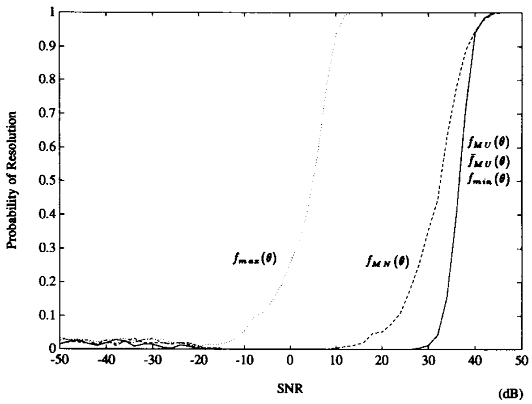

Fig. 4. Curves of probability of resolution versus SNR (in dB) for various null-spectra, $f_{\mathrm{MU}}(\theta), f_{\mathrm{MN}}(\theta), \bar{f}_{\mathrm{MU}}(\theta), f_{\min}(\theta)$, and $f_{\max}(\theta)$, when $N = 100, L = 10, M = 2$ ($\theta_1 = 15°$ and $\theta_2 = 17°$. Each curve was obtained at interval of 2 dB and each point was obtained from 500 trials).

VII. Summary

Using a perturbation matrix, we introduced a hyperplane used to define a generalized null-spectrum, based on both the signal and noise subspaces, while the MUSIC and Min-Norm null-spectra are defined based only on the noise subspace. With the generalized null-spectrum, we derived the upper and lower bounds of a class of the generalized null-spectrum, called the maximum and minimum null-spectra, respectively. From a brief analysis of a class of the generalized null-spectrum and a geometrical interpretation, we expected that the resolution capability of the maximum null-spectrum would be superior to that of other null-spectra. By computer simulation it is shown that the maximum null-spectrum has higher resolution capability than other null-spectra. In addition, it is seen that the asymptotic distributions of the estimates of the direction of arrival by using the maximum and multiple signal classification null-spectra are the same.

References

[1] K. M. Buckley and X. L. Xu, "Spatial-spectrum estimation in a location sector," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1842–1852, Nov. 1990.

[2] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the Minimum-Norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 331–341, Apr. 1986.

[3] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aero. Elect. Syst.*, vol. AES-19, pp. 134–139, Jan. 1983.

[4] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Ant., Propagat.*, vol. AP-34, pp. 276–280, Mar. 1986.

[5] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–740, May 1989.

Computation of a Useful Cramer-Rao Bound for Multichannel ARMA Parameter Estimation

Mrityunjoy Chakraborty and Surendra Prasad

***Abstract*—It has been shown earlier that the problem of multichannel autoregressive moving average (ARMA) parameter estimation can be tackled in a computationally efficient way by converting the given process into an equivalent scalar, periodic ARMA process. This correspondence presents methods to compute the Cramer-Rao bound associated with the identification of the scalar ARMA equivalent of a given multichannel ARMA process. The elements of matrix are obtained by a few very simple operations like periodic AR filtering of certain downsampled versions of the input and output sequences and then cross-correlating the filter outputs. The filter is easily obtainable from the model equation and is common for all the parameters.**

I. Introduction

Autoregressive moving average (ARMA) models have found wide applications in various signal-processing applications like system identification, speech processing, spectrum estimation, etc. While the single-channel ARMA estimation can be extended to the multichannel case in a straightforward manner, the resulting alogrithms, however, employ extensive matrix operations such as inversion and Cholesky factorization and thus become computationally unattractive specially from real-time application point of view. This problem has been considered in [1] where the given multichannel ARMA process is converted into an equivalent scalar, periodic ARMA process. The corresponding estimation algorithms require scalar operations only and are well-suited for implementation on a pipelined processor. In [2], the scalar representation of the multichannel ARMA process has been used to develop fast adaptive least-squares lattice algorithms for the online identification of multichannel ARMA systems. A similar representation for a multichannel autoregressive (AR) process had earlier been derived by Pagano in [3] and identification algorithms based on this had been discussed in [4], [5].

