1999-05-01

The probability density of spectral estimates based on modified periodogram averages

David G. Long
david_long@byu.edu

Paul E. Johnson

The Probability Density of Spectral Estimates Based on Modified Periodogram Averages

Paul E. Johnson, *Student Member, IEEE*, and David G. Long, *Senior Member, IEEE*

Abstract—**Welch's method for spectral estimation of averaging modified periodograms has been widely used for decades. Because such an estimate relies on random data, the estimate is also a random variable with some probability density function. Here, the pdf of a power estimate is derived for an estimate based on an arbitrary number of frequency bins, overlapping data segments, amount of overlap, and type of data window, given a correlated Gaussian input sequence. The pdf's of several cases are plotted and found to be distinctly non-Gaussian (the asymptotic result of averaging frequency bins and/or data segments), using the Kullback–Leibler distance as a measure. For limited numbers of frequency bins or data segments, the precise pdf is considerably skewed and will be important in applications such as maximum likelihood tests.**

Index Terms—**Power spectrum, probability density functions, spectral estimation, Welch's modified periodogram estimates.**

I. INTRODUCTION

THIS paper analyzes the statistics of spectral estimates based on averaging modified periodograms, which are sometimes referred to as Welch's technique for spectrum estimation [1]. For almost 30 years, this technique has been widely applied, but there has been no thorough statistical analysis of the technique. Indeed, only two papers extend beyond the mean and variance of the spectral estimates. The first [2] describes the probability distributions for discrete Fourier spectra based on a single periodogram for data both smoothed and unsmoothed, but does not describe the correlation between frequency bins. The second [3] derives the joint density functions for two frequency bins, including windowed data and averaging over nonoverlapping data periodogram estimates. Here, the pdf of the sum of frequency bins from an average of modified periodogram bins, with nonoverlapping and overlapping data, is derived.

For a data stream $x[j]$, Welch's modified periodogram averaging provides a technique for estimating the spectrum [1]. The sequence is segmented into K overlapping sequences, each of length L, such that the ith sequence is defined by $x[i, n] = x[n + i(1 - r)L]$, where the indices are $i = 0, \ldots, K - 1$ and $n = 0, \ldots, L - 1$. r is the percentage of overlap. The data is windowed to minimize spectral leakage.

The window can be applied through multiplication in the time domain $(w[n])$ or through circular convolution in the frequency domain $(W[k])$. The estimate of the power spectral density of each segment $P[i, k]$ is

$$P[i, k] = \frac{1}{LU} |X[i, k] * W[k]|^2 \tag{1}$$

where U is a scale factor dependent on the window [1]. To obtain the power in a particular frequency band, several bins of the modified periodogram may be summed. Further, the variance of the estimate is reduced by averaging, over i, all of the data segments.

In this paper, the statistical properties of this method of spectral estimation are explored for the case where $x[j]$ is a wide-sense stationary, Gaussian sequence, with distribution $N(0, \mathbf{R})$. The process is formulated in vector space notation from which the resulting power estimate is shown to be a quadratic form in the data vector $\mathbf{x}$. The pdf corresponding to this quadratic form is given in generality, and several representative examples are plotted. Simulations are reported in which sample pdf's of the power estimates based on averaging-modified periodograms match the theoretical functions derived here. Finally, we compute the Kullback–Leibler distance of this density function from a Gaussian.

II. VECTOR SPACE ANALYSIS OF WELCH'S SPECTRUM ESTIMATION

To derive the pdf of a spectral estimate of the power in a frequency band of a data stream, it is useful to consider the vector space formulation of the estimate. For the real, length N, data vector $\mathbf{x} = [x_0, x_1, \ldots, x_{N-1}]^T$, distributed Gaussian with zero-mean and covariance matrix $\mathbf{R}$, we include a data window matrix $\Omega = \text{diag}(w[n])$. Welch's method segments $\mathbf{x}$ into data segments, possibly overlapping, of length L. For an arbitrary data segment i, the periodogram estimate of the power in frequency bin k can be written as

$$P[i, k] = \mathbf{x}^T \Upsilon[i, k] \mathbf{x} \tag{2}$$

$$= \mathbf{x}^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{LU} \Omega(\mathbf{C}_k + \mathbf{S}_k)\Omega & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{x} \tag{3}$$

where the sizes of the $\mathbf{0}$ matrices are adjusted to pick which data segment is being used and the matrices $\mathbf{C}_k$ and $\mathbf{S}_k$ are from the cosine and sine transform kernels. The (m, n)th element of $\mathbf{C}_k + \mathbf{S}_k$ is

$$(\mathbf{C}_k + \mathbf{S}_k)_{m,n} = \cos\left(2\pi \frac{k}{Z}(n - m)\right) \tag{4}$$

where Z is the length of the zero-padded data vector (i.e., if $Z = L$, each data segment has not been padded).

