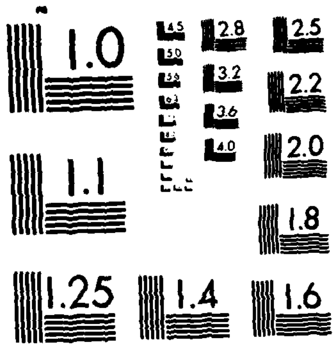
AD-A163 054 TIME DOMAIN ALGORITHMS(U) SYSTEMS CONTROL TECHNOLOGY INC PALO ALTO CA B FRIEDLANDER ET AL SEP 85 PA-6502-01 RADC-TR-85-167 F30602-84-C-0016 1/2

UNCLASSIFIED F/G 17/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A163 054

RADC-TR-85-167
Final Technical Report
September 1985

TIME DOMAIN ALGORITHMS

Systems Control Technology, Inc.

Benjamin Friedlander and J. O. Smith

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-85-167 has been reviewed and is approved for publication.

APPROVED:

RICHARD N. SMITH
Project Engineer

APPROVED: B Beek

BRUNO BEEK
Technical Director
Communications Division

FOR THE COMMANDER: Richard W. Pouliot

RICHARD W. POULIOT
Plans Office

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS N/A
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 6502-01	5. MONITORING ORGANIZATION REPORT NUMBER(S) RADC-TR-85-167
6a. NAME OF PERFORMING ORGANIZATION Systems Control Technology, Inc 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Rome Air Development Center (DCCD)
6c. ADDRESS (City, State, and ZIP Code) 1801 Page Mill Road PO Box 10180 Palo Alto CA 94303-0888	7b. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Rome Air Development Center 8b. OFFICE SYMBOL (If applicable) DCCD	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F30602-84-C-0016
8c. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2305; TASK NO. J8; WORK UNIT ACCESSION NO. 11

AD-A163 054

PA-

11. TITLE (Include Security Classification)
TIME DOMAIN ALGORITHMS

12. PERSONAL AUTHOR(S)
Benjamin Friedlander, J. O. Smith

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Apr 84 TO Apr 85	September 1985	138

16. SUPPLEMENTARY NOTATION
N/A

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	
09	04		Time-Varying Filters; Adaptive Processing
17	02		Constant Modulus Algorithm; Non-Stationary Signals

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The report summarizes research results aimed at addressing signal processing issues related to the non-stationary time-varying nature of signals propagating through communication channels. The Constant Modulus Algorithm is studied and extended and a convergence proof is provided. A class of time-varying waveguide digital filters is derived and its properties analyzed. The application of time-varying autoregressive models to modeling a class of non-stationary signals is presented.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Richard N. Smith	22b. TELEPHONE (Include Area Code) (315) 330-3224 22c. OFFICE SYMBOL RADC (DCCD)

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

TABLE OF CONTENTS

1. INTRODUCTION AND SUMMARY

This report summarizes the work performed on the "Time Domain Algorithms" project, under contract No. F30602-84-C-0016. The objective of this research was to address some of the issues related to the non-stationary, time-varying nature of signals propagating through a communications channel.

Many communications signals (such as frequency modulation and phase modulation formats) are characterized by a constant modulus property. The propagation of these signals through the communication channel introduces various types of "distortions" due to multipath, fading and similar effects. These distortions tend to destroy the constant modulus properties of the signal. Recently, an adaptive processing technique was developed which has the capability of correcting the effects of the channel and restoring the constant modulus property of the signal. This technique has been shown to be very useful in communications applications. We have analyzed this important processing technique and developed some useful extensions.

In Appendix A we prove global convergence of the Constant-Modulus Algorithm (CMA) for the case of a real channel when the model order is equal to or greater than that of the channel (the so-called "model-complete" case). The analysis is based on an exact fourth-order Taylor series representation for the cost function minimized asymptotically by the CMA.

