Correspondence

A Generalized Framework for Power Spectral Estimation
ALBERT H. NUTTALL AND G. CLIFFORD CARTER

Abstract—This correspondence presents a generalized framework for power spectral estimation and shows how three previous estimation methods fit into this framework as special cases. Further, this correspondence clarifies some recent misleading comments made on spectral estimation; in particular, specific references are given that strongly support the use of overlapped weighted segment averaging for spectral estimation.

The purpose of this correspondence is to present a generalized framework for power spectral estimation and to show how three previous estimation methods (to be discussed) fit into this framework. In addition, this correspondence clarifies certain misleading comments [1] recently made on spectral estimation. The criticism [1] of weighted segment averaging was based on an implied computational constraint of utilizing no overlap; therefore, the criticism is misleading. References to empirical and theoretical work which give conclusive support to the widespread use of modern methods for spectral estimation are presented.

In our generalized framework we are concerned with both auto and cross spectral estimation; hence, we consider two discrete random processes. As is often the case in practice, we are limited to a single time-limited realization (TLR) of each random process. Within our generalized framework for power spectral estimation, we first partition each TLR into $N$ segments, where $N$ may be unity or $N$ may be very large, and the segments may be overlapped. Second, each segment is multiplied by a time-weighting function (the weighting function may be unity everywhere within the segment or it may be smooth, as for example, Hannning weighting). Third, the discrete Fourier coefficients (DFC) are computed for each weighted segment via an appropriate algorithm such as the FFT (optionally each segment can be appended with zeros). Further, the DFC's for one TLR segment are multiplied by the complex conjugate of the DFC's for the other segment (or same segment for auto spectra). Fifth, the complex products are averaged over the $N$ available segments (one segment if $N=1$). Next, the result-ant spectral estimates are transformed into the correlation (or lag) domain, where they are multiplied by a lag-weighting function (which may be unity). Finally, the results are transformed back into the frequency domain. (Alternatively, the last two steps can be replaced by a convolution in the frequency domain. Depending on the extent of the frequency domain convolution, the former alternative may be computationally preferable over the latter.) It can be shown that the effective window in this generalized technique, in so far as its effect upon the mean spectral estimate, is equal to the magnitude-squared linear window convolved with the quadratic window.

We now point out how three previous spectral analysis techniques fit into this generalized framework. First, the Blackman and Tukey (BT) method [2] allows for only one segment with rectangular time weighting over the entire record (from each TLR), and it applies a smooth lag-weighting function in the correlation domain, which goes to zero well before the end of the data record. (We note that, historically, the BT approach was not done by transforming into the frequency domain, but since it is faster than the original time domain BT method, it is a viable and equivalent alternative approach.)

In the second method the modern overlapped-segment averaging technique applies a smooth multiplicative time weighting to each of a large number of segments, and averages the DFC products from these overlapped segments to obtain a final spectral estimate, without employing additional lag weighting. We emphasize that the time weighting is typically Hannning, as proposed in [3] and [4] and not a ten percent cosine taper, as proposed in [4]; that is, in this method, we strongly recommend against a time weighting that smoothly rises to a nonzero constant, stays at that constant for a large fraction of the time segment, and then smoothly falls to zero because it yields very poor sidelobe behavior. Overlap is extremely important in the second method in order to realize maximum stability of the spectral estimate.

The third technique of interest, a special case of which is advocated by Yuen [1], recognizes that the number of available data points may be so large as to preclude the normal BT method in practical situations. He segments the data without overlapping and applies rectangular time weighting to each segment (commonly but erroneously thought of as no time weighting). To partially undo the bad sidelobe effects that this rectangular time weighting causes, and to gain additional stability, he then Hanns the segment-averaged power spectral estimate. We would like to point out that, more generally, the segment-averaged power spectrum can be transformed into the correlation (or lag) domain, where a smooth multiplicative lag-weighting function can be applied before transforming back into the frequency domain; while not normally done, this lag-domain "reshaping" is an acceptable method for almost completely undoing the bad sidelobe effects of rectangular time weighting.

Yuen believes that the third method with no overlapping and Hanned spectra is to be preferred. In fact, he states [1] that with time weighting, the second method (which he constrains to use no overlap) is an "unsound idea" with "serious shortcomings" which, while intuitively attractive, is "statistically inconclusive." He also states that "there has been no theoretical or computational demonstration that overlapping produces a significant advantage" and "attempts at mathematical proof usually require a great deal of simplifying assumptions which weaken the result." We disagree with all these quoted statements. We also feel there is not yet sufficient published evidence to support a preference for the third method. Some preliminary investigation by us has revealed that with proper correlation reshaping (lag weighting); this third method may turn out to be the most cost-effective.

