EXTENDING COHERENCE TIME FOR ANALYSIS OF MODULATED RANDOM PROCESSES

Scott Wisdom⋆*, Les Atlas*⋆*, and James Pitton*†⋆

⋆Electrical Engineering Department, University of Washington, Seattle, USA
†Applied Physics Laboratory, University of Washington, Seattle, USA

ABSTRACT

In this paper, we relax a commonly-used assumption about a class of nonstationary random processes composed of modulated wide-sense stationary random processes: that the fundamental frequency of the modulator is stationary within the analysis window. To compensate for the relaxation of this assumption, we define the generalized DEMON ("demodulated noise") spectrum representing modulation frequency, which we use to increase the coherence time of such signals. Increased coherence time means longer analysis windows, which provides higher SNR estimators. We use the example of detection on both synthetic and real-world passive sonar signals to demonstrate this increase.

Index Terms— DEMON spectrum, coherence time, nonstationary random processes, spectral impropriety, modulation frequency

1. INTRODUCTION AND RELATION TO PRIOR WORK

We consider a class of nonstationary signals that are composed of a wide-sense stationary (WSS) random process $w(t)$ modulated by a deterministic, quasi-periodic modulator $m(t)$: $x(t) = m(t)w(t)$. Such signals abound in communications and are also useful for modeling real-world signals such as machinery noise, particularly propeller noise in passive sonar.

Using the assumptions that $w(t)$ is white and that $m(t) = \sum_{k=0}^{\infty} a_k e^{-j2\pi k f_0 t}$, Lourens and du Preez [1] demonstrated that the DEMON spectrum (termed "DEMON" in the sonar literature for "demodulated noise", but can also be used for general analysis of modulation frequency) is an approximate maximum-likelihood estimator (MLE) of the constant fundamental frequency of modulation, f_0. Clark et al. [2] demonstrated that for real-world sonar signals, $w(t)$ is often colored, and relaxed the assumption that $w(t)$ is white. They proposed a multiband version of the DEMON spectrum, which showed improved performance. Tao et al. [3] relaxed the assumption of constant fundamental frequency, assuming a linear chirp model for $m(t)$, and derived a MLE for propeller acceleration rate.

Similar approaches have been taken recently to modeling acoustic frequency variation, with most of the methods focused on speech processing. Omer and Torrésani [4] considered a model that consists of a frequency-modulated complex-valued WSS process and derived an approximate ML estimator using Gabor frames. Kaewtip et al. [5] used time-warping to achieve a similar effect and used the approach to improve automatic speech recognition. Kepési and Weruaga [6] used the fan-chirp transform to fit linear instantaneous frequency (IF) trajectories to short frames of voiced speech.

Our approach differs from these acoustic frequency methods in two ways: first, we are looking at modulation frequency instead of acoustic frequency, and second, our method allows for an arbitrary model of trajectories (i.e., not just constant or linear). Furthermore, our ultimate goal differs somewhat: rather than examine estimator performance, our goal is extension of the coherence time.

In this paper, we generalize the DEMON spectrum to allow for frequency variation in the modulator itself. We also show the relationship between the DEMON spectrum and the theory of spectral impropriety, namely that the DEMON spectrum is equivalent to a scaled spectral impropriety coefficient. Incorporating a model of the frequency variation into a GLRT detection statistic increases the coherence time of the estimator, allowing for the use of longer analysis windows and thus providing greater SNR.

2. BACKGROUND

This section covers the definition of the signal model we are interested in, the DEMON spectrum, its connection to spectral impropriety, and the generalized DEMON spectrum.

2.1. Signal model

In this paper, we will consider a signal of the form

$$x(t) = m(t)w(t), \tag{1}$$

where $m(t)$ is a deterministic, periodic modulator of the form

$$m(t) = \text{Re}\left\{ \sum_{k=1}^{K} a_k \exp\left(j2\pi k\phi_0(t)\right) \right\}, \tag{2}$$

with $a_k \in \mathbb{C}$ and $\phi_0(t) = \int_0^t f_0(w)dw$. $f_0(t)$ is the slowly time-varying fundamental instantaneous frequency (IF), and $w(t)$ is a wide-sense stationary (WSS) random process. The signal model in (1) will be referred to as *amplitude-modulated wide-sense stationary* (AM-WSS).

