Substituting (A4) and (A5) in (A1) and using (A2) to eliminate the variable δ_1, (A1) becomes the equation

$$F(N, \delta_2) = 0. \tag{A7}$$

Eliminating in the same manner δ_1 from (4), the solution of (A7) and (4) gives through (A4) the values N_{opt} and t_{opt}. Also in this case, the problem is well behaved (see Fig. 1) and the solution of the nonlinear system can be obtained quickly by the bisection method.

References

[1] D. S. K. Chan and L. R. Rabiner, "Analysis of quantization errors in the direct form for finite impulse response digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 354–365, Aug. 1973.

[2] R. E. Crochiere and A. V. Oppenheim, "Analysis of linear digital networks," *Proc. IEEE*, vol. 63, pp. 581–595, Apr. 1975.

[3] O. Herrmann, L. R. Rabiner, and D. S. K. Chan, "Practical design rules for optimum finite impulse response lowpass digital filters," *Bell Syst. Tech. J.*, vol. 52, pp. 769–799, July–Aug. 1973.

[4] R. A. Gabel, "On the optimal number of FIR filter coefficients under a memory size constraint," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 366–367, Aug. 1978.

[5] O. Andrisano and L. Calandrino, "Tap weight tolerance effects on CCD transversal filtering," *Alta Freq.*, vol. 45, pp. 739–746, Dec. 1976.

[6] A. Gersho, B. Gopinath, and A. M. Odlyzko, "Coefficient inaccuracy in transversal filtering," *Bell Syst. Tech. J.*, vol. 58, pp. 2301–2316, Dec. 1979.

[7] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975.

[8] R. D. Baertsch, W. E. Engeler, H. S. Goldberg, G. H. Puckette, and J. J. Tiemann, "The design and operation of practical charge-transfer transversal filters," *IEEE J. Solid-State Circuits*, vol. SC-11, pp. 65–74, Feb. 1976.

[9] U. Heute, "Necessary and efficient expenditure for non-recursive digital filters in the direct form," in *Proc. ECCTD '74*, July 1974, pp. 13–19.

[10] F. Grenez, "Reduction of coefficient wordlength for FIR linear phase digital filters," in *Proc. ECCTD '78*, Sept. 1978, pp. 330–334.

[11] —, "Synthése des filtres numeriques non récursifs à coefficients quantifiés," *Ann. Télécommun.*, vol. 34, pp. 33–39, Feb. 1979.

[12] Y. Chen, S. M. Kang, and T. G. Marshall, "The optimal design of CCD transversal filters using mixed-integer programming techniques," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1978, pp. 748–751.

An Approximation to the Cumulative Distribution Function of the Magnitude-Squared Coherence Estimate

ALBERT H. NUTTALL AND G. CLIFFORD CARTER

Abstract—**We investigate a nonlinear distortion that converts the magnitude-squared coherence estimate to a near-Gaussian random variable. In particular, we present simple approximations for the mean and variance of this nonlinearly distorted magnitude-squared coherence estimate and fit a Gaussian cumulative distribution function over a wide range of parameters.**

Manuscript received September 15, 1980; revised January 13, 1981.

The authors are with the New London Laboratory, Naval Underwater Systems Center, New London, CT 06320.

Introduction

In sonar signal processing for detection and estimation the magnitude-squared coherence (MSC) plays an important role. Critical to its proper use is an understanding of its underlying statistical properties. The cumulative distribution function (CDF) of the estimate of MSC, obtained by averaging over N statistically independent pieces of Gaussian data, is available in [1] as a sum of $N-1$ hypergeometric functions. Direct evaluation of this quantity has recently been simplified in [2]; however, it is still a tedious and time-consuming calculation for large N and suffers from numerical overflow and underflow unless special care is taken in programming. Furthermore, no simple result for obtaining confidence limits is available.