In Appendix B we present several extensions to the CMA including IIR equalization, a real-signal version having properties as good as the complex version, use of the Gauss-Newton method in place of gradient descent, interference rejection, and more. Some preliminary simulation results are presented in Appendix F.

The processing and estimation of signals propagating through time-varying channels requires time-varying filter structures. While the area of time-invariant digital filters is well developed, relatively little fundamental work is available on the time-varying case. We investigated some of the basic questions related to time-varying digital filtering and were able to derive a novel filter structure for such applications.

In Appendix C we derive the set of finite-order, linear, time-invariant filters by sampling lossless propagation through a variable-impedance medium. This leads to a flexible class of time-varying filter structures, termed "Waveguide Filters" (WGF) in which signal power is decoupled from changes in the filter parameters. These structures are "balanced" in the sense that the decoupling between signal power and time-varying filter coefficients is maintained for each individual section in the structure. In addition, limit cycles and overflow oscillations are suppressed, even in the time-varying case, when implemented with "passive" arithmetic. Finally, the WGF structures can be interconnected in series or in parallel in a way which does not disturb the signal/coefficient decoupling or the power balance. Thus,the waveguide filters are very useful for modeling physical systems, and the exactness of their physical interpretation enhances their suitability for the time-varying case. All results are obtained for the multi-input/multi-output case.

Another topic which was briefly investigated during this study was the adaptive equalization of rapidly time-varying multipath channels. Due to time limitations only a conceptual study of this topic was possible. In Appendix D we describe on adaptive equalizer for eliminating distortion due to multipath propagation in high-speed digital radio systems. The equalizer is aimed at the case of very fast channel fluctuation, where "fast" is defined relative to the impulse-response duration of the inverse of the instantaneous multipath transfer function. The method is of the decision-feedback type where the demodulated symbols are used to construct an estimate of multipath-induced intersymbol interference. The reconstructed baseband waveform is then delayed and weighted according to the current multipath parameters, and this simulated echo is subtracted from the incoming baseband waveform. The distinguishing feature of our approach is that an explicit model of multipath parameters replaces the tranvsersal equalizer studied previously.

An important counterpart of the filter design prolem is the problelm of modeling non-stationary signals. In Appendix E we consider the problem of estimating sinusoidal or narrowband signals with a time-varying center frequency. The signal parameters are estimated by fitting an autoregressive model with time-varying coefficients to the data. The overdetermined modified

Yule-Walker equations are used to estimate a set of constant model parameters. Some numerical examples illustrating the behavior of the estimator are presented, and its accuracy aspects are briefly discussed. This particular study was performed only in small part under the current contract, but was included for completeness.

The work described here represents an important step towards accomplishing the difficult tasks of modeling and estimation of complex non-stationary signals. Further research is needed into the fundamental properties of time-varying processes and into the development of digital signal processing techniques for handling such processes.

APPENDIX A

GLOBAL CONVERGENCE OF THE CONSTANT MODULUS ALGORITHM

Global Convergence of the Constant Modulus Algorithm

Julius O. Smith
Benjamin Friedlander

Systems Control Technology Inc.
1801 Page Mill Rd., Palo Alto CA, 94303

Abstract

This paper proves global convergence of the Constant-Modulus Algorithm (CMA) for the case of a real channel when the model order is equal to or greater than that of the channel (the so-called "model-complete" case). The analysis is based on an exact fourth-order Taylor series representation for the cost function minimized asymptotically by the CMA.

1. Introduction

The CMA [3,5,6,7,8] adaptively equalizes constant-modulus communications signals such as frequency-modulation (FM) and phase-modulation (PM) formats. The CMA adaptively minimizes a measure of the amplitude-envelope distortion, such as that caused by multipath propagation, using some form of gradient descent [23,4,9,11]. The amplitude distortion measure is typically a weighted time-average of an error of the form $\hat{e}_t^2 = (1 - \hat{m}_{z_t}^2)^2$, where $\hat{m}_{z_t}$ is the modulus of the equalized output signal at time t. By changing the equalizer parameters based on the gradient of this error measure at each time sample, the channel-induced distortions can be eliminated in many situations [3,8].