The signal in (1) is a nonstationary random process, because its time-varying autocovariance function is

$$r_{xx}(t, \tau) = E\left[x(t)x(t-\tau)\right] = m(t)m(t-\tau)r_{ww}(\tau). \tag{3}$$

This work is funded by ONR contract N00014-12-G-0078, delivery order 0013

The $m(t)m(t-\tau)$ factor causes the autocovariance function to vary over time. If $f_0(t)$ is constant over all time, (1) is a periodically-correlated [7], or cyclostationary [8] signal. If $f_0(t)$ is such that $m(t)$ is an almost-periodic function, (1) is an almost-cyclostationary signal [9].

2.2. DEMON spectrum and modulation frequency

The DEMON spectrum is a commonly-used tool in the sonar community to examine the frequency content of $m(t)$ (that is, the modulation frequency content of $x(t)$). The DEMON spectrum of a real-valued signal $x(t)$ is the Fourier transform of the squared signal:

$$D_x(f) = \int x^2(t) e^{-j2\pi f t} dt. \tag{4}$$

The squaring operation is a nonlinearity that performs a demodulation operation (essentially demodulating the signal with itself). If $m(t)$ has constant fundamental frequency f_0 and consists of harmonics, $|D_x(f)|$ exhibits spectral peaks at multiples of f_0.

If $m(t)$ has frequency content that varies over time, a time-variant version of the DEMON spectrum can be used for analysis. The time-varying DEMON spectrum is defined as follows:

$$D_x(d,f) = \int x_d^2(t) e^{-j2\pi f t} dt, \tag{5}$$

with $x_d(t) = g(t)x(t-dT)$ is the dth frame of $x(t)$, where $g(t)$ is a data taper of duration T. Just as a short-time Fourier transform (STFT) is used to examine time-varying acoustic frequency content, the time-varying DEMON spectrum is used to examine time-varying modulation frequency content.

2.3. Connection to spectral impropriety

Clark et al. examined the spectral impropriety of real-valued signals in [10]. Here we do the same, though with a different formulation.

The Cramér-Loève spectral representation of a random process $x(t)$ is [11]

$$x(t) = \int e^{j2\pi f t} d\xi(f), \tag{6}$$

where $d\xi(f)$ is a complex-valued spectral increment process. Since $d\xi(f)$ is complex valued, it requires two second-order statistics [12]. Thus, its increments have Hermitian and complementary correlations [12, p 199]:

$$E[d\xi(f_1)d\xi^*(f_2)] = S_{xx}(f_1,f_2)df_1 df_2 \tag{7}$$

$$E[d\xi(f_1)d\xi(f_2)] = \widetilde{S}_{xx}(f_1,f_2)df_1 df_2. \tag{8}$$

There is some ambiguity as to how the complementary spectral correlation can be defined. Schreier and Scharf [12] define $\widetilde{S}_{xx}(f_1,f_2)df_1 df_2 = E[d\xi(f_1)d\xi(-f_2)]$. Napolitano [9, (1.13)] defines $\widetilde{S}_{xx}(f_1,f_2)$ as we have done in (8), and we define it as such here because of the elegant expression it yields.

If $x(t)$ is a finite energy signal, we can say that $d\xi(f) = X(f)df$, where $X(f)$ is the Fourier transform of $x(t)$. Using this assumption, Clark observed [13, (3.12)] that if $x(t)$ is AM-WSS and $w(t)$ is white noise, the complementary spectral correlation $\widetilde{S}_{xx}(f_1,f_2) = E[X(f_1)X(f_2)]$ is

$$\widetilde{S}_{xx}(f_1,f_2) = \sigma_w^2 \int m^2(t) e^{-j2\pi(f_1+f_2)t}. \tag{9}$$

If $x(t)$ is real-valued and thus $X^*(f) = X(-f)$, the Hermitian spectral correlation $S_{xx}(f_1,f_2) = E[X(f_1)X^*(f_2)]$ is

$$S_{xx}(f_1,f_2) = \sigma_w^2 \int m^2(t) e^{-j2\pi(f_1-f_2)t}. \tag{10}$$

If we assume that the complex-valued random variables $X(f)$ are zero mean, then the spectral correlations $\widetilde{S}_{xx}(f_1,f_2)$ and $S_{xx}(f_1,f_2)$ are equal to the spectral covariances $\widetilde{R}_{xx}(f_1,f_2)$ and $R_{xx}(f_1,f_2)$. If $X(f)$ are distributed as complex-valued Gaussians, then the impropriety coefficient [12] of $X(f)$ is defined as

$$\rho_X(f) = \frac{\widetilde{R}_{xx}(f,f)}{R_{xx}(f,f)} = \frac{\sigma_w^2 \int m^2(t) e^{-j2\pi 2ft} dt}{\sigma_w^2 \int m^2(t) dt}. \tag{11}$$