In this correspondence we investigate the suggestion of Fisher [3] and Jenkins and Watts [4] that the nonlinearity arc tanh $(\sqrt{x})$ converts the MSC estimate to a near-Gaussian random variable. In particular, we present simple approximations for the mean and variance of this nonlinearly distorted MSC estimate and fit a Gaussian CDF over a wide range of: N, the number of pieces averaged in the MSC estimate; C, the true MSC; and P, the value of the CDF. Inversion of the Gaussian CDF affords a simple way of getting confidence limits for specified probabilities of threshold crossings.

Investigation of arc tanh $(\sqrt{x})$ Nonlinearity

Moments

The probability density function of the MSC estimate $\hat{C}$, as given in [1], is

$$p_1(x) = (N-1)\frac{(1-C)^N(1-x)^{N-2}}{(1-Cx)^{2N-1}} F(1-N, 1-N; 1; Cx)$$

$$\text{for } 0 < x < 1 \text{ and } N \geq 2 \tag{1}$$

where N is the number of pieces averaged, C is the true MSC, and F is a Gaussian hypergeometric function. In this case F is a polynomial in Cx.

The nonlinearly distorted version of random variable $\hat{C}$, which we are interested in, is

$$D = \text{arc tanh}\,(\sqrt{\hat{C}}) = \frac{1}{2}\ln\left(\frac{1+\sqrt{\hat{C}}}{1-\sqrt{\hat{C}}}\right). \tag{2}$$

If the nonlinear operation in (2) results in a Gaussian random variable for D, then we will be interested in the mean and variance of D. This problem is considered analytically in [5]; the only closed-form results that we have obtained are listed below. Let the mth moment of D be denoted by

$$\mu_m(N, C) = \overline{D^m} = \int_0^1 dx[\text{arc tanh}(\sqrt{x})]^m p_1(x) \tag{3}$$

where the overbar denotes a statistical average. Then for $C = 0$, the mean of D is

$$\mu_1(N, 0) = \frac{\sqrt{\pi}}{2}\frac{\Gamma(N-1)}{\Gamma(N-1/2)} \sim \frac{\sqrt{\pi}/2}{\sqrt{N-1.25}} \quad \text{as } N \to \infty, \tag{4}$$

the mth moment is

$$\mu_m(N, 0) \sim \frac{\Gamma(1+m/2)}{(N-1)^{m/2}} \quad \text{as } N \to \infty, \tag{5}$$

and the variance of D, for $C = 0$, is

$$\sigma^2(N, 0) = \mu_2(N, 0) - \mu_1^2(N, 0) \sim \frac{1 - \pi/4}{N - 1} \quad \text{as} \quad N \to \infty. \quad (6)$$

In order to deduce the fundamental behavior of the mean and variance of D for $C \neq 0$, we evaluated (3) numerically in [5]. The main results of the numerical investigation are listed below. We find mean

$$\mu_1(N, C) \cong \text{arc tanh}(\sqrt{C + B}) \qquad \text{for } C > 0 \quad (7)$$

where

$$B = \frac{1 - C^2}{2(N - 1)} \quad (8)$$

and variance

$$\sigma^2(N, C) \cong \frac{1}{2(N - 1)} \qquad \text{for } C > 0. \quad (9)$$

The imprecise qualifier $C > 0$ reflects the fact that the probabilistic behavior of $\hat{C}$ and D is distinctly different for $C = 0$ versus $C > 0$; for example, compare (6) and (9). In particular, as N becomes large, (9) becomes a better approximation for smaller C. The precise region where (7)–(9) are valid will become clear in later plots. The result (9) is a slight modification of Jenkins and Watts [4] that better fits the calculated values of variance and the slope of the CDF plots. Its independence of true MSC value C is striking and convenient.