This paper examines the convergence behavior of the CMA based on gradient descent. Three forms of the CMA are considered, corresponding to three error measures. The first is the standard CMA for complex data [3], the second is a novel real-only form [9,11], and the third is the pre-existing real-only form [6,8]. These will be referred to as cases 1,2, and 3, respectively. Various extensions of the CMA discussed in [9,11] are incorporated also.

Section 2 gives the CMA problem formulation, and sections 3 and 4 respectively address asymptotic bias and convergence for gradient-based implementations of the CMA.

2. Problem Formulation

Signal Model

Let y_t denote the received signal for $t = 1, 2, \ldots, T$. In the complex case, y_t is assumed to be of the form

$$y_t = H_{xy}(d)x_t + H_{uy}(d)u_t + H_{xy}(d)\nu_t \qquad (1)$$

This work was supported by the U.S. Air Force (AFSC), Rome Air Development Center, under contract no. F30602-84-C-0016.

where $x_t = m_{x_t}e^{j\phi_t}$ is the transmitted signal (ϕ_t is the real-valued information-bearing signal), ν_t is additive white noise at the channel *input*, u_t is an interference signal "template" (assumed known), and $H_{xy}(d)$ and $H_{uy}(d)$ are the unknown linear time-invariant channel filters associated with x_t and u_t respectively:

$$\begin{aligned} H_{xy}(d) &\triangleq \frac{C(d)}{A(d)} \qquad H_{uy}(d) \triangleq \frac{B(d)}{A(d)} \\ A(d) &\triangleq a_0 + a_1 d + a_2 d^2 + \cdots + a_{n_a} d^{n_a} \\ B(d) &\triangleq \qquad b_1 d + b_2 d^2 + \cdots + b_{n_b} d^{n_b} \\ C(d) &\triangleq 1 + c_1 d + c_2 d^2 + \cdots + c_{n_c} d^{n_c} \end{aligned} \qquad (2)$$

where $\{a_i, b_i, c_i\}$ are real. The unit-delay operator d is defined by $d^n x_t = x_{t-n}$ for an arbitrary signal x_t. We assume the modulus $m_{x_t} = |x_t|$ is either constant or known for all t.

In the "real-only" case, the transmitted signal is assumed to be of the form $x_t = m_{x_t}\cos(\phi_t)$, and u_t and ν_t are real interference and additive noise, respectively. We assume $\phi_t = \omega_c t + \psi_t$ where

$$|\partial\psi_t/\partial t| \ll \omega_c \qquad (3)$$

so that positive and negative frequency components are not aliased together.

Error Criteria

The CMA minimizes

$$J(\hat{\theta}; T) \triangleq \frac{1}{T}\sum_{t=1}^{T} w_t \hat{e}_t^2(\hat{\theta}) \qquad (4)$$

with respect to the equalizer parameter-vector $\hat{\theta}$, where

$$\hat{e}_t(\hat{\theta}) = \frac{1}{2}\begin{cases} \left|\hat{z}_t(\hat{\theta})\right|^2 - m_{x_t}^2, & \text{Case 1: Complex Signals} \\ \dfrac{1}{P_t}\displaystyle\sum_{p=0}^{P_t-1} \hat{z}_{t-p}^2(\hat{\theta}) - \sigma_{x_t}^2, & \text{Case 2: Real Signals} \\ \hat{z}_t^2(\hat{\theta}) - \sigma_{x_t}^2, & \text{Case 3: Simplified Real} \end{cases} \qquad (5)$$

