Noise Reduction for an Estimated Wiener Filter Using Noise References

P. COMON AND J. L. LACOUME

Abstract—**A simple approach to the adaptation of a linear filter to cope with spectral uncertainties is presented. Based on the calculation of the total output error power, including the errors of spectral estimation, a criterion is proposed for deciding whether the estimated filter is effective at any given frequency. By forcing the filter response to unity at all ineffective frequencies, an improvement in performance is obtained.**

I. Introduction

Classical Wiener filtering assumes that power spectral densities (PSD's) are known. In practice, estimates of PSD matrices must be used, leading to suboptimal performance. The purpose of this correspondence, is to show how to reduce the output noise caused by the estimation errors. We study the case of a Wiener filter; a subset of the filter's inputs called the noise references is assumed to contain noise only, correlated with the noises disturbing the signal components. Furthermore, signal and noises are assumed to be stationary Gaussian processes with zero mean, with signal independent of the noises. The estimate of the filter is calculated from one observation record of the discrete-time inputs and then is applied to observation records, independent of the previous one.

II. Formulation of the Problem

A. Filter Structure

We assume throughout this correspondence that we observe realizations of the following. 1) n input data "noisy signals"

$$u(m) \triangleq \begin{bmatrix} u_1(m) \\ u_n(m) \end{bmatrix},$$

where $u_i(m) = s_i(m) + x_i(m)$. The processes s_i and x_i represent signal and additive noise, respectively. 2) p input data called "noise references"

$$y(m) \triangleq \begin{bmatrix} y_1(m) \\ y_p(m) \end{bmatrix}.$$

These inputs y_j are supposed to represent random components correlated with the x_i additive noises. All the data are finite length discrete-time sequences, defined for $1 \le m \le N$.

The data in the spectral domain will be obtained by discrete Fourier transform (DFT) and will be denoted by capital letters. For example $U(K)$ is the DFT of $u(m)$. A^T, A^*, and A^+ will denote, respectively, the transposed, the complex conjugate, and transposed conjugate matrix of A.

The theoretical linear filter giving an estimate $\hat{S}$ of the signal is completely defined by gain matrices G and H so that

$$\hat{S}(k) = H(k) \cdot U(k) - G(k) \cdot Y(k).$$

Taking into account the hypotheses mentioned in the introduction, it is well-known that the least mean square error estimate of the signal reduces to

$$\hat{S}(k) = U(k) - G(k) \cdot Y(k)$$
$$G(k) = N(k) \cdot D^{-1}(k). \tag{1}$$

$N(k)$ and $D(k)$ will be called the numerator and denominator, respectively, of the matrix gain $G(k)$ and are defined by

$$N(k) \triangleq E[U(k) \cdot Y^+(k)] = E[X(k) \cdot Y^+(k)]$$
$$D(k) \triangleq E[Y(k) \cdot Y^+(k)]. \tag{2}$$

Note that the filter defined by (1) and (2) is linear with respect to the variable $U(k)$, and no generality is lost if we assume that $n = 1$. In the following, the variables $U(k)$, $X(k)$, and $S(k)$ will be considered as scalar variables.

B. Estimate of the Gain Matrix

Suppose that the cross-spectrum matrices used are, in fact, estimates $\hat{N}(k)$ and $\hat{D}(k)$ instead of exact values $N(k)$ and $D(k)$, so that one obtains only a suboptimal solution $\hat{G}(k) = \hat{N}(k) \cdot \hat{D}^{-1}(k)$ (Fig. 1).

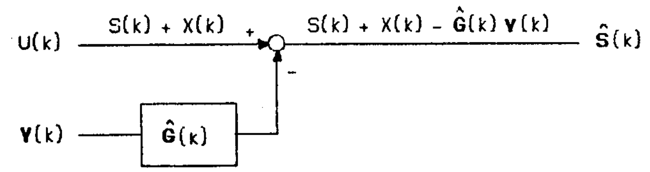

Fig. 1 Filtering diagram.

