. Kirsteins, and L. Atlas, "Complementary envelope estimation for frequency-modulated random signals," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2013, pp. 5363–5367.

[21] J.-P. Delmas, A. Oukaci, and P. Chevalier, "On the asymptotic distribution of GLR for impropriety of complex signals," *Signal Processing*, vol. 91, no. 10, pp. 2259–2267, Oct. 2011.

[22] D. Stowell, D. Giannoulis, E. Benetos, M. Lagrange, and M. Plumbley, "Detection and Classification of Acoustic Scenes and Events," *IEEE Transactions on Multimedia*, vol. 17, no. 10, pp. 1733–1746, Oct. 2015.