

TABLE I
SCER FOR VARIOUS SCHEMES AND N

N (m)	8 (200)	16 (200)	32 (200)	64 (100)	128 (100)	256 (50)	512 (30)	1024 (20)
SM	48.86	44.86	41.22	37.75	34.56	31.42	28.36	25.26
2C (α)	53.81	50.30	46.96	43.83	40.78	37.86	35.08	33.20
2C (β)	53.81	49.93	46.65	43.54	40.53	37.63	34.96	32.13
2C δ 1C	51.32	47.47	43.99	40.76	37.68	34.53	31.63	28.61
2CRM (α)	47.54	44.89	42.08	39.96	37.80	35.73	33.62	31.37
2CRM (β)	47.46	44.58	41.82	39.61	37.48	35.46	33.27	31.27

Note: m —number of FFT's averaged, α —coefficient resolution 12 bits, β —coefficient resolution 10 bits.

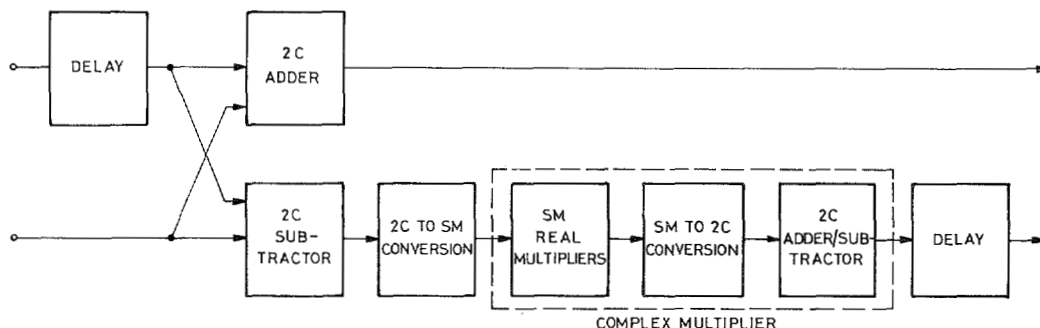


Fig. 3. 2C processing with SM multiplier in an FFT butterfly.

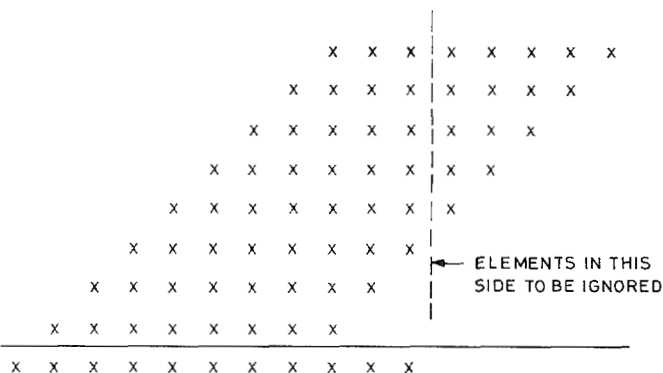


Fig. 4. Reduced multiplier array with rounding off.

If, instead of the 2C to SM conversion in Fig. 3, a 2C to 1C conversion is made, this gives a saving in hardware, though the SCER goes down, as shown in Table I for the scheme 2C δ 1C. However, this SCER is still better than that of SM by about 3 dB. The SM to 2C conversions can be coupled with the a/s which come after the real multipliers, to give some reduction in hardware [2].

Simulation studies were carried out for different lengths of the coefficient word (real and imaginary) namely, 10, 11, and 12 bits. The SCER for the 12-bit case gave very negligible improvement over the 11-bit case. The SCER for 10- and 12-bit cases is given in Table I. These results indicate that the coefficient wordlength need not be greater than 11 for 12-bit processing.