Here the study is limited to the case of spectral estimation by the smoothed periodogram method [1]–[3]:

$$\hat{N}(k) = \sum_{n=0}^{N-1} U(n) \cdot Y^+(n) \cdot F_k(n)$$

$$\hat{D}(k) = \sum_{n=0}^{N-1} Y(n) \cdot Y^+(n) \cdot F_k(n), \tag{3}$$

where the spectral window $F_0(n)$ is real, even, positive, and periodic of period N, normalized by the equation

$$\sum_{n=0}^{N-1} F_0(n) = 1. \tag{4}$$

The function $F_k(n)$ is centered at the frequency $n = k$, given by $F_k(n) = F_0(n - k)$.

C. Calculation Hypotheses

1) Hypothesis of Periodicity: All the discrete-time or discrete-frequency functions are assumed to be periodic, of period N. This hypothesis is usual when applying the DFT to finite-length data.

2) Gaussian Hypothesis: The processes $X(k)$ and $Y(k)$ are obtained by DFT of real Gaussian processes $x(m)$ and $y(m)$, with zero mean and period N. So, $X(k)$ and $Y(k)$ are complex normally distributed, as introduced by Wooding [4], with a zero mean; the variances of $X(k)$ and $Y(k)$ will be denoted, respectively, by $P_X(k)$ and $D(k)$.

3) Local Whiteness Hypothesis: The spectral functions $P_X(n)$, $N(n)$, and $D(n)$ are assumed to be approximately constant over the support of $F_k(n)$. Thus, if $Q(n)$ is a function depending on

Manuscript received September 18, 1984; revised March 22, 1985.

P. Comon is with the Department of Aerospace, CROUZET, 26027 Valence Cedex, France.

J. L. Lacoume is with the Centre d'Etude des Phénomènes Aléatoires et Géophysiques, Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble, BP 46, 38402 Saint-Martin-d'Hères-Cedex, France.

IEEE Log Number 8406632.

0018-9448/86/0300-0310$01.00 ©1986 IEEE

spectral quantities, we get the following results:

$$\sum_{n=0}^{N-1} F_k^2(n) \cdot Q(n) \cong Q(k) \cdot \sum_{n=0}^{N-1} F_k^2(n)$$

$$\sum_{n=0}^{N-1} F_k(n) \cdot F_{N-k}(n) \cdot Q(n) \cong \mathrm{Re}[Q(k)] \cdot \sum_{n=0}^{N-1} F_k(n) \cdot F_{N-k}(n). \quad (5)$$

Equation (5) is a result of the property $Q(N-k) = Q^*(k)$.

D. Statistical Properties of the Gain

The estimates $\hat{N}(k)$ and $\hat{D}(k)$, defined by (3), are matrix-valued random variables. From (5), it may be seen that their mean values are approximately $N(k)$ and $D(k)$. Consequently, $\hat{N}$ and $\hat{D}$ are approximately unbiased estimates.

More precisely, $\hat{N}$ and $\hat{D}$ are complex Wishart random variables as defined in Brillinger [1] and are not independent in the general case. Therefore, the statistical design of $\hat{G}$ requires complicated calculations.

A simple suboptimal design is obtained by linearizing the estimates about their theoretical optimal values (the statistical mean). Let the approximately zero-mean variables be defined as

$$\delta N \triangleq \hat{N} - N \qquad \delta D \triangleq \hat{D} - D,$$

and let $\delta G \triangleq \hat{G} - G$. This yields, to first order,

$$\delta G \cong \delta N \cdot D^{-1} - N \cdot D^{-1} \cdot \delta D \cdot D^{-1}. \quad (6)$$

Hence the variable δG is zero mean to first order.

III. Error Power

The disturbing noise $X(k)$ can be split into a noise correlated with the references $G(k) \cdot Y(k)$ and an independent noise $Z(k)$, giving

$$U(k) = S(k) + X(k) = S(k) + Z(k) + G(k) \cdot Y(k)$$
$$\hat{S}(k) \triangleq S(k) + \epsilon(k) = S(k) + Z(k) - \delta G(k) \cdot Y(k).$$

With these notations, the output residual noise reduces to

$$\epsilon(k) = Z(k) - \delta G(k) \cdot Y(k).$$

$Z(k)$ represents the theoretical minimum error, and $\delta G(k) \cdot Y(k)$ the extra noise due to estimation errors in the filter.