Equation (11) is interesting, because it provides a connection between spectral impropriety and the DEMON spectrum. That is, when $x(t)$ is a real-valued, finite energy, AM-WSS signal and $w(t)$ is white noise, the expected value of the DEMON spectrum of $x(t)$ at double frequencies, $E[D_x(2f)]$, is essentially an unnormalized spectral impropriety coefficient scaled by $\|m(t)\|_2^2$.

2.4. Generalized DEMON spectrum

Most prior work has assumed that $f_0(t) = f_0$; that is, the fundamental modulation frequency is constant over the duration of the signal $x(t)$, and thus $m(t)$ is perfectly periodic. Real-world data, however, often has time-varying modulation frequency, and thus rarely has perfectly periodic modulation.

Time-varying frequency content in $m(t)$ smears energy in the DEMON spectrum. This effect is illustrated in figure 1, where the DEMON spectrum of two types of AM-WSS signals are shown. For both AM-WSS signals, $w(t)$ is unit-variance white noise. For the pure-tone case, $m(t)$ is a sinusoid with $f_0 = 10$ Hz. This results in a narrow peak in the DEMON spectrum at $2f_0 = 20$ Hz. In the second case, $m(t)$ is a frequency-modulated (FM) tone with starting frequency of 10 Hz and total frequency change of 5 Hz over the duration. Because of the FM in $m(t)$, the DEMON spectrum is smeared out, resulting in a reduced magnitude at the spectral location corresponding $2f_0 = 20$ Hz.

To correct the "smearing" effect caused by time-varying frequency content in $m(t)$, a generalized DEMON spectrum can be defined that is more coherent with the signal. If the DEMON spectrum is taken along an instantaneous frequency trajectory that is approximately equal to the true instantaneous frequency of $m(t)$, then the peaks at multiples of this trajectory in the generalized DEMON spectrum will be sharper and maximized. For example, using the true instantaneous frequency trajectory, the generalized DEMON spectrum of the

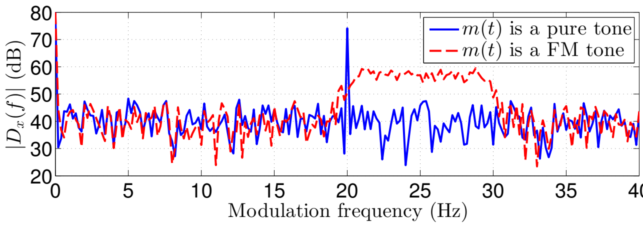

Fig. 1. Illustration of smearing of the DEMON spectrum when $m(t)$ is a frequency-modulated tone. In both cases $w(t)$ is white noise.

AM-WSS when $m(t)$ an FM tone would be equal to the blue line in figure 1, which would restore the concentrated peak at 20 Hz.

Using the connection to spectral impropriety from §2.3, the smearing in the DEMON spectrum also corresponds to smearing of the spectral impropriety. Thus, if a measurement of spectral impropriety does not account for time-varying modulation frequency, the estimate will be smeared out across adjacent frequencies and will be reduced. This indicates the importance of coherence when measuring spectral impropriety.

Formally, define a generalized DEMON spectrum along an IF trajectory $f(t)$, where $f(t)$ is a function over the duration T of $x(t)$. Let $\phi(t) = \int_0^t f(w)dw$ be the instantaneous phase corresponding to the trajectory $f(t)$. Then the generalized DEMON spectrum is defined to be

$$D_x^{(G)}(\phi) = \int x^2(t) e^{-j2\pi\phi(t)} dt. \tag{12}$$

This formulation can use an arbitrary model of trajectories. However, to simplify derivation and implementation, the remainder of this paper will use a linear model of IF trajectories, where $f(t) = f_c + \beta t$. In this model, f_c is a center frequency, and β is a chirp rate.

3. DETECTION OF AM-WSS SIGNALS

To illustrate the advantage of using the generalized DEMON spectrum, we consider the problem of detecting an AM-WSS signal $x(t)$ embedded in additive white Gaussian noise $v(t)$. The noisy measurements are

$$y(t) = x(t) + v(t) = m(t)w(t) + v(t). \tag{13}$$

Let the variance of $v(t)$ be σ_v^2 and let $w(t)$ be white noise with variance 1. Assume $m(t)$, $w(t)$, and $v(t)$ are zero-mean.