is the instantaneous error for the three cases being considered, and

$$
\begin{aligned}
\hat{x}_t(\hat{\theta}) &\triangleq \hat{H}_{xy}^{-1}(d)\left[y_t - \hat{H}_{uy}(d)u_t\right] \\
\hat{H}_{xy}(d) &\triangleq \frac{\hat{C}(d)}{\hat{A}(d)} \qquad \hat{H}_{uy}(d) \triangleq \frac{\hat{B}(d)}{\hat{A}(d)} \\
\hat{A}(d) &\triangleq \hat{a}_0 + \hat{a}_1 d + \hat{a}_2 d^2 + \cdots + \hat{a}_{\hat{n}_a} d^{\hat{n}_a} \\
\hat{B}(d) &\triangleq \qquad \hat{b}_1 d + \hat{b}_2 d^2 + \cdots + \hat{b}_{\hat{n}_b} d^{\hat{n}_b} \\
\hat{C}(d) &\triangleq 1 + \hat{c}_1 d + \hat{c}_2 d^2 + \cdots + \hat{c}_{\hat{n}_c} d^{\hat{n}_c} \\
\hat{\theta} &\triangleq \left[\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{\hat{n}_a}, \hat{b}_1, \ldots, \hat{b}_{\hat{n}_b}, \hat{c}_1, \ldots, \hat{c}_{\hat{n}_c}\right] \\
\theta_0 &\triangleq \left[a_0, a_1, \ldots, a_{\hat{n}_a}, b_1, \ldots, b_{\hat{n}_b}, c_1, \ldots, c_{\hat{n}_c}\right]
\end{aligned}
\tag{6}
$$

If $\hat{n}_a > n_a$, then $a_{n_a+k} \triangleq 0 \forall k > 0$, and similarly for $\hat{n}_b, \hat{n}_c$. In (the real) cases 2 and 3, the modulus m_{x_t} is replaced by the root-mean-square σ_{x_t} (assumed constant or known). The sequence w_t provides a non-negative weight function; typically, $w_t/w_{t+1} \equiv \lambda =$ constant for a fixed exponential decay of past data influence. The signals $\hat{x}_t, y_t, u_t$ are complex in case 1, and real in cases 2 and 3, while the equalizer parameters $\hat{a}_i, \hat{b}_i$, and $\hat{c}_i$ are real in every case.

Case 2 sums the "instantaneous error" $\hat{e}_t$ over the "instantaneous period" P_t. This error definition for the real-signal case is essentially the same as the complex-signal case *if* the imaginary part is constructed at the receiver by delaying the real part one quarter of the instantaneous period. Measurement of the instantaneous period is feasible as long as (3) holds. We assume that the sum of signal samples x_t from time t to $t - P_t + 1$ is exactly zero; in practice, one may wish to interpolate $\hat{x}_t$ to achieve this [13].

Case 3 is the previously studied real-only algorithm [6,8], and here it is viewed as an approximation to case 2. The simplification introduces asymptotic bias and a changed asymptotic parameter variance. The next section contains formulas which can be used to compute the asymptotic bias and variance (the variance being inversely proportional to the curvature of the error surface at the minimum point [4]).

3. Asymptotic Bias

Definition 1. The weighted time-average expectation operator is defined by

$$
\mathrm{E}_t y_t \triangleq \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w_t y_t \tag{7}
$$

where w_t is a fixed non-negative weight function.

Definition 2. The *model-complete case* of the CMA is defined as the case $\hat{n}_a \geq n_a$, $\hat{n}_b \geq n_b$, $\hat{n}_c \geq n_c$, and $\nu_t \equiv 0$. That is, the model coefficients can be set to exactly represent the true channel and interference given y_t and u_t.

Lemma 3 (Bias). For any gradient-based CMA, cases 1 and 2 are locally asymptotically unbiased in the model-complete case, the minimum error being $J(\theta_0) = 0$. The Simplified Real CMA (case 3) is generally asymptotically biased.