As stated in the Introduction, we assume that $\hat{G}$ is calculated from an observation record of inputs independent of those to which it is applied. So, the realizations $S(k)$, $X(k)$, $Z(k)$, and $Y(k)$ are independent of $\delta G(k)$ in the foregoing expressions. Then, the power of the residual noise can be expressed in terms of the input noise power as follows:

$$P_\epsilon(k) = P_X(k) - G(k) \cdot D(k) \cdot G^+(k) + E[\delta G(k) \cdot D(k) \cdot \delta G^+(k)]. \quad (7)$$

The last term of the right side represents the decrease of performance resulting from the estimation errors. Let

$$L(k) \triangleq \sum_{n=0}^{N-1} F_k(n) \cdot F_k(N-n). \quad (8)$$

The shape of the function $L(n)$ is presented in Fig. 2. In this example, $F_0(n)$ is the DFT of a Hann window $f(m)$:

$$f(m) \triangleq \left(\frac{1}{2} + \frac{1}{2}\cos\frac{2\pi m}{N}\right)^8.$$

From the hypotheses explained in Section II-C and taking into account the result of the Appendix, (7) yields

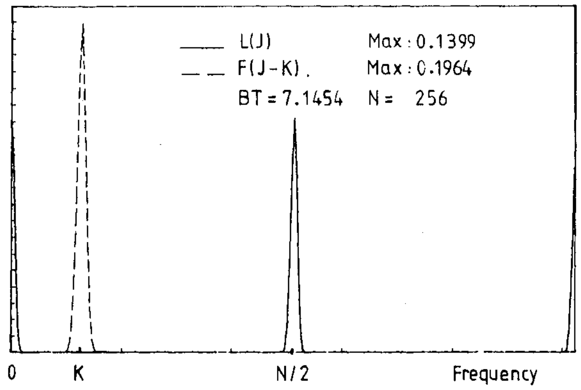

Fig. 2.

$$P_\epsilon(k) \cong P_X(k) + p \cdot L(0) \cdot [P_U(k) - N(k) \cdot D^{-1}(k) \cdot N^+(k)] - G(k) \cdot D(k) \cdot G^+(k)$$

for $k \notin \sup(L)$; $\sup(L)$ denotes the support of L. From (1), the right side simplifies to

$$P_\epsilon(k) \cong P_X(k) + p \cdot L(0) \cdot P_U(k) - (1 + p \cdot L(0)) \cdot N(k) \cdot D^{-1}(k) \cdot N^+(k) \quad (9)$$

for $k \notin \sup(L)$.

IV. Performance Criterion

A. Definition of the Criterion

We consider the filtering operation as effective for frequency k if the output noise PSD is lower than the input noise PSD: $P_\epsilon(k) < P_X(k)$. That is, from (9)

$$N(k) \cdot D^{-1}(k) \cdot N^+(k)/P_U(k) > p \cdot L(0)/[1 + p \cdot L(0)].$$

The left side represents the multiple coherence function $|C_{UY}(k)|^2$ [2] between the inputs $U(k)$ and $Y(k)$.

Furthermore, from (4) and (8), the value of the bandwidth duration (BT) product η of the spectral window $F_0(k)$ is $1/L(0)$. Hence another simpler form of the criterion appears:

$$|C_{UY}(k)|^2 > \frac{1}{1 + \eta/p}; \qquad k \notin \sup(L). \quad (10)$$

We could obtain the general case ($n > 1$) using the criterion

$$|C_{U_iY}(k)|^2 > \frac{1}{1 + \eta/p}; \qquad k \notin \sup(L);\ 1 \le i \le n.$$

B. Utilization of the Criterion

Let $\hat{C}_{UY}$ be an estimate of the multiple coherence function C_{UY}. Let $[C_{UY} - \Delta_{\inf}, C_{UY} + \Delta_{\sup}]$ be a confidence interval for $\hat{C}_{UY}$. In practice, in order to take into account the estimation errors in the multiple coherence function, (10) can be modified to

$$|\hat{C}_{UY}(k)| > \Delta_{\sup}(k) + \frac{1}{\sqrt{1 + \eta/p}}. \quad (11)$$

The upper bound $\Delta_{\sup}(k)$ of the confidence interval can be calculated by assuming that the random variable $\tanh^{-1} \hat{C}_{UY}(k)$ is normal [1], [2].