Averaging the power from the data segments and summing over several frequency bins yields a quadratic form in the normal random vector $\mathbf{x}$

$$P = \frac{1}{K}\sum_{i=0}^{K-1}\sum_{k=k_1}^{k_2} P[i,k] \qquad (5)$$

$$= \mathbf{x}^T \mathbf{\Upsilon} \mathbf{x} \qquad (6)$$

where the matrix $\mathbf{\Upsilon}$ depends on the length of the data segment, the amount of zero-padding, the data window, the frequency bins of interest, the number of periodograms to average, and the amount of overlap of the data segments

$$\mathbf{\Upsilon} = \frac{1}{K}\sum_{i=0}^{K-1}\sum_{k=k_1}^{k_2} \mathbf{\Upsilon}[i,k]. \qquad (7)$$

Formation of the matrix $\mathbf{\Upsilon}$ can be facilitated for arbitrary overlap by partitioning the fundamental matrix of (3) $\frac{1}{LU}\Omega(\mathbf{C}_k + \mathbf{S}_k)\Omega$ such that overlapping partitions from different data segments can be added. For example, with nonoverlapping data ($r = 0$), partitioning is unnecessary, and $\mathbf{\Upsilon}$ is a block diagonal matrix that is formed as the Kronecker product

$$\mathbf{\Upsilon} = \mathbf{I}_K \otimes \left[\Omega\left(\frac{1}{KLU}\sum_{k=k_1}^{k_2}\mathbf{C}_k + \mathbf{S}_k\right)\Omega\right]. \qquad (8)$$

For 50% overlap ($r = 0.5$), the fundamental matrix can be partitioned into four submatrices (each of $\frac{L}{2}$ elements square).

The moment-generating function of a quadratic form in a zero-mean Gaussian vector, such as (6), can be expressed as [4, p. 65]

$$M(t) = \prod_{i=1}^{D}(1 - t2\eta_i)^{-\frac{\nu_i}{2}} \qquad (9)$$

where the η_i are the D distinct, nonzero eigenvalues of $\mathbf{R}\mathbf{\Upsilon}$. Each eigenvalue has multiplicity ν_i.

The mean and variance of the power estimate P are, respectively

$$\mu_P = \sum_{i=1}^{D}\eta_i\nu_i \qquad (10)$$

$$\sigma_P^2 = \sum_{i=1}^{D}2\eta_i^2\nu_i. \qquad (11)$$

Defining some convenient variables

$$d_i = -\frac{1}{2\eta_i} \qquad (12)$$

$$g = \prod_{i=1}^{D}(2\eta_i)^{-\frac{\nu_i}{2}} \qquad (13)$$

the pdf corresponding to (9) is (see the Appendix)

$$f(y) = \frac{g}{\Gamma(\frac{J}{2})}y^{\frac{J}{2}-1}\Phi_2\left(\frac{\nu_1}{2},\ldots,\frac{\nu_D}{2};\frac{J}{2};d_1y, d_2y,\ldots,d_Dy\right) \qquad (14)$$

where Φ_2 is the generalized hypergeometric function, and J is the total number of nonzero eigenvalues.

While (14) provides a completely general solution for the pdf, calculation of the hypergeometric function is computationally intensive. However, if all of the nonzero eigenvalues have even multiplicities, $M(s)$ can be expanded with a partial fraction expansion. This yields a simple and practical density function (see Appendix)

$$f_P(p) = g\sum_{i=1}^{D}\sum_{j=1}^{h_i} A_{ij}\frac{1}{(j-1)!}p^{j-1}e^{d_iP}U(p). \qquad (15)$$

Even when these eigenvalues are distinct, they are typically in near pairs for commonly used windows. Clustering the nonzero eigenvalues into groups with even multiplicities to use with (15) provides an accurate numerical approximation for practical computation of the density function.

We note in passing that for complex data, $\mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I$, with the assumption that the vectors $\mathbf{x}_R$ and $\mathbf{x}_I$ are independent and distributed $N(0, \mathbf{R})$ so that $\mathbf{x}$ is complex circular $N(0, 2\mathbf{R})$, the moment-generating function for the power estimate is the square of (9)

$$M(t) = \prod_{i=1}^{D}(1 - t2\eta_i)^{-\nu_i}. \qquad (16)$$

This ensures even multiplicities of the eigenvalues; therefore, the partial fraction expansion always yields the exact pdf without grouping eigenvalues.

In practice, the cumulative distribution function is frequently desired; this is easily found by integration of (15), resulting in a sum of incomplete Gamma functions

$$F_P(p) = g\sum_{i=1}^{D}\sum_{j=1}^{h_i} A_{ij}\frac{(-d_i)^{-j}}{(j-1)!}\gamma(j, -d_i p)U(p) \qquad (17)$$

where the incomplete Gamma function is defined as [5]

$$\gamma(a, x) = \int_0^x e^{-t}t^{a-1}\,dt. \qquad (18)$$

III. THE PDF'S FOR SPECIAL CASES

In this section, (15) is developed for several special cases. In particular, note that the pdf of the power in a frequency band based on averaging modified periodograms requires computation of the eigenvalues of $\mathbf{R}\mathbf{\Upsilon}$, where $\mathbf{R}$ is the covariance matrix of the Gaussian sequence $\mathbf{x}$. For simplicity, in the examples presented throughout most of this section, the signal $\mathbf{x}$ will be assumed to be a white $\mathbf{R} = \sigma^2\mathbf{I}$; we present one example of a colored signal. For white noise, the eigenvalues of $\mathbf{R}\mathbf{\Upsilon}$, labeled η, are simply σ^2 times the eigenvalues of $\mathbf{\Upsilon}$, labeled λ; that is, $\eta_i = \sigma^2\lambda_i$. We also discuss the distance between a Gaussian density and the theoretical density we have derived.