Proof. Let $J(\hat{\theta}) \triangleq \lim_{T\to\infty} \frac{1}{2} J(\hat{\theta}; T)$, where $J(\hat{\theta}; T)$ is given by (4). (The factor of 1/2 is introduced to simplify later expressions.) An unbiased minimizer can be obtained only if $J'(\theta_0) = 0$. We need to show that $J'(\theta_0) = 0$ in cases 1 and 2, and that $J'(\theta_0) \neq 0$ in case 3. We have

$$
J(\hat{\theta}) \triangleq \frac{1}{2}\mathrm{E}_t \hat{e}_t^2(\hat{\theta}) \qquad J'(\hat{\theta}) \triangleq \mathrm{E}_t \hat{e}_t(\hat{\theta}) \hat{e}_t'(\hat{\theta}) \tag{8}
$$

where $\hat{e}_t(\hat{\theta})$ is given by (5), and $\hat{e}_t'(\hat{\theta}) \triangleq g(\hat{x}_t \hat{x}_t')$, where

$$
g(\hat{x}_t \hat{x}_t') \triangleq \begin{cases} Re\{\hat{x}_t \hat{x}_t'\}, & \text{Case 1} \\ \dfrac{1}{P_t} \displaystyle\sum_{p=0}^{P_t-1} \hat{x}_{t-p} \hat{x}_{t-p}', & \text{Case 2} \\ \hat{x}_t \hat{x}_t', & \text{Case 3} \end{cases} \tag{9}
$$

and

$$
\begin{aligned}
\hat{x}_t &= \hat{\varphi}_t^T \hat{\theta}, \qquad \hat{x}_t' = \hat{\varphi}_t^f, \qquad \hat{\varphi}_t^f = \hat{\varphi}_t / \hat{C}(d) \\
\hat{\varphi}_t &= [y_t, y_{t-1}, \ldots, y_{t-\hat{n}_a}, -u_{t-1}, \ldots, -u_{t-\hat{n}_b}, \\
&\qquad - \hat{x}_{t-1}, \ldots, -\hat{x}_{t-\hat{n}_c}]^T
\end{aligned}
\tag{10}
$$

Note that g is a linear operator.

Now, $J'(\hat{\theta}) = 0$ iff $\mathrm{E}_t \hat{e}_t(\hat{\theta}) g(\hat{\varphi}_t \hat{\varphi}_t^T \hat{\theta}) = 0$. In cases 1 and 2, $\hat{e}_t(\theta_0) \equiv 0 \Rightarrow J'(\theta_0) = 0$.

In case 3, we have

$$
\begin{aligned}
J'(\theta_0) &= \mathrm{E}_t [\hat{x}_t^2(\theta_0) - \sigma_{x_t}^2] \hat{\varphi}_t \hat{\varphi}_t^T \theta_0 \\
&= \mathrm{E}_t x_t^3 \hat{\varphi}_t - \sigma_{x_t}^2 \mathrm{E}_t x_t \hat{\varphi}_t
\end{aligned}
\tag{11}
$$

It suffices to show $J'(\theta_0) \neq 0$ for constant $\sigma_{x_t} = \sigma_x$, no interference, and degenerate channel $\theta_0 = 1$. In this case, $y_t = x_t = m_x \cos(\phi_t)$, $m_x = \sqrt{2}\sigma_x$, and $\varphi_t = x_t$. The error-surface gradient is then $J'(\theta_0) = \mathrm{E}_t x_t^4 - \sigma_{x_t}^2 \mathrm{E}_t x_t^2$. Setting $m_x = 1$ and $\phi_t = \omega_c t$, we find $\mathrm{E}_t x_t^2 = \mathrm{E}_t \cos^2(\omega_c t) = 1/2$ while $\mathrm{E}_t x_t^4 = \mathrm{E}_t \cos^4(\omega_c t) = 3/8 \neq \sigma_{x_t}^2 \mathrm{E}_t x_t^2 = (1/2)^2$. Thus, case 3 necessarily yields a biased solution under gradient descent.

4. Asymptotic Convergence

Definition 4. An m-dimensional matrix R of order n is defined as any scalar function over a set of m-tuples $R[i_1, i_2, \ldots, i_m]$ where each index i_j ranges from 1 to n, $j = 1, 2, \ldots, m$. Such a matrix will be called an (m, n) matrix.