To improve the performance of the filtering method, the filter defined by (1) and (3) should not be employed for the frequencies that do not satisfy (11). For these frequencies, we set the output $S(k)$ equal to the input $U(k)$. We call this processing the modified filtering.

Furthermore, (11) can be used for any frequency k, $0 \leq k \leq N-1$, even if $k \in \sup(L)$, as long as the BT product η of the smoothing window is not too large. In that case, the support of $L(k)$ reduces to two narrow bands around the frequencies $k = 0$ (band of the trends) and $k = N/2$ (Nyquist's frequency) as shown in Fig. 2. In fact, spectral power densities often vanish around these frequencies if the inputs are correctly prefiltered (trends and high frequencies have to be removed [2]). On the other hand, for large value of η, an exact criterion must be determined according to Theorem B stated in the Appendix.

V. Example

For the sake of simplicity, we present simulation results obtained in the scalar case. The general case ($n > 1$, $p > 1$) leads to similar results for each output $S_i(k)$, but n thresholds $t_i(k)$ must be computed. In this example, we have computed two independent white zero-mean normal random processes $r_i(m)$, and the data have been simulated as follows:

$$U(k) = S(k) + X(k)$$
$$Y(k) = R_1(k),$$

where

$$S(k) = 0$$
$$X(k) = R_1(k) + P(k) \cdot R_2(k)$$

and $P(k)$ is a real positive unimodal function.

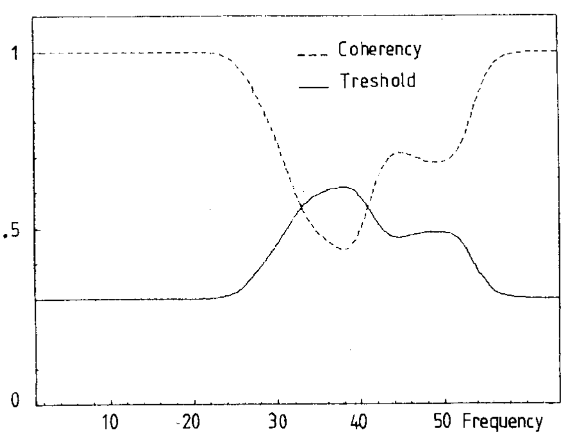

Fig. 3.

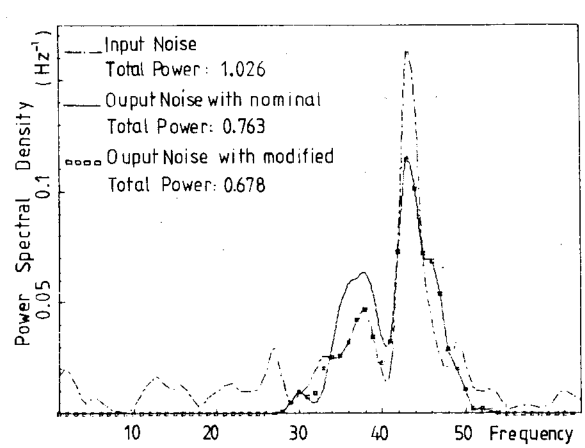

Fig. 4.

From (11), choosing a 90 percent confidence interval, the following threshold is defined (cf. Fig. 3):

$$|\hat{C}(k)| > t(k)$$

$$t(k) = \frac{1}{\sqrt{1+\eta}} + \frac{\left[1 - |\hat{C}(k)|^2\right] \tanh\left(z/\sqrt{\eta}\right)}{1 + |\hat{C}(k)| \tanh\left(z/\sqrt{\eta}\right)}; \qquad z = 1.65.$$

The PSD of $(S - \hat{S})$ is plotted in Fig. 4, for both the nominal and the modified filter, with $\eta = 10.0$. From this figure, we observe that the modified filter output PSD coincides with the nominal in the effective ranges, and with the input noise PSD in the ineffective range. In this example, the mean power of the output noise is reduced by 11 percent when the modified filter is employed instead of the nominal.