The key assumptions for an investigation of spectral analysis are stationarity, Gaussian random processes, and a large product of observation time and desired resolution bandwidth. Since we are interested in spectral estimates, second-order stationarity is required. Since one must investigate variances of second-order quantities for stability determination, fourth-order moments of the random processes are required; hence, for mathematical tractability, the Gaussian assumption is needed. And in practice, to have meaningful spectral estimates with any method, large observation-time resolution-bandwidth

U.S. Government work not protected by U.S. copyright
products are required. These assumptions do not significantly weaken the results; rather, they are the practical essentials of any statistical analysis of spectral estimation.

Under these assumptions, Nuttall [5] has shown that the second method (time-weighted overlapped-segment averaging method) can achieve, in particular, the number of equivalent degrees of freedom (EDF) as the BT spectral estimation method for both auto and cross spectral estimation if the proper overlap is used for each weighting, when both methods operate on the same amount of data and are constrained to the same frequency resolution. For many practical time weightings, most of the maximum EDF can be attained by a computationally reasonable amount of overlap. For example, with Hanning weighting, 92 percent of the maximum EDF can be realized with 50 percent overlap. And for Parzen (cubic) weighting, 93 percent of the maximum EDF can be realized with 62.5 percent overlap. Furthermore, the number of FFT's required is virtually independent of the particular time weighting employed (with its optimum overlap), but depends only upon the observation-time resolution-bandwidth product [5]. The only tradeoff between the various weightings, each with its inherent sidelobe structure, is that those weightings with better sidelobes require larger size FFT's. But even here, the larger size FFT's are very reasonable; for example, the Hanning and cubic weightings require 1.63 and 2.05 times the size of the FFT required for rectangular weighting [5]. Thus the computational load of the second method, even for smooth time weightings with good sidelobe structure and near-optimum overlap, is very reasonable.

The Yuen method, which segments the data but employs no overlapping, must lose at least a small amount of EDF due to edge effects. That is, the data points at the end of one segment never interact with those at the beginning of the next segment; hence, for all nonzero lags, there are not as many degrees of freedom in the Yuen method as in the BT method. This could be almost completely overcome by using overlap or by taking larger size segments.

For more complicated spectral measures, such as coherence, the analysis becomes unwieldy, and one is driven to simulation. (However, some recent analysis in this area has been accomplished by Lugannani [6].) In particular, Carter, Knapp, and Nuttall [3] empirically investigated the effect of overlap on the coherence estimate via the second method. There was a pronounced improvement (about a factor of two in variance reduction) with overlap as opposed to no overlap. These results, smooth weightings and significant overlap, explain the apparent discrepancy presented by Yuen in his 1978 Technometrics work [7]. In particular, (6) of [7] does not allow for overlap; this in turn explains why Fig. 4 of [7] showed that linear weighting (second method) produced twice the variance of the BT method. In another demonstration, Carter and Knapp [8] showed empirically the effect of estimating coherence between a flat broad-band input to a second-order digital filter and its output using Hanning and rectangular weightings. Conclusively, and without a doubt, smooth weighting functions were required with the second method. (Recall that the second method does not rely on additional lag weighting.)

The weighted overlapped-segment averaging WOSA method with proper overlap can attain the EDF of the BT method, when both methods operate on the same amount of data and are constrained to have the same frequency resolution. Further, for good time weightings, reasonable amounts of overlap achieve most of the available EDF. However, it is likely that there exists some tradeoff between lag weighting and time weighting in the generalized framework in terms of speed and storage requirements; further study of this tradeoff is required. It appears, based on preliminary independent analytic work, that the third method can virtually attain the EDF of the BT method and yield very good sidelobes through the use of proper lag weighting. It may also be faster to implement than the second method in practice, and therefore deserves consideration. However, Yuen's dismissal of the WOSA method, for reasons of greater instability when no overlap was used, is very misleading. Any preference for the third method, based on this type of rejection of the second method, is incorrect; rather, any preference must be based on a thorough analytic and simulation comparison.

REFERENCES

The Power Spectral Density of a Random Sequence of Zeros and Ones

GYORGY HENXER

Abstract—The statistical properties of a randomly sampled sequence are conveniently studied by expressing the randomly sampled sequence as the time term product of uniformly spaced signal values with a random phase.

Manuscript received September 19, 1978; revised March 9, 1979 and November 13, 1979.

The author is with the Department of Electrical Engineering, Ben Gurion University of the Negev, Be'er Sheva, Israel.

0096-3518/80/0600-0335$00.75 © 1980 IEEE