The performance of these algorithms can be evaluated by computing the Cramer-Rao lower bound (CRLB) on the error covariance matrices of the parameters to be estimated. The error variance associated with any estimation technique that produces unbiased estimates of the parameters can be compared with the CRLB to obtain a measure of the estimation accuracy of that technique. In this correspondence, we present a method to compute the CRLB associated with the estimation of the scalar, periodic ARMA equivalent of a multichannel ARMA process. The derivation is based on the approach followed by Friedlander [6] for a single channel, stationary ARMA process. We show here that the elements of matrix (the CRLB is given by its inverse), for the scalar, periodic ARMA model under consideration, can be obtained by a few very simple operations. These involve constructing certain sequences by delaying and downsampling the scalar input and output sequences, filtering them by a periodic AR filter and cross-correlating the filter outputs. The filter is easily obtainable from the model equation and the same filter is employed for evaluating all the CRLB's.

Manuscript received April 17, 1992; revised July 8, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. John Goutsias.

The authors are with the Department of Electrical Engineering, Indian Institute of Technology, New Delhi, India.

IEEE Log Number 9214186.

II. Derivation of the Cramer-Rao Lower Bound

A. Review of the Scalar, Periodic Representation of a Vector ARMA Process

Consider a d-variate ARMA (p,q) process given by

$$\mathbf{x}(k)+\sum_{i=1}^{p}\mathbf{A}(i)\mathbf{x}(k-i)=\mathbf{z}(k)+\sum_{j=1}^{q}\mathbf{B}(j)\mathbf{z}(k-j) \quad (1)$$

where $\mathbf{A}(i)$'s and $\mathbf{B}(j)$'s are $d\times d$ matrices and $\mathbf{z}(k)$ is a d-variate, zero-mean, white input process with positive definite covariance matrix $\mathbf{Q}$, i.e., $E[\mathbf{z}(k)\mathbf{z}^{\dagger}(k)]=\mathbf{Q}$ (where the superscript '$\dagger$' denotes Hermitian transposition). Then, Cholesky factorization of $\mathbf{Q}$ yields $\mathbf{Q}=\mathbf{LDL}^{\dagger}$ where $\mathbf{L}$ is a unit lower triangular matrix and $\mathbf{D}$ is a diagonal matrix with real, positive diagonal entries. Introducing $\mathbf{u}(k)=\mathbf{L}^{-1}\mathbf{z}(k)$ in (1), we obtain

$$\mathbf{L}^{-1}\mathbf{x}(k)+\sum_{i=1}^{p}\mathbf{L}^{-1}\mathbf{A}(i)\mathbf{x}(k-i) = \mathbf{u}(k)+\sum_{j=1}^{q}\mathbf{L}^{-1}\mathbf{B}(j)\mathbf{L}\mathbf{u}(k-j). \quad (2)$$

Consider two scalar processes $y(n)$ and $e(n)$ related to $\mathbf{x}(k)$ and $\mathbf{u}(k)$, respectively, in the following manner: $y(j+(k-1)d)=x_j(k); e(j+(k-1)d)=u_j(k)$. Replacing the components of $\mathbf{x}(k)$ and $\mathbf{u}(k)$ in (2) by the corresponding elements of $y(n)$ and $e(n)$, respectively, and equating the rows of the L.H.S. and the R.H.S. separately, we observe that $y(n)$ is governed by the following periodically time-varying difference equation

$$y(n)+\sum_{i=1}^{p_n}a_{n,i}y(n-i)=e(n)+\sum_{j=1}^{q_n}b_{n,j}e(n-j) \quad (3)$$

where $a_{n,i}=a_{n\pm d,i}, b_{n,j}=b_{n\pm d,j}, p_n=p_{n\pm d}, q_n=q_{n\pm d}$, and $e(n)$ is a periodically white process, i.e., $E[e(n+m)e^{*}(n)]=\sigma_n^2\,\delta(m), \sigma_n^2=\sigma_{n\pm d}^2$ (The superscript '$*$' denotes complex conjugation and $\delta(m)=1$ if $m=0$ and zero otherwise). The index 'n' is related to the index 'k' in the following manner: $n=j+(k-1)d, j=1,2,\cdots,d$, and the coefficients of the scalar, periodic model (3), i.e., $a_{n,i}$'s and $b_{n,j}$'s are given by the elements of the matrices $\mathbf{L}^{-1}, \mathbf{L}^{-1}\mathbf{A}(i)$'s and $\mathbf{L}^{-1}\mathbf{B}(j)\mathbf{L}$'s. The relationship between the two models (1) and (3) is one-to-one, and one can be obtained from the other in a very simple manner. Further, if one is given to be an innovation model, meaning that it is causal, stable and causally invertible, the other also turns out to be an innovation model [1].