Probabilistic Statements

If D is nearly a Gaussian random variable with mean $\mu_1(N, C)$ and variance $\sigma^2(N, C)$, then the CDF of D, i.e., the probability that D is less than some threshold T, is given by the approximation

$$P_2(T) \equiv \text{Prob}\{D < T\} \cong \int_{-\infty}^{T} \frac{dx}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma^2}\right)$$

$$= \Phi\left(\frac{T - \mu_1}{\sigma}\right) \quad (10)$$

where

$$\Phi(y) = \int_{-\infty}^{y} \frac{du}{\sqrt{2\pi}} \exp(-u^2/2) \quad (11)$$

is the CDF for a zero-mean unit-variance Gaussian random variable. A plot of (10) on normal probability paper, with T as the linear abscissa, yields a straight line with abscissa-intercept μ_1 and slope $1/\sigma$, since the nonlinear transformation to plot ordinates in this case is the inverse function $\Phi^{-1}(\ )$.

The exact CDF of D is given by

$$P_2(T) = \text{Prob}\{D < T\} = \text{Prob}\{\text{arc tanh}(\sqrt{\hat{C}}) < T\}$$

$$= \text{Prob}\{\hat{C} < \tanh^2(T)\} = P_1(\tanh^2(T)) \quad (12)$$

where we employed (1) and (2) and defined P_1 as the CDF of $\hat{C}$

$$P_1(A) = \text{Prob}\{\hat{C} < A\} = \int_0^A dx\, p_1(x). \quad (13)$$

Thus, we can relate CDF P_2 to the known CDF P_1 given in [1]. Computer programs for the exact evaluation of (12) and (13) are given in [5].

The availability of approximation (10) for the CDF of D enables us to give an approximation to the CDF P_1 of $\hat{C}$, namely

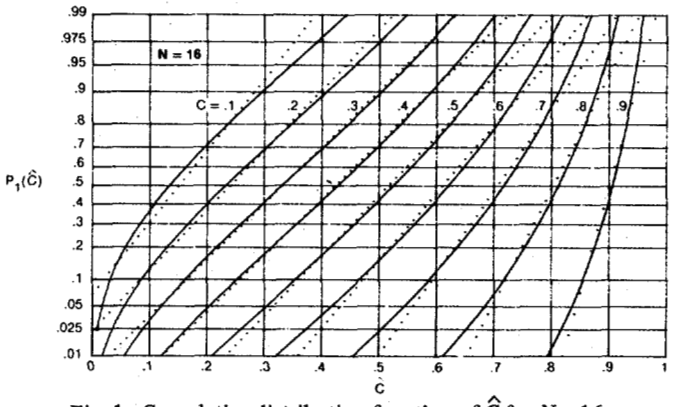

Fig. 1. Cumulative distribution function of $\hat{C}$ for $N = 16$.

$$P_1(A) = \text{Prob}\{\hat{C} < A\} = \text{Prob}\{D < \text{arc tanh}(\sqrt{A})\}$$

$$= P_2(\text{arc tanh}(\sqrt{A}))$$

$$\cong \Phi\left(\frac{\text{arc tanh}\sqrt{A} - \mu_1}{\sigma}\right) \quad (14)$$

where μ_1 and σ are given by (7)–(9) and Φ is as previously defined in (11).

Furthermore, we can now solve the inverse problem of determining a threshold A for specified values of CDF P_1 and for given parameter values C and N. From (14) we have

$$A \cong \tanh^2(\mu_1 + \sigma\Phi^{-1}(P_1)) \quad (15)$$

where μ_1 and σ are in (7)–(9). Simplification of (15) results in

$$A \cong \left(\frac{\alpha + \beta}{1 + \alpha\beta}\right)^2 \quad (16)$$

where

$$\alpha = \left[C + \frac{1 - C^2}{2(N - 1)}\right]^{1/2} \quad (17)$$

$$\beta = \tanh\left[\frac{\Phi^{-1}(P_1)}{\sqrt{2(N - 1)}}\right]. \quad (18)$$

Plots

Before embarking on the plots of exact CDF P_2 of distorted random variable D, as given by (12), we plot the exact CDF P_1 in (13) of the original random variable $\hat{C}$, as given by [1]. A plot on normal probability paper for $N = 16$ is given in Fig. 1. Plots for other values of N are given in [5]. In this figure each dotted straight line corresponds to a Gaussian random variable with mean and variance as derived for MSC estimate $\hat{C}$ in [6, p. 20]. The discrepancy with the exact CDF (in solid lines) indicates that $\hat{C}$ is not well approximated by a Gaussian random variable, especially for small N, and for the small and large values of C and the extreme values of probability near 0.01 and 0.99.