The following derivation represents discrete data sampled at f_s as N-length vectors, for example $\mathbf{y}$ where the nth element is $y[n]$ for $n \in [0, N-1]$. Here, we have the following two hypotheses:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y} = \mathbf{v} \\ \mathcal{H}_1 &: \mathbf{y} = \mathbf{x} + \mathbf{v} \end{aligned} \tag{14}$$

We will use a generalized likelihood ratio test (GLRT) [14] as a detection statistic, which is defined as

$$\mathcal{L}(\mathbf{y}) \triangleq \frac{\max_{\theta_1} p(\mathbf{y}; \boldsymbol{\theta_1}, \mathcal{H}_1)}{\max_{\theta_0} p(\mathbf{y}; \boldsymbol{\theta_0}, \mathcal{H}_0)} \tag{15}$$

where $\boldsymbol{\theta_i}$ are the unknown parameters under the hypothesis $\mathcal{H}_i$. Here, under $\mathcal{H}_1$, we take $\boldsymbol{\theta_1}$ to be the parameters that specify $f_0(t)$. Under $\mathcal{H}_0$, we have no unknown parameters.

If we examine the natural log of (15), we get

$$\ln \mathcal{L}(\mathbf{y}) = \ln \max_{\boldsymbol{\theta_1}} p(\mathbf{y}; \boldsymbol{\theta_1}, \mathcal{H}_1) - \ln p(\mathbf{y}; \mathcal{H}_0). \tag{16}$$

Since $\ln(\cdot)$ is a monotonically increasing function, we can push the max outside the log. Using the assumption $m^2[n] \ll \sigma_v^2$ for all n, and taking $p(\cdot)$ to be a Gaussian distribution, simplification of (16) yields the following:

$$\begin{aligned} \ln \mathcal{L}(\mathbf{y}) \approx \max_{\boldsymbol{\theta_1}} \Bigg[&-\frac{1}{2} \sum_{n=0}^{N-1} \ln \left(1 + \frac{m^2[n]}{\sigma_v^2}\right) \dots \\ &+ \frac{1}{2\left(\sigma_v^2\right)^2} \sum_{n=0}^{N-1} m^2[n] y^2[n] \Bigg]. \end{aligned} \tag{17}$$

The first term inside the max of (17) can be approximated by $-\frac{N}{2}$SNR using the assumption that $\frac{m^2[n]}{\sigma_v^2} \ll 1$ for all n and the Taylor series expansion for $\ln(1+x)$. Since $-\frac{N}{2}$SNR is a constant term, it can be pushed outside the max and absorbed into the likelihood threshold. The $\frac{1}{2(\sigma_v^2)^2}$ scaling factor on the second term in (17) is also data-independent and can be absorbed into the likelihood threshold. Thus, we can define the modified log-GLRT

$$L(\mathbf{y}) \triangleq \max_{\boldsymbol{\theta_1}} \sum_{n=0}^{N-1} m^2[n] y^2[n]. \tag{18}$$

Under the conventional assumption that $f_0(t) = f_0$, Lourens and du Preez derived an expression for (18) [1, (17)-(21)] in terms of the DEMON spectrum, which yields

$$L(\mathbf{y}) = \max_{f_0} \sum_{k=0}^{\infty} |D_{\mathbf{y}}(kf_0)|^2, \tag{19}$$

where $D_{\mathbf{y}}(f) = \sum_{n=0}^{N-1} y^2[n] e^{-j2\pi f \frac{n}{f_s}}$ is the discrete version of (4). Likewise, using a linear model $f_0(t) = f_c + \beta t$, Tao et al. derived a similar expression for (18) in [3], which we can write in terms of the discrete version of the generalized DEMON spectrum (12): $D_{\mathbf{y}}^{(G)}(\phi) = \sum_{n=0}^{N-1} y^2[n] e^{-j2\pi\phi(t)}$, where $\phi(t)$ is evaluated at $t = \frac{n}{f_s}$, yielding

$$L^{(G)}(\mathbf{y}) = \max_{f_c, \beta} \sum_{k=0}^{\infty} \left| D_{\mathbf{y}}^{(G)}\left(k \left[f_c + \frac{1}{2}\beta t\right] t\right) \right|^2. \tag{20}$$

Note that the k=0 terms in (19) and (20) are simply $\|\mathbf{y}\|_2^2$, so they are constant terms that are not affected by $\boldsymbol{\theta_1}$ and can thus be moved outside the maximization.