We will consider, as examples, two useful and interesting examples for $w[n]$: the rectangular window and the Hann window. For a rectangular data window, $w[n] = 1$, so that the window matrix is the identity $\Omega = \mathbf{I}$ and $U = 1$. The Hann window is defined as $w[n] = \sin^2(\pi\frac{n}{L})$ and $U = \frac{3}{8}$.
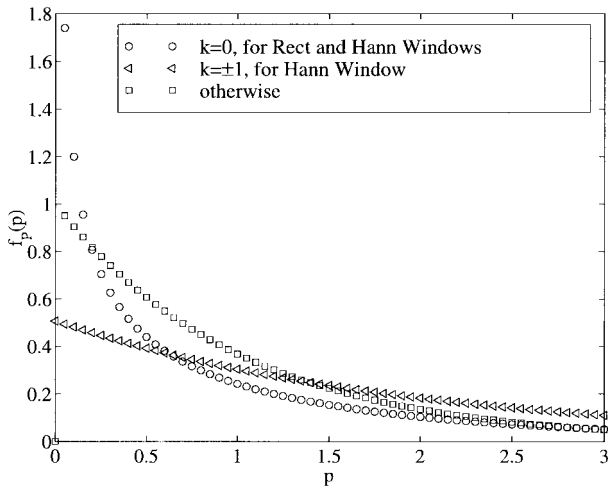
Fig. 1.   Probability density function of the power in a frequency bin estimated through a single periodogram estimate.

First, the well-known result for a single frequency bin based on a single data segment is developed for rectangular and Hann data windows. The more complicated case of averaging nonoverlapping data segments with multiple frequency bins is then found. The case of 50% overlapping data segments and multiple bins is considered. We also include an example of spectral estimation for a colored sequence. The section concludes with an examination of the Kullback–Leibler distance as a measure of how different the actual densities are from a Gaussian with the same first and second moments.

A. One Frequency Bin and a Single Data Segment

As a simple example, consider the case of a single frequency bin $(k_1 = k_2 = k)$ and a single data segment $(K = 1)$. The examples we present use even data segments $L = 2q$ with no zero-padding of the segments, although the theory of the previous section includes these possibilities.

For the rectangular window, the nonzero eigenvalues of $\Upsilon = \frac{1}{LU}\Omega(\mathbf{C}_k + \mathbf{S}_k)\Omega$ are easily found to be

$$\lambda(\Upsilon) = \begin{cases} 1, & \text{for } k = 0, L/2 \\ \frac{1}{2}, \frac{1}{2}, & \text{otherwise.} \end{cases} \tag{19}$$

The resulting density function is Gamma

$$f_P(p) = \begin{cases} \dfrac{1}{(2\sigma^2)^{\frac{1}{2}}\Gamma(\frac{1}{2})} p^{-\frac{1}{2}} \exp\left(-\dfrac{p}{2\sigma^2}\right) U(p), & \text{for } k = 0, \frac{L}{2} \\ \dfrac{1}{\sigma^2} \exp\left(-\dfrac{p}{\sigma^2}\right) U(p), & \text{otherwise.} \end{cases} \tag{20}$$

Note that for $k = 0, \frac{L}{2}$, with $\sigma^2 = 1$, the distribution is a Chi-square distribution with one degree of freedom; for other frequency bins, the distribution is exponential, which is a well-known result.

Similarly, for the Hann window, the nonzero eigenvalues of Υ are

$$\lambda(\Upsilon) = \begin{cases} 1, & k = 0, L/2 \\ \dfrac{5}{12}, \dfrac{7}{12}, & k = \pm 1, L/2 \pm 1 \\ \dfrac{1}{2}, \dfrac{1}{2}, & \text{otherwise.} \end{cases} \tag{21}$$

The pdf is the same as for the rectangular data window except at $k = \pm 1, \frac{L}{2} \pm 1$, where the solution is in terms of the modified Bessel function

$$f_P(p) = \begin{cases} \dfrac{1}{(2\sigma^2)^{\frac{1}{2}}\Gamma(\frac{1}{2})} p^{-\frac{1}{2}} \exp\left(-\dfrac{p}{2\sigma^2}\right) U(p), & \text{for } k = 0, \frac{L}{2} \\ \dfrac{1}{2\sigma^2}\sqrt{\dfrac{36}{35}} e^{-\frac{18}{35\sigma^2}p} I_0\left(\dfrac{3}{35\sigma^2}p\right) U(p), & \text{for } k = \pm 1, \frac{L}{2} \pm 1 \\ \dfrac{1}{\sigma^2} \exp\left(-\dfrac{p}{\sigma^2}\right) U(p), & \text{for otherwise.} \end{cases} \tag{22}$$

Fig. 1 displays the pdf's of the power in a single frequency bin, based on a single periodogram using a rectangular data window and a Hann window. The two data windows yield identical pdf's, except at $k = \pm 1, \frac{L}{2} \pm 1$, where the Hann window uses a Bessel function. For the plot, $\sigma^2 = 1$ is used. For all of the plots, the mean of P is 1, i.e., the estimate is an unbiased estimate of the signal variance σ^2.