Plots of the exact CDF P_2 of D, given by (12), are presented in Figs. 2–4 for $N = 16$, 32, and 64. Dotted straight lines corresponding to a Gaussian random variable satisfying (7)–(10) have been drawn for every C and N value being considered; however, they have been overdrawn by the exact CDF (12) in some cases and are not visible. The agreement between exact and Gaussian CDF's is extremely good except for very low values of C and N. Plots for additional values of N are given in [5]; however, for $N \geq 128$ the agreement between the approximation and the exact result are so close as to be overlapping for the range of parameters considered here.

Notice that the curves in Figs. 2–4 are approximately parallel

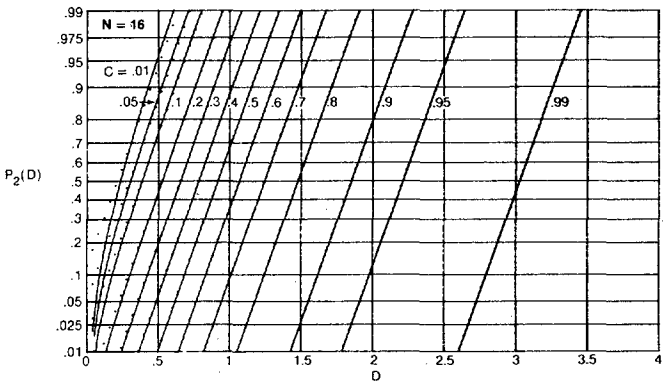

Fig. 2. Cumulative distribution function of D for $N = 16$.

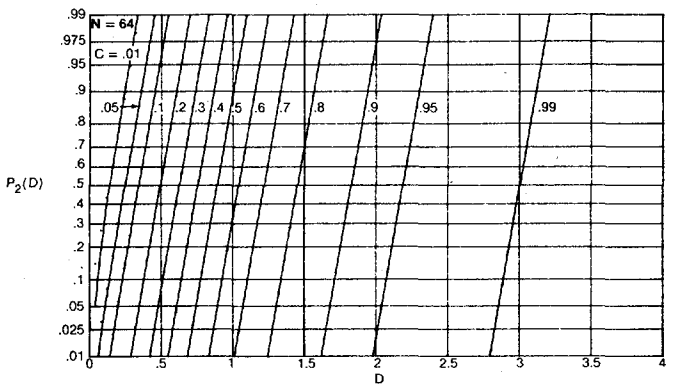

Fig. 3. Cumulative distribution function of D for $N = 32$.

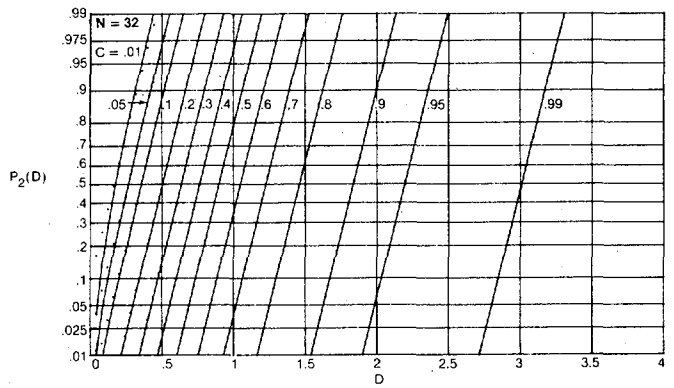

Fig. 4. Cumulative distribution function of D for $N = 64$.

straight lines over a wide range and can, therefore, be interpolated more easily. Thus, approximate probabilistic relation (10) and its inverse (16) are very useful and accurate for a wide range of C and N, encompassing most of the useful values of these parameters.