The $(2b - 1)$ -bit multiplier output is rounded off to b bits. Due to this, the contribution of the least significant columns (lsc) of the multiplier array to the output is very negligible. Therefore, some of the lsc's can be simply removed from the multiplier array as indicated in Fig. 4. Detailed computer simulation has shown that the removal of as many lsc's as to constitute 25 percent of the array elements causes a reduc-

tion in SCER by at most, only 0.3 dB. Any further reduction affects SCER severely, as seen from Table I for the scheme 2CRM (2C with reduced multiplier) with 8 columns (about 30 percent array elements) removed from a 12×12 -bit SM multiplier. This loss of SCER reduces with increasing N . It is felt that this is due to increased interleaving of errors for larger N .

Thus, the techniques described here will help the designer of the FFT hardware in choosing various schemes considering hardware cost and computational accuracy.

REFERENCES

- [1] G. C. O'Leary, "Non-recursive digital filtering using cascade FFT transformers," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 177-183, June 1970.
- [2] J. P. Agrawal and J. Ninan, "Hardware considerations in FFT processors," presented at the IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Philadelphia, PA, Apr. 1976.
- [3] A. V. Oppenheim and C. J. Weinstein, "Effects of finite register length in digital filtering and the fast Fourier transform," *Proc. IEEE*, pp. 957-976, Aug. 1975.
- [4] J. P. Agrawal and V. U. Reddy, "Log-sum multiplier," presented at the Nat. Computer Conf., New York, NY, June 1976.

Improved Approximation of Bias in Squared Coherence Estimates for Weakly Smooth Spectra

PETER LÖHNBERG

Abstract—Bias in squared coherence estimates for normal processes is approximated by a function of coherence and second derivatives divided

Manuscript received February 17, 1977; revised September 29, 1977 and November 23, 1977.

The author is with the Department of Electrical Engineering, Twente University of Technology, Enschede, The Netherlands.

by values of spectral densities which may vary slightly over the window width. The result approximates more accurately than a known formula results from nonrefuted literature concerning very smooth spectra.

INTRODUCTION

The squared coherence at a certain frequency is defined as

$$C = \frac{G_{xy}G_{yx}}{G_{xx}G_{yy}} \quad (1)$$

with the *spectra*: cross spectral density G_{xy} with complex conjugate G_{yx} and power spectral densities G_{xx} and G_{yy} of stationary real random processes x and y .

Spectra can be estimated under ergodicity assumption from a realization of x and y over time T . An efficient procedure consists of dividing this interval into N equally long segments, forming products per segment of FFT's and their complex conjugates, multiplication of inverse FFT's by a lag window, and averaging over FFT's of all N results. Substitution of estimates \hat{G}_{xy} , \hat{G}_{xx} , and \hat{G}_{yy} into (1) gives estimate \hat{C} .

For optimal window selection and bias correction it is desirable that bias $[\hat{C}]$ and var $[\hat{C}]$ are known for arbitrary spectra. Bias $[\hat{C}]$ was approximated¹ in [1] with weak restrictions on spectral variations over the *window width*, defined as the width of the range of frequencies f for which the Fourier transform $W(f)$ of the lag window has significant values. Stronger restrictions were made in an exact derivation [2] and apparently white processes were used in experiments [3], [4]. References [2]–[4] showed bias $[\hat{C}]$ to be a decreasing function of C reaching 0 for $C = 1$. But according to [1]¹, for these stronger restrictions bias $[\hat{C}]$ would be independent of C . So use of [1]¹ might cause serious errors.

Therefore, an improved expression for bias $[\hat{C}]$ is required which indicates effects of spectral variations over the window width. Extension of the exact derivation to weaker restrictions causes mathematical problems, whereas repetition of experiments for a sufficient number of combinations of spectra is hardly feasible. A more accurate approximation for even stronger spectral variations than [1]¹ is possible however; for in [1]¹ only a first-order Taylor approximation of $\hat{C}(\hat{G}_{xy}, \hat{G}_{yx})$ at point G_{xy}, G_{yx} was used. It is only slightly more difficult to approximate \hat{C} as a function of all spectral estimates by a second-order Taylor series. This expresses bias $[\hat{C}]$ into first and second moments of spectral estimates, for which expressions are known. These moments were already used maximally in [1] to approximate² var $[\hat{C}]$. Therefore, only bias $[\hat{C}]$ will be approximated here such as to make maximal use of these moments.