B. Averaging Nonoverlapping Data Segments

As described in Section II, the case of nonoverlapping data segments Υ is formed as a Kronecker product. The eigenvalues of a Kronecker product are all the products of the eigenvalues of the two matrices. Applied to the problem at hand, averaging K segments increases the multiplicity of each eigenvalue for a single data segment by a factor of K.

For the rectangular data window with no overlap, each combination of b bins and K data segments can be characterized as having $D = 1$ distinct nonzero eigenvalues with multiplicity $\nu = 2Kb$ (as long as the frequency bins do not span across $k = 0$ or $\frac{L}{2}$, where the value of the eigenvalue is different). Summing several frequency bins yields the single nonzero eigenvalue $D = 1$ with multiplicity $\nu = 2b$; averaging over independent data segments increases the multiplicity of this eigenvalue to $\nu = 2Kb$. Further, the value of this eigenvalue is $\lambda = \frac{b}{\nu} = \frac{1}{2K}$. The pdf of a power estimate based on b frequency bins, averaging K nonoverlapping periodograms, with a rectangular data window is a Gamma distribution with $\alpha = Kb$ and $\beta = \frac{\sigma^2}{K}$:

$$f_P(p) = \left(\dfrac{K}{\sigma^2}\right)^{Kb} \dfrac{1}{\Gamma(Kb)} p^{Kb-1} e^{-p\frac{K}{\sigma^2}} U(p). \tag{23}$$

The Hann window with no overlap introduces some distinctions. Specifically, while the Kronecker product increases all the multiplicities by a factor of K, the window introduces correlation between the frequency bins such that summing multiple frequency bins yields more distinct nonzero eigenvalues. Again, avoiding frequency bins $k = 0, \pm 1, \frac{L}{2}$ and $\frac{L}{2} \pm 1$, the number of distinct nonzero eigenvalues is $D = b$, each with multiplicity $\nu = 2K$, that is, h_i of (15) is K. Note that for a single bin, this is identical to the case of the rectangular data window.

Some examples of the density function for rectangular and Hann windows are displayed in Fig. 2 for nonoverlapping data
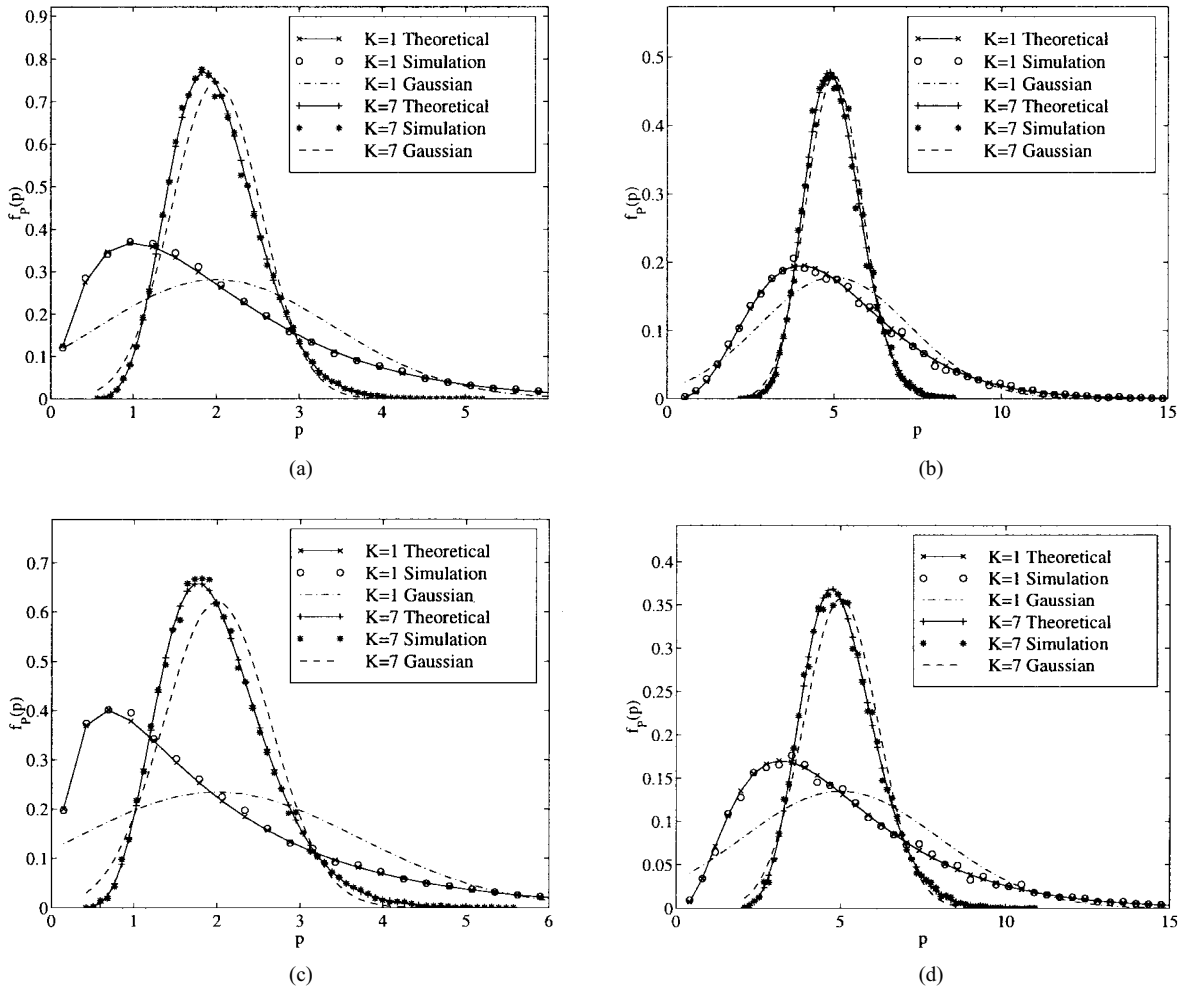
Fig. 2. Probability density functions for nonoverlapping data segments based on the sum of two and five frequency bins (adjacent frequency bins were chosen well away from $k = 0, \frac{L}{2}$). The solid lines plot the theoretical pdf's based on one and seven data segments, the dashed lines indicate the corresponding Gaussian densities with identical first and second moments, and the circles and asterisks are the results of Monte Carlo simulations.