4. EXAMPLES

We now consider the effect of the length of the analysis window on detection performance and demonstrate that when fundamental modulation frequency varies within a window, $L^{(G)}(\mathbf{y}) \geq L(\mathbf{y})$ for both synthetic and real data. This result shows that our proposed method increases the coherence time and thus allows longer analysis windows.

4.1. Synthetic data

A Monte Carlo analysis using synthetic data is performed. Realistic parameters that correspond to real-world passive sonar signals are used. The WSS noise carrier $w(t)$ is unit-variance white noise. The modulator $m(t)$ has 4 harmonics, center frequency $f_c = 2$ Hz, and we consider four possible chirp rates spaced equally from 0.05 to 0.2 Hz/s. For each chirp rate, we will consider a range of analysis window durations $T \in [3, 8]$ seconds. The sampling rate is $f_s = 4$ kHz.

For each T, $L(\mathbf{y})$ and $L^{(G)}(\mathbf{y})$ are computed and averaged over 400 trials, 100 trials for each β. Figure 2 shows the results. Notice that $L^{(G)}(\mathbf{y})$ remains approximately the same for all β for all analysis durations T, while $L(\mathbf{y})$ decreases as T increases. The difference in performance between $L(\mathbf{y})$ and $L^{(G)}(\mathbf{y})$ happens because $L^{(G)}(\mathbf{y})$ accounts for linear frequency variations of $m(t)$, which means $L^{(G)}(\mathbf{y})$ is coherent with the signal over a longer duration. Said another way, $L^{(G)}(\mathbf{y})$ adds up the concentrated, higher-amplitude peaks produced by the generalized DEMON spectrum, while $L(\mathbf{y})$ adds up the smeared, lower-amplitude peaks of the conventional DEMON spectrum.

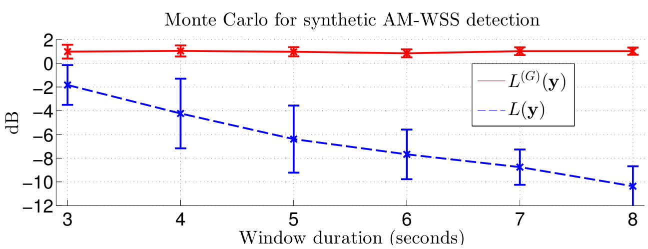

Fig. 2. Monte Carlo experiment with synthetic data demonstrating increased coherence time using $L^{(G)}(\mathbf{y})$, the log-GLRT for our proposed model, versus $L(\mathbf{y})$, the log-GLRT for the conventional model.

4.2. Real-world sonar data

For a real-world example, we consider propeller noise that exhibits time-varying frequency content in its modulator. The example is a 12 second recording $\mathbf{z}$ of a Zodiac boat starting up its engine and moving away[1]. The time-varying DEMON spectrum given by (5) of this signal is shown in figure 3. The first 6 seconds are labeled as "startup", during which the propeller rate is increasing as the boat starts its engine and moves away. During the last 6 seconds, labeled "steady", the Zodiac boat is underway and the propeller rate is relatively constant. We choose this example because the startup portion illustrates a real-world case when $f_0(t)$ is varying over the analysis window, while the steady portion illustrates the conventional assumption that $f_0(t)$ is constant over the analysis window.

Figure 4 shows a comparison between $L(\mathbf{z}_d)$ and $L^{(G)}(\mathbf{z}_d)$ for varying analysis window durations T. The expression $\mathbf{z}_d$ is the dth frame of duration T of $\mathbf{z}$. The analysis windows are not overlapping. Notice that for all T and d, and for

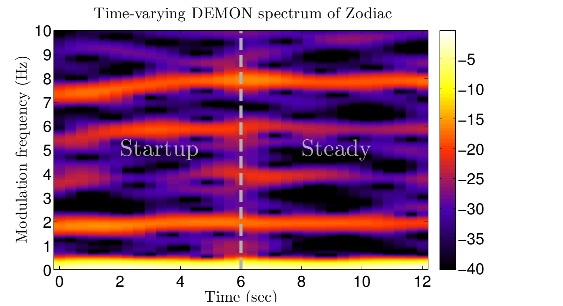

Fig. 3. Time-varying DEMON spectrum of Zodiac boat data. The startup and steady portions are labeled. Notice the increasing modulation frequency in the startup portion.