APPROXIMATE BIAS OF SQUARED COHERENCE ESTIMATE

For notational simplicity, define for estimates of arbitrary quantities X and Y the normalizations $B[\hat{X}] = \text{bias}[\hat{X}/X]$ and³ $V[\hat{X}, \hat{Y}] = \text{cov}[\hat{X}/X, \hat{Y}/Y]$.

Then, by second-order Taylor expansion of $\hat{C}(\hat{G}_{xy}, \hat{G}_{yx}, \hat{G}_{xx}, \hat{G}_{yy})$ at point $G_{xy}, G_{yx}, G_{xx}, G_{yy}$ follows

$$\begin{aligned} B[\hat{C}] &\simeq B[\hat{G}_{xy}] + B[\hat{G}_{yx}] - B[\hat{G}_{xx}] - B[\hat{G}_{yy}] \\ &+ B[\hat{G}_{xy}]B[\hat{G}_{yx}] + V[\hat{G}_{xy}, \hat{G}_{yx}] \\ &- B[\hat{G}_{xy}]B[\hat{G}_{xx}] - V[\hat{G}_{xy}, \hat{G}_{xx}] \\ &- B[\hat{G}_{xy}]B[\hat{G}_{yy}] - V[\hat{G}_{xy}, \hat{G}_{yy}] \end{aligned}$$

¹Formula (9.3.25) in [1].

²Formula (9.2.19) in [1].

³Here, contrary to [1], the definition of covariance for complex random variables is used.

$$\begin{aligned} &- B[\hat{G}_{yx}]B[\hat{G}_{xx}] - V[\hat{G}_{yx}, \hat{G}_{xx}] \\ &- B[\hat{G}_{yx}]B[\hat{G}_{yy}] - V[\hat{G}_{yx}, \hat{G}_{yy}] \\ &+ B^2[\hat{G}_{xx}] + V[\hat{G}_{xx}, \hat{G}_{xx}] + B^2[\hat{G}_{yy}] \\ &+ V[\hat{G}_{yy}, \hat{G}_{yy}] + B[\hat{G}_{xx}]B[\hat{G}_{yy}] \\ &+ V[\hat{G}_{xx}, \hat{G}_{yy}]. \end{aligned} \quad (2)$$

Bias of spectral estimates may be derived similarly to [5]. This requires small variation of the spectra over frequency width N/T and of second derivatives to frequency G''_{xy}, G''_{xx} , and G''_{yy} over the window width. Then

$$B[\hat{G}_{xy}] \simeq SG''_{xy}/G_{xy} \quad (3)$$

and analogous expressions for \hat{G}_{xx} and \hat{G}_{yy} with systematic error constant

$$S = \frac{1}{2} \int_{-\infty}^{\infty} f^2 W(f) df.$$

S is obtainable for usual windows by substitution of the values of " D " from [5] into $S = "D"/(2\pi)^2$.

Covariance of spectral estimates also has been derived [1].⁴ For that it is required that G_{ac} and G_{bd} of arbitrary normal processes a, b, c , and d vary little over the window width. Then

$$V[\hat{G}_{ab}, \hat{G}_{cd}] \simeq RG_{ac}G_{db}/G_{ab}G_{dc} \quad (4)$$

with random error constant

$$R = \frac{1}{T} \int_{-\infty}^{\infty} W^2(f) df < 1.$$

R is obtainable for usual windows by substitution of the values of " I " from [5] into $R = "I"/T$. If no window is applied [2]–[4], $R = 1/N$.

Substitution of (3), (4) into (2) yields

$$\begin{aligned} B[\hat{C}] &\simeq SG''_{xy}/G_{xy} + SG''_{yx}/G_{yx} - SG''_{xx}/G_{xx} - SG''_{yy}/G_{yy} \\ &+ S^2G''_{xy}G''_{yx}/G_{xy}G_{yx} + R/C \\ &- S^2G''_{xy}G''_{xx}/G_{xy}G_{xx} - R - S^2G''_{xy}G''_{yy}/G_{xy}G_{yy} - R \\ &- S^2G''_{yx}G''_{xx}/G_{yx}G_{xx} - R - S^2G''_{yx}G''_{yy}/G_{yx}G_{yy} - R \\ &+ S^2G''_{xx}/G_{xx}^2 + R + S^2G''_{yy}/G_{yy}^2 + R \\ &+ S^2G''_{xx}G''_{yy}/G_{xx}G_{yy} + RC. \end{aligned} \quad (5)$$