segments. Here, we have assumed that the data stream **x** is white Gaussian noise, with variance $\sigma^2 = 1$ and summed two and five frequency bins. The solid lines display the theoretical density, whereas the broken lines are the corresponding Gaussian densities with the same mean and variance, and the circles and asterisks display the results of a Monte Carlo simulation in which Welch's method was applied to over 13 000 random data segments of length $L = 2^{17}$ to estimate the pdf of the power spectrum for each case. Although this is not an exhaustive set, these plots demonstrate the general behavior of the pdf of the power. In every case, the Monte Carlo simulation corresponds well to the theoretical distribution. Further, the mode of the distribution occurs at a value of p considerably less than the mean—the tail on the left side of the distribution is much lower than that on the right side. As the number of frequency bins added together increases, and/or as the number of data segments increases, the density becomes less skewed and closer to a Gaussian.

C. PDF for 50% Overlapping Data Segments

Having considered the cases of a single data segment and K nonoverlapping segments, we now consider the case of over-

lapping segments. In Section II, we described how overlapping data segments can be readily analyzed by partitioning the matrix for each periodogram and summing the corresponding submatrices. With this method, it is straightforward to construct such a matrix for any overlap and numerically evaluate the eigenvalues.

Examples of the densities based on 50% overlapping data segments, summing two and five frequency bins, are displayed in Fig. 3. As before, the data stream **x** is white Gaussian noise, with variance $\sigma^2 = 1$. The dashed lines, indicating the corresponding Gaussian density, are quite different from the theoretical pdf. The general form of the pdf of the power is shown to include considerable skew, with the right tail being much higher. The plot also shows the agreement of Monte Carlo simulation results with the theoretical density function as circles and asterisks.

D. A Correlated Data Segment

While the previous examples described the pdf of power estimates from Gaussian, white data sequences, the derivation allowed a more general, correlated sequence with arbitrary
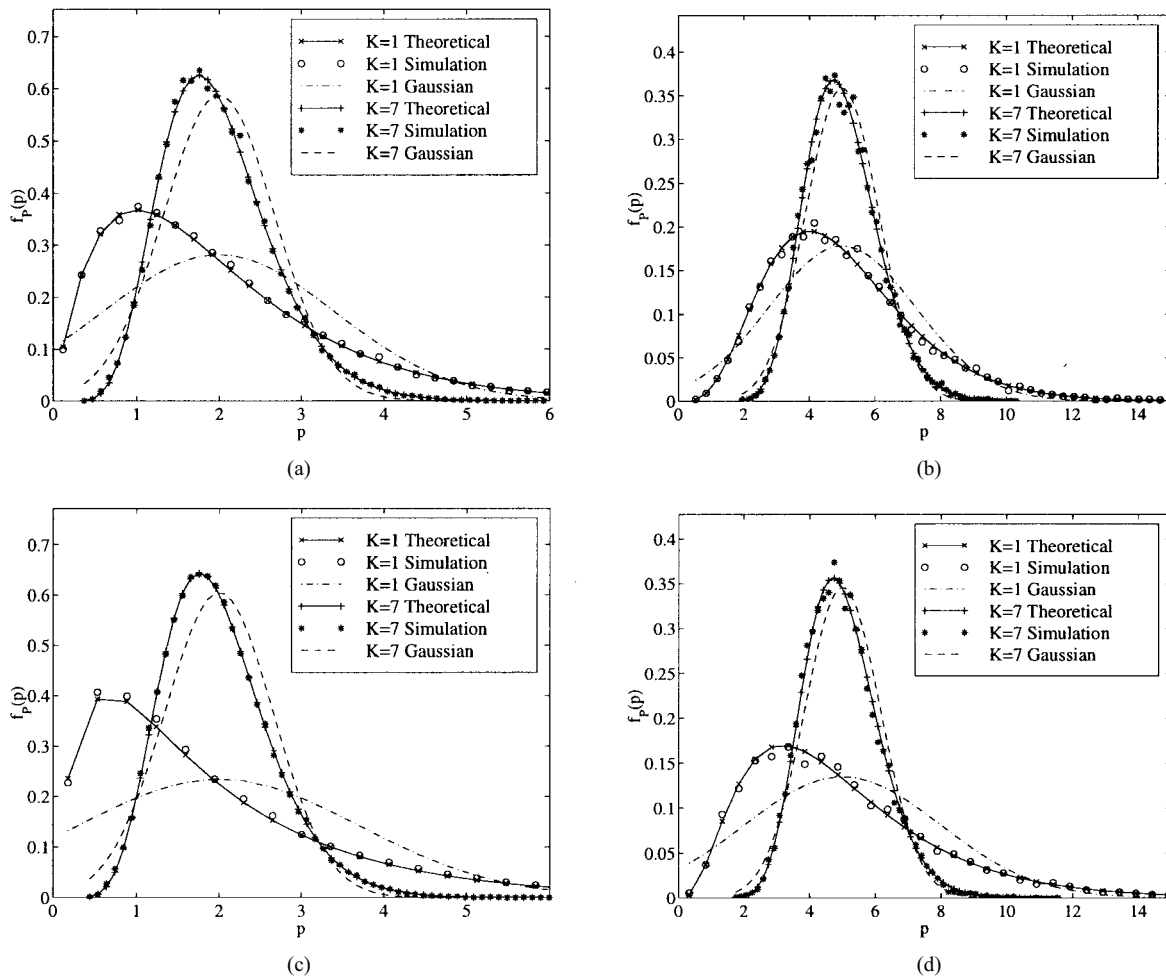
Fig. 3. Probability density functions for 50% overlapping data segments based on the sum of two and five frequency bins (adjacent frequency bins were chosen well away from $k = 0, \frac{L}{2}$). The solid lines indicate the theoretical pdf's, the dashed lines indicate the corresponding Gaussian densities, and the circles and asterisks are the results of Monte Carlo simulations.