VI. Conclusion

This work has been carried out within the framework of adaptive filtering using noise references. The aim of the study was to determine a simple performance criterion in the spectral domain, in order to reduce the contribution of the estimation errors in the residual output noise. Based on reasonable assumptions described in Section II, we have defined a criterion in Section IV that can easily be applied in experimental data processing because of its simplicity. This criterion allows the construction of a modified filter that performs better than the nominal filter.

Appendix

Theorem A

The following theorem is a direct result of the properties of the complex normal random variables.

Let $\boldsymbol{H}$ be an Hermitian matrix. Let $\boldsymbol{A}(n)$, $\boldsymbol{B}(n)$, $\boldsymbol{C}(n)$, $\boldsymbol{D}(n)$ be complex vector-valued random processes, periodic with period N, obtained by taking the DFT of real zero-mean normal random processes (they are therefore complex normal random variables for each n, see [1] and [4]).

Let $\boldsymbol{S}(m, n)$ be the deterministic periodic function with period N:

$$\boldsymbol{S}(m,n) = \boldsymbol{E}[\boldsymbol{A}(n) \cdot \boldsymbol{B}^+(n) \cdot \boldsymbol{H} \cdot \boldsymbol{C}(m) \cdot \boldsymbol{D}^+(m)].$$

Then, expressing the second-order moments, $\boldsymbol{S}(m, n)$ reduces to

$$\begin{aligned} \boldsymbol{S}(m,n) = {} & \boldsymbol{E}[\boldsymbol{A}(n) \cdot \boldsymbol{B}^+(n)] \cdot \boldsymbol{H} \cdot \boldsymbol{E}[\boldsymbol{C}(m) \cdot \boldsymbol{D}^+(m)] \\ & + \boldsymbol{E}[\boldsymbol{A}(n) \cdot \boldsymbol{D}^+(m)] \cdot \boldsymbol{E}[\boldsymbol{B}^+(n) \cdot \boldsymbol{H} \cdot \boldsymbol{C}(m)] \\ & \cdot \delta(m-n) + \boldsymbol{E}[\boldsymbol{A}(n) \cdot \boldsymbol{C}^+(n)] \cdot \boldsymbol{H}^T \\ & \cdot \boldsymbol{E}[\boldsymbol{B}^*(n) \cdot \boldsymbol{D}^T(n)] \cdot \delta(m+n-N). \end{aligned}$$

Theorem B

$$\begin{aligned} \boldsymbol{E}[\delta\boldsymbol{G}(k) \cdot \boldsymbol{D}(k) \cdot \delta\boldsymbol{G}^+(k)] & \\ \cong p \cdot L(0)\left[\boldsymbol{P}_U(k) - \boldsymbol{N}(k) \cdot \boldsymbol{D}^{-1}(k) \cdot \boldsymbol{N}^+(k)\right] & \\ + L(k)\left[\mathrm{Re}\left[\boldsymbol{N}(k) \cdot \boldsymbol{D}^{-1}(k) \cdot \boldsymbol{N}^T(k)\right]\right. & \\ \left. - \boldsymbol{N}(k) \cdot \boldsymbol{D}^{-1}(k) \cdot \boldsymbol{N}^+(k)\right]. & \end{aligned}$$

Proof: Dropping index k for simplicity, the linear approximation (6) yields

$$\begin{aligned} \boldsymbol{E}[\delta\boldsymbol{G} \cdot \boldsymbol{D} \cdot \delta\boldsymbol{G}^+] & \\ \cong \boldsymbol{E}[\delta\boldsymbol{N} \cdot \boldsymbol{D}^{-1} \cdot \delta\boldsymbol{N}^+] + \boldsymbol{G} \cdot \boldsymbol{E}[\delta\boldsymbol{D} \cdot \boldsymbol{D}^{-1} \cdot \delta\boldsymbol{D}^+] \cdot \boldsymbol{G}^+ & \\ - \boldsymbol{G} \cdot \boldsymbol{E}[\delta\boldsymbol{D} \cdot \boldsymbol{D}^{-1} \cdot \delta\boldsymbol{N}^+] - \boldsymbol{E}[\delta\boldsymbol{N} \cdot \boldsymbol{D}^{-1} \cdot \delta\boldsymbol{D}^+] \cdot \boldsymbol{G}^+. & \end{aligned}$$