B. CRLB Evaluation

The exact likelihood function of an ARMA data record is known to be a very complicated function of the ARMA parameters and is rarely used in applications like CRLB computation or ML estimation. Instead, an asymptotic likelihood function is used, which approaches the exact likelihood function asymptotically as the data size increases, under the conditions that the poles and zeros of the model do not lie close to the unit circle in the complex z plane. Assume that $\mathbf{z}(k)$ is a real, multivariate Gaussian process and a data record $\mathbf{x}(k),\mathbf{z}(k), k=1,2,\cdots,N$, is given. Then the asymptotic log-likelihood function $H=\ln P(\mathbf{x}(1),\cdots\mathbf{x}(N))$ is given by [7]

$$H=-\frac{1}{2}\sum_{k=1}^{N}\mathbf{z}^{t}(k)\mathbf{Q}^{-1}\mathbf{z}(k)-\frac{N}{2}\ln(\det[\mathbf{Q}]) -\frac{Nd}{2}\ln(2\pi). \quad (4)$$

(The superscript 't' denotes simple transposition and '$\det[\mathbf{Q}]$' denotes the determinant of the matrix $\mathbf{Q}$). The dependence of H in (4) on the ARMA parameters comes through $z(k)$'s i.e., if the linear system (1) is described by the operator T_μ with μ denoting the set of ARMA parameters, then $z(k)$ in (4) is to be treated as $T_\mu^{-1}\mathbf{x}(k)$. Replacing $\mathbf{Q}$ by $\mathbf{LDL}^t$ and introducing $\mathbf{u}(k)=\mathbf{L}^{-1}\mathbf{z}(k)$ in (4), we get

$$H=-\frac{1}{2}\sum_{k=1}^{N}\mathbf{u}^{t}(k)\mathbf{D}^{-1}\mathbf{u}(k)-N\sum_{i=1}^{d}\ln\sigma_i-\frac{Nd}{2}\ln(2\pi), \quad (5)$$

which is the same as

$$H=-\frac{1}{2}\sum_{n=1}^{Nd}\frac{1}{\sigma_n^2}e^2(n)-N\sum_{i=1}^{d}\ln\sigma_i-\frac{Nd}{2}\ln(2\pi). \quad (6)$$

It is easy to verify that (6) gives the asymptotic log-likelihood function corresponding to the scalar, periodic ARMA process $y(n)$ for a data record of length Nd. (This also follows from the equivalence between the two data sets $\{\mathbf{x}(k)|1\le k\le N\}$ and $\{y(n)|1\le n\le Nd\}$ and also from the equivalence between the parameters of (1) and (3).) The parameters of the model (3) to be estimated are given by the set $\{a_{m,r}, b_{m,s}, \sigma_m | m=1,2,\cdots,d, r=1,2,\cdots,p_m, s=1,2,\cdots,q_m\}$.

Let $\theta_i=[a_{i,1},\cdots,a_{i,p_i}]^t, \boldsymbol{\phi}_i=[b_{i,1},\cdots,b_{i,q_i}]^t, i=1,2,\cdots,d$, and $\boldsymbol{\sigma}=[\sigma_1,\cdots,\sigma_d]^t$. Define $\boldsymbol{\Psi}=[\theta_1^t,\cdots,\theta_d^t,\boldsymbol{\phi}_1^t,\cdots,\boldsymbol{\phi}_d^t]^t$. Also, for two vectors $\mathbf{x}=[x_1,\cdots,x_n]^t, \mathbf{y}=[y_1,\cdots,y_m]^t$ and a scalar function $f(\mathbf{x},\mathbf{y})$, denote by $(\partial^2 f/\partial\mathbf{x}\partial\mathbf{y})$ a $n\times m$ matrix such that

$$\left[\frac{\partial^2 f}{\partial\mathbf{x}\partial\mathbf{y}}\right]_{i,j}=\frac{\partial^2 f}{\partial x_i\partial y_j}.$$

Then the Fisher information matrix $\mathbf{J}$ associated with the estimation of $\boldsymbol{\Psi}$ and $\boldsymbol{\sigma}$ is given by

$$\mathbf{J}=-E\begin{bmatrix}\frac{\partial^2 H}{\partial\boldsymbol{\Psi}^2} & \frac{\partial^2 H}{\partial\boldsymbol{\Psi}\partial\boldsymbol{\sigma}}\\ \frac{\partial^2 H}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\Psi}} & \frac{\partial^2 H}{\partial\boldsymbol{\sigma}^2}\end{bmatrix}. \quad (7)$$