autocorrelation matrix $\mathbf{R}$. Here, we consider an example of this more general case.

One constraint on the autocorrelation matrix, due to the fact that the data sequence is assumed stationary, is that $\mathbf{R}$ is Toeplitz; further, a common type of sequence results from a first-order Markov process with covariance function [6]

$$r(n) = \rho^{|n|}, \quad |\rho| < 1, \quad \forall n. \tag{24}$$

For an L element data segment $\mathbf{x}$, this results in a symmetric autocorrelation matrix with a value for the qth subdiagonal as ρ^q.

Equation (15) provides the pdf of the power in a frequency bin based on the eigenvalues of $\mathbf{R}\Upsilon$. The basic form of the pdf will be the same for correlated sequences as it was for white sequences; the only difference will be that sequence correlation may modify the eigenvalues, changing the details of the pdf.

In Fig. 4, the pdf of the estimated power spectrum is plotted. The x axis is the normalized frequency, plotted from 0 to π (the spectrum is symmetric). The pdf is displayed at four frequencies, with the mean of each estimate (indicated by small circles) corresponding well with the assumed spectrum of a first-order Markov process with correlation coefficient, $\rho = 0.3$. For this plot, each pdf was based on a single
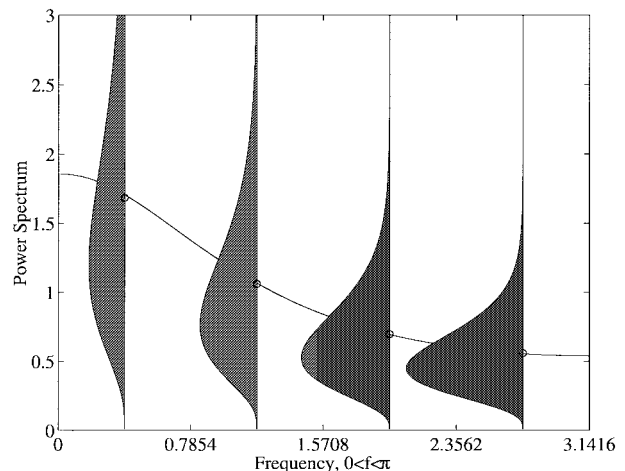

Fig. 4. For a correlated data sequence, the power spectrum (the Fourier transform of the autocorrelation function) varies with frequency. The solid line represents the spectrum of a first-order Markov process with correlation coefficient $\rho = 0.3$. The filled figures describe the pdf of the estimated spectrum at the four selected frequencies. Note that each pdf is skewed toward low power, whereas the mean of the pdf (indicated by the small circles) corresponds to the spectrum established by the Markov process.

frequency bin estimated with $K = 5$ data segments, each of length $L = 64$, with 50% overlap (the data window choice makes no effect for a single frequency bin).

E. Relative Entropy

The plots of probability density functions presented to this point suggest a wide range of behavior, depending on the particular parameters of how many frequency bins, data segments, amount of overlap, and type of data window used. The pdf asymptotically approaches a Gaussian in the limit as more frequency bins and/or data segments are used.

The "distance" between two probability density functions f and g can be quantified with the concept of relative entropy or Kullback–Leibler distance [7]

$$D(f \| g) = \int f(x) \log \frac{f(x)}{g(x)} \, dx. \tag{25}$$

Because the asymptotic behavior of the pdf is Gaussian and because it is not uncommon to compute the first two moments of the density and assume a Gaussian, we consider the distance between the theoretical pdf and a Gaussian with the same mean and variance.