Summary and Discussion

We have investigated the suggestion of Fisher that the nonlinearity arc $\tanh(\sqrt{x})$ converts the MSC estimate to a near-Gaussian random variable. In particular, we presented simple approximations for the mean and variance of this nonlinearly distorted MSC estimate and fit a Gaussian CDF over a wide range of: N, the number of pieces averaged in the MSC estimate; C, the true MSC; and P, the values of the CDF. Inversion of the Gaussian CDF affords a simple way of getting confidence limits for specified probabilities of threshold crossings.

Since the arc $\tanh(\sqrt{x})$ nonlinear distortion takes no account of the known number N of pieces entering the MSC estimate, an improved nonlinear distortion which utilizes this information is possible which will convert the MSC estimate to more nearly a Gaussian random variable over a wider range of parameters N, C, and P. Evaluation of confidence limits requires the inversion of this nonlinearity; analytic inversion is not possible, but a numerical procedure converges rapidly. This more complicated procedure is the subject of a report by the first author [5].

References

[1] G. C. Carter, C. H. Knapp, and A. H. Nuttall, "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 337–344, Aug. 1973.

[2] P. F. Lee, "An algorithm for computing the cumulative distribution function for magnitude-squared coherence estimates," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 117–119, Feb. 1981.

[3] R. A. Fisher, *Contributions to Mathematical Statistics*. New York: Wiley, 1950; also "The general sampling distribution of the multiple correlation coefficient," *Proc. Roy. Soc., Ser. A*, vol. 121, pp. 654–673, 1928.

[4] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*. San Francisco, CA: Holden-Day, 1968.

[5] A. H. Nuttall, "Approximations to the cumulative distribution function of the magnitude-squared coherence estimate," NUSC Tech. Rep. 6327, August 25, 1980.

[6] A. H. Nuttall and G. C. Carter, "Approximations for statistics of coherence estimators," NUSC Tech. Rep. 5291, Mar. 9, 1976.

Intersection Filters for General Decimation/Interpolation

I. PAUL AND J. W. WOODS

Abstract—**Optimum multiple stopband FIR filters for decimation or interpolation by integer factors have been presented in the literature. In this correspondence we extend these results to the case of general decimation/interpolation by rational factors. The resulting filters are called intersection filters because their stopbands are intersections of the stopbands required in the case of integer factors. Significant improvement in stopband attenuation is found in many cases for band-limited signals.**

I. Introduction: The Efficient Interpolator

The use of FIR digital filters in interpolation is well known and has been well reported in the literature [1]–[3]. In these papers the frequency domain representation of the interpolation process is used to designate appropriate pass and stopbands for the FIR digital filters to be designed. In our analysis we retain the same approach.

Let us start with the initial sequence $x(n)$ and interpolate it by a rational factor $R = L/M$, where L and M are positive integers. Let the interpolated sequence be $y(n)$. Let $X(e^{j\omega})$ and $Y(e^{j\omega})$ be the Fourier transforms of $x(n)$ and $y(n)$, respectively. If we assume that $x(n)$ is band limited to ω_s, i.e., $X(e^{j\omega}) = 0$ for $\omega_s < |\omega| \leq \pi$, then $X(e^{j\omega})$ will have a form similar to that shown in Fig. 1(a) (specialized to $\omega_s = \pi/2$). The standard method for interpolating a sequence $x(n)$ by a rational factor consists of interpolating $x(n)$ by the integer factor L and then decimating the resulting sequence $w(n)$ by the integer factor M (schematically shown in Fig. 2). $W(e^{j\omega})$, the Fourier transform of the intermediate sequence $w(n)$, is of the form shown in Fig. 1(b) (with $L = 9$ in this example). To specify the filter $H(e^{j\omega})$ (Fig. 2) we need to define two parameters: the highest frequency ω_1 in $X(e^{j\omega})$ which has to

Manuscript received September 5, 1980; revised March 2, 1981.

The authors are with Rensselaer Polytechnic Institute, Troy, NY 12181.

0096-3518/81/0800-0934$00.75