Now, for any two parameters $\alpha,\beta\,(\alpha,\beta\neq\sigma_i, i=1,2,\cdots,d)$, we have, from (6),

$$\frac{\partial^2 H}{\partial\beta\,\partial\alpha}=-\sum_{n=1}^{Nd}\frac{1}{\sigma_n^2}\frac{\partial e(n)}{\partial\beta}\frac{\partial e(n)}{\partial\alpha}-\sum_{n=1}^{Nd}\frac{1}{\sigma_n^2}e(n)\frac{\partial^2 e(n)}{\partial\beta\,\partial\alpha} \quad (8)$$

$$\frac{\partial^2 H}{\partial\alpha\,\partial\sigma_i}=\frac{2}{\sigma_i^3}\sum_{n=0}^{N-1}e(i+nd)\frac{\partial e(i+nd)}{\partial\alpha}, \qquad i=1,2,\cdots,d. \quad (9)$$

In addition

$$\frac{\partial^2 H}{\partial\sigma_i^2}=-\frac{3}{\sigma_i^4}\sum_{n=0}^{N-1}e^2(i+nd)+\frac{N}{\sigma_i^2}, \qquad i=1,2,\cdots,d. \quad (10)$$

$$\frac{\partial^2 H}{\partial\sigma_i\partial\sigma_j}=0, \qquad i\neq j, i,j=1,2,\cdots,d. \quad (11)$$

Clearly

$$E\left(\frac{\partial^2 H}{\partial\sigma_i^2}\right)=-\frac{2N}{\sigma_i^2}. \quad (12)$$

To obtain $(\partial e(n)/\partial\alpha)$ and $(\partial e(n)/\partial\beta)$, we differentiate the L.H.S. and the R.H.S. of (3) to get

$$\frac{\partial e(n)}{\partial a_{m,r}}+\sum_{j=1}^{q_n}b_{n,j}\frac{\partial e(n-j)}{\partial a_{m,r}} =\begin{Bmatrix} y(n-r), & \text{if } n=m\pm ld, \text{ for any integer } l\\ 0, & \text{otherwise}\end{Bmatrix} \quad (13)$$

and

$$\frac{\partial e(n)}{\partial b_{m,s}} + \sum_{j=1}^{q_n} b_{n,j} \frac{\partial e(n-j)}{\partial b_{m,s}} = \left\{ \begin{array}{ll} -e(n-s), & \text{if } n = m \pm ld, \text{ for any integer } l \\ 0, & \text{otherwise} \end{array} \right\} \quad (14)$$

Next define a unit sample train $s(n) = \Sigma_{r=-\infty}^{\infty} \delta(n - rd)$. In addition, let $\mathcal{L}$ denote a periodic AR filter with coefficients given by the MA part of (3), i.e., for an input sequence $u(n)$, the output $v(n) = \mathcal{L}(u(n))$ is given by

$$v(n) + \sum_{j=1}^{q_n} b_{n,j} v(n-j) = u(n). \quad (15)$$

The stability of the filter $\mathcal{L}$ comes from the requirement on the model (1) (or, equivalently, on model (3)) to be an innovation model (Note that the innovation model is the only model uniquely identifiable from the second-order statistics of the concerned ARMA process.) Then, from (13) and (14), it follows that $(\partial e(n)/\partial a_{m,r})$ and $(\partial e(n)/\partial b_{m,s})$ can be generated by driving the filter $\mathcal{L}$ with the following sequences: $(z^{-r}y(n))(z^{-m}s(n))$ and $(-z^{-s}e(n))(z^{-m}s(n))$, respectively, where z^{-1} is the unit delay operator. Clearly $(\partial e(n)/\partial a_{m,r})$ and $(\partial e(n)/\partial b_{m,s})$ depend on the past values of $y(n)$ and $e(n)$, respectively, and are therefore orthogonal to $e(n)$. From (8) and (9), we then have

$$E\left(\frac{\partial^2 H}{\partial \beta \, \partial \alpha}\right) = -\sum_{n=1}^{Nd} \frac{1}{\sigma_n^2} E\left(\frac{\partial e(n)}{\partial \beta} \frac{\partial e(n)}{\partial \alpha}\right) \quad (16)$$