For the special case where $\mathbf{R}\Upsilon$ has only a single nonzero eigenvalue η with multiplicity ν, the resulting pdf is Gamma with parameters $\alpha = \frac{\nu}{2}$ and $\beta = \eta$. This case corresponds to an uncorrelated data stream for which the power is estimated using a rectangular data window and nonoverlapping data; this is a useful case because the relative entropy between Gamma and Gaussian densities with mean $\alpha\beta$ and variance $\alpha\beta^2$ can be written in closed form as

$$D(f \| g) = \log\left(\frac{\sqrt{2\pi\alpha}}{\Gamma(\alpha)}\right) - \alpha + \frac{1}{2} - (\alpha - 1)\gamma$$
$$+ (\alpha - 1)\sum_{j=1}^{\alpha-1} \frac{1}{j}. \tag{26}$$

In more general cases, the relative entropy can be computed numerically. For several situations, $D(f \| g)$ is plotted in Fig. 5. Data points are calculated for $K = 1, 4, 7$, and 10 data segments (with linear interpolation between the data points). In every case, there is a nearly exponential decay as the pdf's asymptotically approach Gaussian as the number of data segments and/or frequency bins increases.

IV. CONCLUSION

Welch's modified periodogram averaging has served as a simple, common technique for spectral estimation for three decades. However, the statistical structure of the estimate has never been fully reported; in this paper, we present the pdf of a spectral estimate of a frequency band based on Welch's method with Gaussian input. The pdf of an estimate will be useful in many situations. For example, maximum-likelihood estimations approximating the distribution as Gaussian overestimate the mode because of the skewed structure of the correct pdf.

The critical component of the analysis requires the eigenvalues of a matrix, $\mathbf{R}\Upsilon$, which can be computed numerically. Several examples of the pdf have been plotted for illustration purposes. Simulations were presented in which the sample pdf's of the spectral estimates were compared with the theoretical functions derived here. Finally, the Kullback–Leibler
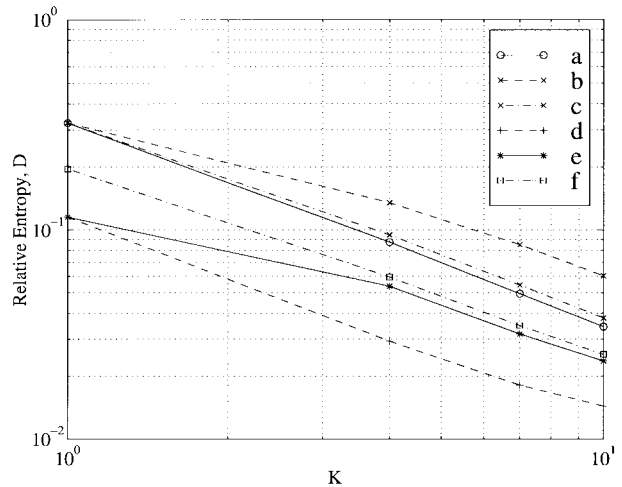

Fig. 5. Relative entropy, or Kullback–Leibler distance, between a Gamma distribution and a Gaussian distribution with the same mean and variance as a function of K, which is the number of data segments used in the periodogram estimate. (a) One bin, nonoverlapping segments, rect window. (b) One bin, overlapping segments, rect window. (c) One bin, overlapping segments, Hann window. (d) Three bins, nonoverlapping segments, rect window. (e) Three bins, overlapping segments, rect window. (f) Three bins, nonoverlapping segments, Hann window.

distance between the correct pdf and the corresponding Gaussian density was plotted for several representative cases, displaying the monotonically converging behavior of the pdf to a Gaussian as the frequency bins and/or the number of data segments is increased.

APPENDIX A
THE PDF FROM THE MOMENT-GENERATING FUNCTION

The pdf is the inverse Laplace transform of the moment-generating function with the negative of the argument, that is, the kernel of the moment-generating function is e^{tx}, whereas that of the Laplace transform is e^{-sx} [8]. In this section, the pdf is found from the moment-generating function in generality and in the simpler case when all eigenvalue multiplicities are even.

Defining $s = -t$, the Laplace transform of the pdf of the random variable $P = \mathbf{x}^T \Upsilon \mathbf{x}$, which is a central quadratic form in $\mathbf{x}$, where $\mathbf{x}$ is distributed $N(0, \mathbf{R})$, can be written as (with g and d_i defined in the text)

$$M(s) = \frac{g}{s^{\frac{J}{2}}} \prod_{i=1}^{D} \left(1 - \frac{d_i}{s}\right)^{-\frac{\nu_i}{2}}. \tag{27}$$

The inverse Laplace transform of (27) is [9]

$$f(y) = \frac{g}{\Gamma\left(\frac{J}{2}\right)} y^{\frac{J}{2}-1} \Phi_2\left(\frac{\nu_1}{2}, \ldots, \frac{\nu_D}{2}; \frac{J}{2}; d_1 y, d_2 y, \ldots, d_D y\right). \tag{28}$$

The generalized hypergeometric function Φ_2 is defined with a sum over D-dimensional space

$$\Phi_2(b_1, \ldots, b_D; c; x_1, \ldots, x_D)$$
$$= \sum \frac{(b_1)_{m_1} \cdots (b_D)_{m_D}}{(c)_{m_1+\cdots+m_D} m_1! \cdots m_D!} x_1^{m_1} \cdots x_D^{m_D} \tag{29}$$

where the summation Σ is a D-fold summation with m_1 through m_D each running from 0 to ∞, and the Pochhammer symbol is defined as

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}. \tag{30}$$

Although (28) provides a general solution, calculation of the generalized hypergeometric function is computationally restrictive. We therefore seek special cases to provide more practical solutions.