Substitution into (5) of the result from (1) $G_{xy} = \sqrt{CG_{xx}G_{yy}} e^{j\text{Arg}(G_{xy})}$ where $j = \sqrt{-1}$ expresses $B[\hat{C}]$ into independent variables. This was partly used for [1].¹

Equation (5) shows that relative, spectral second and not first derivatives influence bias. As expected, bias is decreased by decreasing S or R . However, for fixed T and window type decrease of either increases the other. So for each T , an optimal window for minimal coherence bias might be derived from (5) with appropriate spectral information. By also taking into account var $[\hat{C}]$, the mean-square error in \hat{C} might be minimized instead.

COMPARISON WITH LITERATURE FOR VERY SMOOTH SPECTRA

Several simplifications are needed for comparison of (5) and [1]¹ to [2]–[4]. For [1]¹ G_{xy} was required to be smooth, that is $S|G''_{xy}/G_{xy}| \ll 1$. Then the fifth term in (5) can be

⁴Formula (A.9.1.28) in [1].

omitted. Furthermore, by the rough approximation, only the first, second, and sixth terms in (5) were considered for [1].¹

References [2]–[4] concern *very smooth* spectra, that is $S|G''_{xy}/G''_{xy}| \ll R$ and G''_{xx} and G''_{yy} similar. Then all systematic error terms containing S and second derivatives may be omitted, and only the random error terms in (5) remain. So according to (5)

$$\text{bias}[\hat{C}]_{\text{very smooth}} = CB[\hat{C}]_{\text{very smooth}} \simeq R(1-C)^2 = F(C). \quad (6)$$

The version of [1]¹ for very smooth spectra is

$$\text{bias}[\hat{C}]_{\text{very smooth}} \simeq R. \quad (7)$$

Approximation $F(C)$ in (6) is R for $C=0$ and 0 for $C=1$. This agrees with theoretical bias $[\hat{C}]$, [2], with simulation results for $R = \frac{1}{4}$, [4], and even with degenerate procedure effect $\hat{C} = 1$ for $R = 1$. But for $0 < C < 1$, $F(C)$ is lower than bias according to all of these. These lower values must be due to approximation (2) and possibly (4). Nevertheless, $F(C)$ differs less from bias $[\hat{C}]$ according to [2], [4] than approximation R in (7) does for any R and $C > 0$.

Reference [3] resulted in empirical estimates of bias $[\hat{C}]$ with fit⁵ $R(1-\hat{C})$. For very low C these results are slightly lower, and for high $C < 1$, higher than bias $[\hat{C}]$ according to [2], [4] and (6). These deviations increase with R . They must at least partly be due to the estimation in [3] of bias $[\hat{C}]$ by the inverse Z^{-1} of the bias in $Z(\hat{C}) = \frac{1}{2} \ln \{(1 + \sqrt{\hat{C}})/(1 - \sqrt{\hat{C}})\}$. This procedure implies use of the approximation $E[\hat{C}] \simeq Z^{-1}\{E[Z(\hat{C})]\}$. The resulting approximation error in bias $[\hat{C}]$ increases with the second derivative of $Z(\hat{C})$ to \hat{C} and with $\text{var}[\hat{C}]$, which is proportional to R [1].² These expected increases agree with the indicated ones of the actual deviations. Therefore, [3] will not be used for comparison.

CONCLUSION

Equation (5) approximates bias $[\hat{C}]$ of normal processes for values and second derivatives of spectra varying slightly over the window width. The derivation uses a more extensive and higher order Taylor approximation than [1]. Therefore, result (5) is expected to be more accurate than [1].¹ This might be supported by very smooth spectral version (6) of (5) agreeing closer with [2], [4] than corresponding version (7) of [1]¹ does. A reason for the deviation of [3] from [2], [4], and (6) has been indicated.