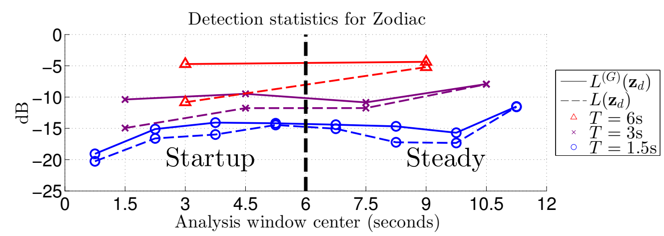

Fig. 4. Detection statistics for real-world Zodiac boat data using different analysis window lengths. Analysis windows are non-overlapping. Notice that the statistic using the proposed generalized DEMON spectrum, $L^{(G)}(\mathbf{z}_d)$, is consistently higher than the conventional statistic $L(\mathbf{z}_d)$, especially as the analysis window duration T increases.

both startup and steady, $L^{(G)}(\mathbf{z}_d)$ consistently achieves a higher value than $L(\mathbf{z}_d)$. During startup, $L^{(G)}(\mathbf{z}_d)$ is especially greater (by over 6dB) using a long analysis window ($T = 6$s), because the propeller rate is far from constant.

5. DISCUSSION AND CONCLUSION

In this paper, we have examined a particular class of non-stationary signals, AM-WSS, and how its spectral statistics relate to the DEMON spectrum. Conventional approaches assume that the modulation frequency content is stationary within the window. However, this assumption often does not hold. By accounting for the variation, longer analysis windows can be used, providing higher SNR estimators.

We presented a method to extend the coherence time of such signals using the generalized DEMON spectrum. We validated this method using a linear model for modulation frequency trajectories to perform detection on synthetic data in a Monte Carlo experiment as well as on real-world passive sonar data. On real-world passive sonar data, our proposed method achieved up to a 6dB gain in the detection statistic over the conventional method.

Future work will investigate other models for time-varying modulation frequency (such as quadratic and higher-order polynomials), applying the method to speech signals to extend analysis windows, and incorporating our method into other statistical estimators used in beamforming, detection, and classification.

[1]Thanks to Brad Hanson and the NOAA Northwest Fisheries Science Center Marine Mammal Program for providing this data.

6. REFERENCES

[1] J. Lourens and J. du Preez, "Passive sonar ML estimator for ship propeller speed," *IEEE Journal of Oceanic Engineering*, vol. 23, no. 4, pp. 448–453, 1998.

[2] P. Clark, I. Kirsteins, and L. Atlas, "Multiband analysis for colored amplitude-modulated ship noise," in *2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2010, pp. 3970–3973.

[3] R. Tao, Y. Feng, and Y. Wang, "Theoretical and experimental study of a signal feature extraction algorithm for measuring propeller acceleration in a port surveillance system," *IET Radar, Sonar Navigation*, vol. 5, no. 2, pp. 172–181, 2011.

[4] H. Omer and B. Torrésani, "Estimation of frequency modulations on wideband signals; applications to audio signal analysis," arXiv e-print 1305.3095, May 2013.

[5] K. Kaewtip, L. N. Tan, and A. Alwan, "A pitch-based spectral enhancement technique for robust speech processing," *Interspeech*, 2013.

[6] M. Kepési and L. Weruaga, "Adaptive chirp-based time-frequency analysis of speech signals," *Speech Communication*, vol. 48, no. 5, pp. 474–492, May 2006.

[7] E. G. Gladyshev, "Periodically and almost-periodically correlated random processes with a continuous time parameter," *Theory of Probability & Its Applications*, vol. 8, no. 2, pp. 173–177, Jan. 1963.

[8] W. A. Gardner, *Statistical spectral analysis: a nonprobabilistic theory*, Prentice Hall, Englewood Cliffs, N.J., 1988.

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

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

[11] M. Loève, *Probability theory*, Springer-Verlag, New York, 1977.

[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] P. Clark, *Coherent Demodulation of Nonstationary Random Processes*, Ph.D. thesis, University of Washington, 2012.

[14] S. M. Kay, *Detection theory*, Prentice Hall PTR, Upper Saddle River, NJ, 1998.