If all of the nonzero eigenvalues have an even number of multiplicities, the pdf can be developed as a partial fraction expansion [10]; this approach was demonstrated by [11]. Defining $h_i = \frac{\nu_i}{2}$, (27) can be written as

$$M(s) = g \sum_{i=1}^{D} \sum_{j=1}^{h_i} \frac{A_{ij}}{(s - d_i)^j} \tag{31}$$

where

$$A_{ij} = \frac{1}{(h_i - j)!} \left[\frac{d^{h_i - j}}{ds^{h_i - j}} \left[(s - d_i)^{h_i} \frac{M(s)}{g} \right] \right]\Bigg|_{s = d_i}. \tag{32}$$

The derivatives required in A_{ij} can be be written as

$$D_i^{(m)}(s) = \frac{d^m}{ds^m} \prod_{k=1, k \neq i}^{D} (s - d_k)^{-h_k}, \quad m \geq 0 \tag{33}$$

where the higher order derivatives can be computed recursively from $D_i^{(0)}(s)$

$$D_i^{(m+1)}(s) = \sum_{n=0}^{m} \binom{m}{n} D_i^{(n)}(s) \sum_{k=1, k \neq i}^{D} h_k \frac{(m - n)!}{(d_k - s)^{m - n + 1}}$$
$$m \geq 0. \tag{34}$$

Thus, the coefficient A_{ij} can be written as

$$A_{ij} = \frac{1}{h_i - j} \sum_{n=0}^{h_i - j - 1} \frac{D_i^{(n)}(d_i)}{n!} \sum_{k=1, k \neq i}^{D} \frac{h_k}{(d_k - d_i)^{h_i - j - n}}$$
$$i = 1 \ldots D, \quad j = 1 \ldots h_i. \tag{35}$$

Taking the inverse Laplace transform of (31), the pdf of a power estimate is

$$f_P(p) = g \sum_{i=1}^{D} \sum_{j=1}^{h_i} A_{ij} \frac{1}{(j-1)!} p^{j-1} e^{d_i P} U(p). \tag{36}$$

REFERENCES

[1] P. D. Welch, "The use of the fast Fourier transform for the estimation of power spectra," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, pp. 70–73, June 1967.
[2] T. S. Durrani and J. M. Nightingale, "Probability distributions for discrete Fourier spectra," *Proc. Inst. Elect. Eng.*, vol. 120, no. 2, pp. 299–311. Feb. 1973.
[3] T. S. Durrani, "Joint density functions for digital spectra," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 314–320, Oct. 1974.
[4] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*.   Reading, MA: Addison-Wesley, 1991.
[5] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*.   New York: Academic, 1994.
[6] A. K. Jain, *Fundamentals of Digital Image Processing*.   Englewood Cliffs, NJ: Prentice-Hall, 1989.
[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*.   New York: Wiley, 1991.
[8] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*.   New York: McGraw-Hill, 1991.
[9] G. E. Roberts and H. Kaufman, *Table of Laplace Transforms*.   Philadelphia, PA and London, U.K.: W. B. Saunders, 1966.
[10] A. V. Oppenheim and A. S. Willsky, *Signals and Systems*.   Englewood Cliffs, NJ: Prentice-Hall, 1983.
[11] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrika*, vol. 48, nos. 3 and 4, pp. 419–426, 1961.

Paul E. Johnson (S'96) received the B.S. and M.S. degrees in electrical and computer engineering in 1993 and 1994, respectively, from Brigham Young University (BYU), Provo, UT. He is pursuing the Ph.D. degree in electrical engineering from BYU, with emphases in estimation theory, electromagnetics, signal processing, and remote sensing.

He is currently working at Ball Aerospace & Technologies Corporation, Boulder, CO, in antenna system design and algorithm development.

David G. Long (SM'98) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1989.

From 1983 to 1990, he worked for NASA's Jet Propulsion Laboratory (JPL), Pasadena, CA, as a Radar Systems Engineer. He was the Project Engineer (senior Technical Manager) for the NASA Scatterometer (NSCAT) project and was responsible for the high-level design, analysis, and technical management of the overall NSCAT Project. His responsibilities included developing and overseeing the instrument design and fabrication, algorithm development and coding for the ground processing system, mission operations, and data analysis. As a Group Leader, he was responsible for system performance analysis, high-level design, development and maintenance of system requirements, and supervision of system engineers working on several JPL flight projects. He was the Experiment Manager and Project Engineer for the SCANSCAT scatterometer (now known as SeaWinds), which was part of NASA's Mission to Planet Earth. Since 1990, he has been a faculty member with the Electrical and Computer Engineering Department, Brigham Young University, Provo, UT, where he teaches upper-division and graduate courses and conducts research in microwave remote sensing and is an active member of the microwave remote sensing community. He is a member of a number of NASA Science Teams including the NSCAT Science Working Team, the SeaWinds Science Team, and the TRMM science team. He is the primary investigator on a number of NASA-funded projects and is Head of the Microwave Earth Remote Sensing Laboratory. His research interests include microwave remote sensing, polar ice, signal processing, estimation theory, and mesoscale atmospheric dynamics. He has well over 100 publications in the open literature.

Dr. Long has received the NASA Certificate of Recognition several times.