ACKNOWLEDGMENT

The author wishes to thank Prof. E. W. Gröneveld, G. H. L. M. Heideman, and J. de Haas for their sincere interest in this subject, and Miss G. G. M. Steijlen for typing the manuscript.

REFERENCES

- [1] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*. San Francisco: Holden-Day, 1968.
- [2] G. C. Carter and A. H. Nuttall, "Statistics of the estimate of coherence," *Proc. IEEE (Lett.)*, vol. 60, pp. 465–466, Apr. 1972.
- [3] V. A. Benignus, "Estimation of the coherence spectrum and its confidence interval using the fast Fourier transform," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 145–150, June 1969.
- [4] A. H. Nuttall and G. C. Carter, "Bias of the estimate of magnitude-squared coherence," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 582–583, Dec. 1976.
- [5] A. Papoulis, "Minimum-bias windows for high-resolution spectral estimates," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 9–12, Jan. 1973.

⁵Incorrectly cited in [4] as $B_2 = R(1-C)$.

Notes on Linear Image Restoration by Maximizing the *A Posteriori* Probability

H. J. TRUSSELL

Abstract—A distinction is made between the minimum mean-square error estimate derived for the linear case of maximum *a posteriori* restoration and the common Wiener estimate. The linear MAP filter is shown to produce good results with less *a priori* knowledge than is required for the Wiener filter.

Minimum mean-square error filters can be derived from both *a priori* and *a posteriori* knowledge. The minimum mean-square error filter derived from *a priori* knowledge is usually referred to as a Wiener filter [1]. This filter is a standard against which most other image restoration/enhancement filters are compared. The minimum mean-square error filter derived from *a posteriori* knowledge is the same as the filter which maximizes the *a posteriori* probability for a linear imaging system. Because of its relation to the maximum *a posteriori* (MAP) estimation method, we will refer to this filter as the linear MAP filter. This correspondence shows that restorations of quality comparable to Wiener restorations can be obtained by using the linear MAP filter. In actual implementation, the linear MAP restoration uses less *a priori* knowledge than the Wiener restoration; furthermore, it is less sensitive to perturbations of that *a priori* knowledge.

Hunt [2] derived the maximum *a posteriori* (MAP) equation for nonlinear image restoration. For the linear case (1), using lexicographic notation [3], the imaging system is described by

$$g = Hf + n \quad (1)$$

where

- f is the original image drawn from a Gaussian ensemble with mean \bar{f} and covariance R_f ;
- H is the point-spread function matrix;
- n is zero-mean Gaussian noise with variance/covariance R_n ;
- g is the recorded image.

The linear MAP estimate is then given by

$$\hat{f}_{\text{MAP}} = (R_f^{-1} + H^T R_n^{-1} H)^{-1} (H^T R_n^{-1} g + R_f^{-1} \bar{f}). \quad (2)$$

The Wiener estimate with ensemble mean \bar{f} included can be shown [see [1] and model (1)] to be

$$\hat{f}_W = (R_f H^T + \bar{f} \bar{f}^T H^T) (H R_f H^T + H \bar{f} \bar{f}^T H^T + R_n)^{-1} g. \quad (3)$$

If $\bar{f} = 0$, as is usually assumed in Wiener estimation, then $\hat{f}_{\text{MAP}} = \hat{f}_W$.

The difference between the linear MAP filter and the usual form of the Wiener filter is the nonzero mean \bar{f} . The model used for the derivation of the Wiener filter assumes a zero-mean (or a constant-mean which can be subtracted). This implies that all spatial variation of the image ensemble is accounted for by the covariance matrix R_f . Since the mean \bar{f} has been assumed to have an extremely simple structure, the covariance must necessarily have a great deal of structure.

The model used for deriving the linear MAP filter assumes a nonstationary mean \bar{f} . For computation it is assumed that $f_D = f - \bar{f}$ is a stationary process. Since some of the ensemble

Manuscript received February 14, 1977; revised May 17, 1977 and November 1, 1977.

The author is with Los Alamos Scientific Laboratory, University of California, Los Alamos, NM 87545.