$$E\left(\frac{\partial^2 H}{\partial \alpha \, \partial \sigma_i}\right) = 0 \quad (17)$$

where $\alpha, \beta \in \{a_{m,r}, b_{m,s} | m = 1, 2, \cdots, d, r = 1, 2, \cdots, p_m, s = 1, 2, \cdots, q_m\}$. The Fisher information matrix $\mathbf{J}$ can then be written as

$$\mathbf{J} = N \begin{bmatrix} E\left(-\frac{1}{N}\frac{\partial^2 H}{\partial \psi^2}\right) & 0_{I \times d} \\ 0_{d \times I} & 2\mathbf{D}^{-1} \end{bmatrix} \quad (18)$$

where $I = \Sigma_{i=1}^{d} p_i + \Sigma_{j=1}^{d} q_j$ and $\mathbf{D}$ is the diagonal matrix obtained by the Cholesky factorization of $\mathbf{Q}$. A typical element of $E((-1/N)(\partial^2 H/\partial \psi^2))$ is given by

$$E\left(-\frac{1}{N}\frac{\partial^2 H}{\partial \beta \, \partial \alpha}\right) = \sum_{i=1}^{d} \frac{1}{\sigma_i^2} \sum_{n=0}^{N-1} E\left(\frac{\frac{\partial e(i+nd)}{\partial \beta}\frac{\partial e(i+nd)}{\partial \alpha}}{N}\right) \quad (19)$$

which is the same as

$$E\left(-\frac{1}{N}\frac{\partial^2 H}{\partial \beta \, \partial \alpha}\right) = \sum_{i=1}^{d} \frac{1}{\sigma_{n+i-1}^2} E\left(\frac{\partial e(n+i-1)}{\partial \beta} \cdot \frac{\partial e(n+i-1)}{\partial \alpha}\right) \quad (20)$$

i.e., the summation over a period of the cross correlations between the two periodic processes $(\partial e(n)/\partial \alpha)$ and $(\partial e(n)/\partial \beta)$, normalized to the respective variances of $e(n)$. For sufficiently long data, the cross correlations between $(\partial e(n)/\partial \alpha)$ and $(\partial e(n)/\partial \beta)$ can be approximated by computing the sample cross correlations at the filter outputs.

TABLE I
CRAMER-RAO LOWER BOUNDS FOR THE PARAMETERS OF THE SCALAR, PERIODIC ARMA MODEL GIVEN BY (21a) AND (21b) AND THE MEAN SQUARE ESTIMATION ERROR FOR EACH PARAMETER

Parameter	Actual Value of Parameter	CRLB	M.S. Estimation Error
σ_1^2	1.5625	0.0026	0.0911
σ_2^2	0.1225	0.0002	0.0074
$a_{1,1}$	0.75	0.0002	0.0005
$a_{1,2}$	−0.85	0.0001	0.0007
$a_{2,1}$	−0.96	0.0004	0.0464
$a_{2,2}$	−1.27	0.0003	0.0262
$a_{2,3}$	0.166	0.0003	0.3380
$b_{1,1}$	1.889	0.0110	0.0410
$b_{1,2}$	−0.1699	0.0016	0.0022
$b_{2,2}$	0.1699	0.0032	0.1602
$b_{2,3}$	0.3670	0.0002	0.0012

III. SIMULATION RESULTS

The proposed method was simulated on the following scalar, periodic ARMA model with period 2 (i.e., the scalar, periodic equivalent of a 2-channel ARMA model):

$$\begin{aligned} y(n) + 0.75y(n-1) - 0.85y(n-2) &= e(n) + 1.889e(n-1) \\ &\quad - 0.1699e(n-2), \qquad n = 1 + 2l \end{aligned} \quad (21a)$$

$$\begin{aligned} y(n) - 0.96y(n-1) - 1.27y(n-2) + 0.166y(n-3) &= e(n) + 0.1699e(n-2) \\ &\quad + 0.367e(n-3), \qquad n = 2 + 2l \end{aligned} \quad (21b)$$

where $\sigma_{1+2l}^2 = 1.5625, \sigma_{2+2l}^2 = 0.1225$, and l is any integer. Data records of 600 points for $y(n)$ and $e(n)$ were used. The filter (15), constructed with the MA coefficients of (21a)–(21b), was excited by two sequences, obtained by delaying and then downsampling $y(n)$ and $e(n)$ by a factor of 2. For example, for $n = 1 \pm ld$, (i.e., for $m = 1$), $(\partial e(n)/\partial b_{1,2})$ was obtained by exciting the filter with $(-z^{-2}e(n))(z^{-1}s(n))$, i.e., by the sequence $\cdots e(-3), 0, e(-1), 0, e(1), 0, e(3), \cdots$, with $e(-1)$ placed at $n = 1$. The simulation results are presented in Table I, where, using the periodicity of (3), the index n is restricted to $\{1, 2\}$. The table shows the CRLB and the mean square estimation error (MSEE) for each parameter (based on 50 experiments), the estimation being carried out by the Modified Yule Walker method presented in [1].