Definition 5. The j-product of an (m, n) matrix R times a $(1, n)$-matrix x (n-vector) is defined as the $(m-1, n)$-

matrix

$$R^j x_j = \sum_{i_j=1}^{n} R[i_1, i_2, \ldots, i_m] x[i_j], \quad 1 \le j \le m \qquad (12)$$

Definition 6. An (m, n) matrix R is said to be *nonsingular* if its j-product with any nonzero n-vector is nonzero for all j, i.e., $R^j x_j = 0$ for some j implies $x = 0$.

Definition 7. A signal y_t is said to be *persistently exciting (PE) of order* (m, n) if the (m, n)-matrix

$$R_y \triangleq \bar{E}_t y_{t-i_1} y_{t-i_2} \cdots y_{t-i_m}, \; i_j = 1, \ldots, n \qquad (13)$$

is nonsingular. This definition can be regarded as applying to the received realization y_t of an underlying random process. Normally we expect such a random process to be PE with probability one (wp1). Note also that the weighting w_t used in the time-averaging operator $\bar{E}_t$ can affect whether y_t is PE.

Definition 8. A signal y_t is said to be *persistently exciting of order* $(4, n)$ *with respect to the scalar function* g if the $(4, n)$-matrix

$$R_y^g[i, j, k, l] \triangleq \bar{E}_t g(y_{t-i} y_{t-j}) g(y_{t-k} y_{t-l}) \qquad (14)$$

is nonsingular.

Definition 9. A *convergent gradient descent* algorithm is any iterative algorithm for $\hat{\theta}$ (of the form $\hat{\theta}_{k+1} = f_k(\hat{\theta}_k)$) [2] which converges to either a stationary point $\hat{\theta}^*$ of the error surface $J(\hat{\theta})$ (in which case $J'(\hat{\theta}^*) = 0$), or to a point on the constraint boundary for $\hat{\theta}$ (if any). Normally this property is obtained by using a diminishing non-summable step-size (such as $1/k$) in the gradient descent iteration [4].

Theorem 10 (Global convergence). In the model-complete case, y_t persistently exciting of order $(4, \hat{n}_a)$ with respect to g (of (9)) wp1, no interference, $\hat{n}_b = \hat{n}_c = 0$, then any unconstrained convergent gradient descent CMA will converge with probability one to

$$\hat{\theta}^* = \pm\theta_0 = \pm[a_0, a_1, a_2, \ldots, a_{n_a}, 0, \ldots, 0]^T \qquad (15)$$

in cases 1 and 2, for any nonzero initial parameter vector $\hat{\theta}(0) \ne 0$.

Proof. The stationary points of a gradient descent algorithm occur at points $\hat{\theta}$ where the gradient $J'(\hat{\theta})$ is zero. The first goal is to find all such points.

Let $\hat{\theta}^*$ denote any local minimizer of the cost function $J(\hat{\theta})$ given by equation (8). By definition, $J'(\hat{\theta}^*) = 0$. By lemma 3, $\hat{\theta}^* = \theta_0$ is one such minimizer. Let

$$\begin{aligned} \hat{x}_t &\triangleq \varphi_t^T \hat{\theta}, \qquad \hat{x}_t^* \triangleq \varphi_t^T \hat{\theta}^* \\ \tilde{x}_t &\triangleq \varphi_t^T(\hat{\theta} - \hat{\theta}^*) = \hat{x}_t - \hat{x}_t^* \\ \varphi_t^T &\triangleq [y_t, y_{t-1}, \ldots, y_{t-\hat{n}_a}] \end{aligned} \qquad (16)$$

and define $g(\cdot)$ as in (9). Expanding $J(\hat{\theta})$ in a Taylor series about $\hat{\theta}^*$ gives (exactly)