From definitions (3) and (8), result (5), and Theorem A, every term of the right side can be calculated.

a) Numerator–numerator moment:

$$\begin{aligned} &\boldsymbol{E}\left[\delta\boldsymbol{N}(k)\cdot\boldsymbol{D}^{-1}(k)\cdot\delta\boldsymbol{N}^{+}(k)\right] \\ &= \sum_{n=0}^{N-1} F_k^2(n)\cdot\boldsymbol{E}\left[U(n)U^{*}(n)\right] \\ &\quad\cdot\boldsymbol{E}\left[\boldsymbol{Y}^{+}(n)\cdot\boldsymbol{D}^{-1}(k)\cdot\boldsymbol{Y}(n)\right] \\ &\quad+\sum_{n=0}^{N-1} F_k(n)\cdot F_k(N-n)\cdot\boldsymbol{E}\left[U(n)\cdot\boldsymbol{Y}^{+}(n)\right] \\ &\quad\cdot\boldsymbol{D}^{-1}(k)\cdot\boldsymbol{E}\left[\boldsymbol{Y}^{*}(n)\cdot U(n)\right] \\ &\cong P_U(k)\cdot p\cdot L(0) \\ &\quad+\operatorname{Re}\left[\boldsymbol{N}(k)\cdot\boldsymbol{D}^{-1}(k)\cdot\boldsymbol{N}^{T}(k)\right]\cdot L(k). \end{aligned}$$

b) Denominator–denominator moment:

$$\begin{aligned} &\boldsymbol{E}\left[\delta\boldsymbol{D}(k)\cdot\boldsymbol{D}^{-1}(k)\cdot\delta\boldsymbol{D}^{+}(k)\right] \\ &= \sum_{n=0}^{N-1} F_k^2(n)\cdot\boldsymbol{E}\left[\boldsymbol{Y}(n)\boldsymbol{Y}^{+}(n)\right] \\ &\quad\cdot\boldsymbol{E}\left[\boldsymbol{Y}^{+}(n)\cdot\boldsymbol{D}^{-1}(k)\cdot\boldsymbol{Y}(n)\right] \\ &\quad+\sum_{n=0}^{N-1} F_k(n)\cdot F_k(N-n)\cdot\boldsymbol{E}\left[\boldsymbol{Y}(n)\cdot\boldsymbol{Y}^{+}(n)\right] \\ &\quad\cdot\boldsymbol{D}^{-T}(k)\cdot\boldsymbol{E}\left[\boldsymbol{Y}^{*}(n)\cdot\boldsymbol{Y}^{T}(n)\right] \\ &\cong \boldsymbol{D}(k)\left[p\cdot L(0)+L(k)\right]. \end{aligned}$$

c) Numerator–denominator moment:

$$\begin{aligned} &\boldsymbol{E}\left[\delta\boldsymbol{N}(k)\cdot\boldsymbol{D}^{-1}(k)\cdot\delta\boldsymbol{D}^{+}(k)\right] \\ &= \sum_{n=0}^{N-1} F_k^2(n)\cdot\boldsymbol{E}\left[U(n)\boldsymbol{Y}^{+}(n)\right] \\ &\quad\cdot\boldsymbol{E}\left[\boldsymbol{Y}^{+}(n)\cdot\boldsymbol{D}^{-1}(k)\cdot\boldsymbol{Y}(n)\right] \\ &\quad+\sum_{n=0}^{N-1} F_k(n)\cdot F_k(N-n)\cdot\boldsymbol{E}\left[U(n)\cdot\boldsymbol{Y}^{+}(n)\right] \\ &\quad\cdot\boldsymbol{D}^{-1}(k)\cdot\boldsymbol{E}\left[\boldsymbol{Y}^{*}(n)\cdot\boldsymbol{Y}^{T}(n)\right] \\ &\cong \boldsymbol{N}(k)\left[p\cdot L(0)+L(k)\right]. \end{aligned}$$

The theorem follows by using these expressions.