IV. CONCLUSION

A method to compute the CRLB's for the parameter estimates of the scalar periodic ARMA equivalent of a multivariate ARMA model is presented. The main operation involves periodic AR filtering of certain signals, constructed by delaying and decimating the input and output sequences. While the filter is common for all the parameters, the amount of delay in the input/output sequences and also in the decimator determines the particular parameter whose CRLB is being evaluated.

REFERENCES

[1] M. Chakraborty and S. Prasad, "Multivariate ARMA modeling by scalar algorithms," *IEEE Trans. Signal Processing*, vol. 41, no. 4, pp. 1692–1697, Apr. 1993.

[2] M. Chakraborty and S. Prasad, "Multichannel time-varying ARMA model identification by least squares circular latice structures," *Proc. ICASSP'91* (Toronto, Canada), 1991.

[3] M. Pagano, "Periodic and Multiple Autoregressions," *Ann. Statist.*, vol. 6, pp. 1310–1317, Nov. 1978.

[4] H. Sakai, "Circular lattice filtering using Pagano's method," *IEEE Trans. Acoust. Speech Signal Processing,* vol. ASSP-30, no. 2, pp. 279–287, Apr. 1982.
[5] T. Kawase, H. Sakai, and H. Tokumaru, "Recursive least squares circular lattice and escalator estimation algorithms," *IEEE Trans. Acoust. Speech Signal Processing,* vol. ASSP-31, no. 1, pp. 228–231, Feb. 1983.
[6] B. Friedlander, "On the computation of the Cramer-Rao bound for ARMA parameter estimation," *IEEE Trans. Acoust. Speech Signal Processing,* vol. ASSP-32, no. 4, pp. 721–727, Aug. 1984.
[7] S. M. Kay, *Modern Spectral Estimation: Theory and Application.* Englewood Cliffs, NJ: Prentice-Hall, 1988, pp. 291–293.

Conditions for Positivity of an Energy Operator

Alan C. Bovik and Petros Maragos

***Abstract*—We present necessary and sufficient conditions such that the output from the Teager-Kaiser energy operator** $[\dot{s}(t) = ds(t)/dt]$

$$\Psi_c\,[s(t)] = \dot{s}^2(t) - s(t)\ddot{s}(t) \tag{1}$$

for continuous-time signals $s(t)$ and the output from the corresponding discrete-time energy operator

$$\Psi_d\,[s(n)] = s^2(n) - s(n+1)s(n-1) \tag{2}$$

be non-negative everywhere.

I. Introduction

The nonlinear signal operators in (1) and (2) were developed by Teager [1] in his work on speech modeling and were introduced recently by Kaiser [2], [3]. These operators have been shown to be effective for AM and FM demodulation in several useful classes of signals, such as speech and image signals [4]–[10]. Ψ_c owes its energy-tracking capability to the fact that when it is applied to the output signal from a simple harmonic oscillator, it tracks the energy of the source generating the signal. A more general and particularly useful property of the energy operators (1), (2) are their behavior when applied to AM-FM signals of the form

$$s(t) = a(t)\cos[\phi(t)] \tag{3}$$

in the continuous case, and

$$s(n) = a(n)\cos[\phi(n)] \tag{4}$$

in the discrete case. Here we have

$$\Psi_c[s(t)] \approx a^2(t)\omega_i^2(t) \tag{5}$$

the squared product of the amplitude $a(t)$ and the time-varying instantaneous frequency $\omega_i(t) = \dot{\phi}(t)$. Similarly in the discrete case

$$\Psi_d[s(n)] \approx a^2(n)\sin^2[\Omega_i(n)] \tag{6}$$

where $\Omega_i(n) \equiv d\phi(n)/dn$.

The approximations (5) and (6) hold quite well under very useful conditions expressed in terms of the smoothness or bandlimitedness of the amplitude modulation functions and the instantaneous frequencies. Detailed analyses are presented in [4]–[9], which show that the relative error is quite small for realistic signals in speech and other communications applications. These observations have led to the development of *energy separation algorithms* [5], [7], which attempt to separate the amplitude and frequency modulations in the products (5), (6) as distinct useful pieces of information.

The positivity of the energy operator output is a desired property for at least three fundamental reasons: (1) the interpretation of the output as some (normalized) physical energy; (2) the positive nature of the approximations in (5) and (6) needed for AM and/or FM demodulation [4], [6]; (3) the fact that the energy separation algorithms [5], [7] operate under this positivity assumption. In [6], [7] several sufficient conditions have been developed for the positivity of the energy operators. For example, $\Psi_c[s(t)] \geq 0$ if the signal $s(t)$ is any finite product of cosines, real exponentials, and linear trends. Alternatively, $\Psi_c[s(t)$ 0 if $s(t)$ is an AM-FM signal and the amounts of amplitude/frequency modulation are not excessively large and the bandwidths of the amplitude/frequency modulating signals are reasonably smaller than the carrier frequency.

In the current paper, we explore general conditions on the arbitrary signal s such that $\Psi_c(s)$ or $\Psi_d(s)$ be non-negative over the entire domain of analysis. Some of these conditions are necessary and sufficient and have an interesting geometric meaning, since they are expressed in terms of the concavity of the logarithm of the signal magnitude. We treat the continuous and discrete cases separately.

II. Continuous Case

Henceforth, we suppose that the signal $s\colon \mathbf{D} \to \mathbf{R}$ has finite second derivatives everywhere (and hence is continuous) on some arbitrary set $\mathbf{D} \subseteq \mathbf{R}$. Lemma 1 is a simple, easily-tested sufficient condition for non-negativity of $\Psi_c(s)$. Lemma 2 is used to prove Theorem 1, although it also supplies interesting conditions for the non-negativity of the operator Ψ_c for the special case of nonzero signals.

Lemma 1: At any $t \in \mathbf{D}, \Psi_c[s(t)] \geq 0$ if any of the following conditions hold:

(a) $s(t) = 0$. (b) $\ddot{s}(t) = 0$. (c) $s(t) > 0$ and $\ddot{s}(t) < 0$. (d) $s(t) < 0$ and $\ddot{s}(t) > 0$.

Proof: If any of (a)–(d) is true, then $s(t)\ddot{s}(t) \leq 0 \Rightarrow \Psi_c[s(t)] \geq 0$. ♦

Thus, if $\mathbf{I} \subseteq \mathbf{D}$ is some interval whose endpoints are either zeroes or inflection points of $s(t)$, then $\Psi_c[s] \geq 0 \forall t \in \mathbf{I}$ if either s is positive and concave or if s is negative and convex in the interior of $\mathbf{I}$. The next lemma gives necessary and sufficient conditions for everywhere nonzero signals, which are only slightly more complicated.

Lemma 2: Suppose that $s(t) \neq 0 \forall t \in \mathbf{D}$. Then the following three statements are equivalent:

(a) $\Psi_c[s(t)] \geq 0 \forall t \in \mathbf{D}$. (b) $\log[s^2(t)]$ is concave on $\mathbf{D}$. (c) $\log|s(t)|$ is concave on $\mathbf{D}$.

Proof: Let $g(t) = \log[s^2(t)]$. Then (a) $\Leftrightarrow$ (b) since $g(t)$ is concave on $\mathbf{D} \Leftrightarrow \ddot{g} \leq 0 \forall t \in \mathbf{D} \Leftrightarrow 2(s(t)\ddot{s}(t) - \dot{s}^2(t)/s^2(t)) \leq$

Manuscript received May 31, 1992; revised March 28, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. Douglas Williams. A. C. Bovik was supported by a University of Texas faculty research assignment while on sabbatical in the Division of Applied Sciences, Harvard University, Cambridge, MA, and by a grant from Texas Instruments. P. Maragos was supported by the National Science Foundation under Grant NSF MIP-9120624 and by a NSF Presidential Young Investigator Award under NSF Grant MIP-86-58150 with matching funds from Xerox.

A. C. Bovik is with the Laboratory for Vision Systems, University of Texas, Austin, TX 78712-1084.
P. Maragos is with the Division of Applied Sciences, Harvard University, Cambridge, MA 02138.

IEEE Log Number 9214171.