References

[1] D. R. Brillinger, *Time Series, Data Analysis and Theory*. San Francisco, CA: Holden-Day, 1981.

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

[3] A. H. Nuttal and C. G. Carter, "A generalized framework for power spectral estimation," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-28, pp. 334–335, June 1980.

[4] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, pp. 212–215, 1956.

Spectral Analysis of *Q*-ary Digital Signals Encoded by *P*-ary Convolutional Codes

JOÃO B. BEZERRA, MEMBER, IEEE, AND DALTON S. ARANTES

Abstract—A method for spectral calculation of q-ary digital signals encoded by p-ary convolutional codes is developed. We have assumed a parallel encoding process such that for a code with rate k/n and constraint length m, the encoder can be modeled as a Mealy machine of k inputs, n outputs, and $(m-1)k$ q-ary memory cells. From this model and from the generator matrix of the code, the characterizing matrices of the Mealy machine are found in a straightforward way. Two assumptions are made on the statistics of the input signal. 1) The input sequence is a wide-sense stationary vector process of dimension k. 2) The input vectors are statistically independent. With these assumptions and from the characterizing matrices of the Mealy machine, the method developed by Cariolaro and Tronca can be readily used. The continuous and discrete components of the spectrum of the coded signal are expressed in matrix form, which is particularly appropriate for computer programming.

I. Introduction

Until recently, the power spectral density of binary signals encoded by convolutional codes has been assumed to be a scaled version of the spectrum before encoding. In a recent paper, Divsalar and Simon [2] gave a theoretical account of a class of convolutional codes, called correlated codes, for which this assumption is not true. In fact, these codes have been shown to exhibit some interesting properties of bandwidth compression.

The object of the present correspondence is to present an extension of the results given in [1] for block codes, so as to include the case of multilevel convolutional codes. We consider the broad class of convolutional codes with q-ary input symbols and p-ary output symbols, where $q \leqq p$ and all symbols are elements of the Galois field GF(p).

As in [3], we assume a parallel encoder that can be modeled as a Mealy machine whose matrices are obtained from the generator matrix of the code with p-ary elements. We also assume that the input signal is a wide-sense stationary vector process with statistically independent vectors. The work is divided as follows. In Section II we present the mathematical tools that model the encoder as a Mealy machine. In Section III we present a short review of the results developed in [1], and finally in Section IV we give examples of spectral calculations for binary and ternary convolutional codes, both with binary input symbols.

II. Modeling the Encoder

Let the generator matrix of a p-ary convolutional code of rate k/n and constraint length m be given by

$$\boldsymbol{G} = \left[\boldsymbol{G}_0 \quad \boldsymbol{G}_1 \quad \cdots \quad \boldsymbol{G}_{m-1}\right] \tag{1}$$

where $\boldsymbol{G}_l$ $(l = 0,1,\cdots,m-1)$ are $n \times k$ submatrices with elements $g_l(i,j)$ from the Galois field of p elements, GF(p).

A parallel encoder is assumed with $M = k(m-1)$ q-ary memory cells [4, pp. 392–394]. At the rth time slot, the encoder accepts k q-ary numbers $x_i^r \in \mathrm{GF}(q)$, $q \leqq p$, $i = 1,2,\cdots,k$,

Manuscript received April 12, 1985; revised July 25, 1985. This work was supported in part by CNPq-Conselho Nacional de Desenvolvimento Científico e Tecnológico and CPqD, Telebrás, Brazil.

J. B. Bezerra is with the Electrical Engineering Department, UFRN, Campus Universitário, 59000-Natal-RN-Brazil.

D. S. Arantes is with the Electrical Engineering Department, UNICAMP, 13100-Campinas-SP-Brazil.

IEEE Log Number 8406648.

0018-9448/86/0300-0313$01.00 ©1986 IEEE