$$\begin{aligned} J(\hat{\theta}) = J(\hat{\theta}^*) + 0 + \frac{1}{2}\bar{E}_t\Big[g^2(\hat{x}_t^* \tilde{x}_t) + \hat{e}_t(\hat{\theta}^*) g(\tilde{x}_t^2)\Big] \\ + \frac{1}{2}\bar{E}_t g(\hat{x}_t^* \tilde{x}_t) g(\tilde{x}_t^2) + \frac{1}{8}\bar{E}_t g^2(\tilde{x}_t^2) \end{aligned} \qquad (17)$$

and

$$\begin{aligned} J'(\hat{\theta}) &= \bar{E}_t\Big[g(\hat{x}_t^* \tilde{x}_t) g(\hat{x}_t^* \varphi_t) + \hat{e}_t(\hat{\theta}^*) g(\tilde{x}_t \varphi_t)\Big] \\ &\quad + \frac{1}{2}\bar{E}_t g(\tilde{x}_t^2) g(\hat{x}_t^* \varphi_t) + 2g(\hat{x}_t^* \tilde{x}_t) g(\tilde{x}_t \varphi_t) \\ &\quad + \frac{1}{2}\bar{E}_t g(\tilde{x}_t^2) g(\tilde{x}_t \varphi_t) \\ &= \frac{1}{2}\bar{E}_t\Big[\left(\hat{\theta} - \hat{\theta}^*\right)^T g(\varphi_t \varphi_t^T)\left(\hat{\theta} + \hat{\theta}^*\right)\Big]\Big[g(\varphi_t \varphi_t^T)\hat{\theta}\Big] \\ &\quad + \bar{E}_t g\Big[\hat{e}_t(\hat{\theta}^*)\varphi_t \varphi_t^T\Big]\left(\hat{\theta} - \hat{\theta}^*\right) \end{aligned} \qquad (18)$$

Consider first cases 1 and 2. Without loss of generality we can set $\hat{\theta}^* = \theta_0$ in (18). This gives immediately $J'(\theta_0) = J'(-\theta_0) = J'(0) = 0$. Thus we have found three stationary points for cases 1 and 2. We now show that these are the only stationary points. The lth row of equation (18) with $\hat{\theta}^*$ set to θ_0 (valid for cases 1 and 2 only) can be rewritten as

$$J'(\hat{\theta})[l] = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_i \beta_j \gamma_k R_y^g[i, j, k, l] \qquad (19)$$

where $\alpha_i = (\hat{\theta} + \theta_0)[i]$, $\beta_j = \hat{\theta}[j]$, and $\gamma_k = (\hat{\theta} - \theta_0)[k]$. The $(4, \hat{n}_a)$-matrix $R_y^g[i, j, k, l] \triangleq \bar{E}_t g(y_{t-i} y_{t-j}) g(y_{t-k} y_{t-l})$ can be interpreted as a type of four-dimensional covariance matrix. We see that a sufficient condition for the three solutions $\hat{\theta} \in \{0, \theta_0, -\theta_0\}$ to be the only solutions is to have R_y^g be nonsingular. But this holds whenever y_t is persistently exciting of order $(4, \hat{n}_a)$ with respect to g. Hence, $0, \theta_0$, and $-\theta_0$ are the only stationary points in cases 1 and 2 with probability one.

It remains to be shown that $\hat{\theta} = 0$ is an *unstable* stationary point, while $\hat{\theta} = \pm\theta_0$ are *stable* stationary points for the gradient. A stationary point $\hat{\theta}^*$ is stable if the matrix $J''(\hat{\theta}^*)$ is positive definite, and unstable if $J''(\hat{\theta}^*)$ is negative definite [2]. It is straightforward to show that

$$J''(0) = -J''(\pm\theta_0)/2 \qquad (20)$$

That is, the curvature of the error surface at $\hat{\theta} = 0$ is equal to $-1/2$ the curvature at $\hat{\theta} = \theta_0$, and the curvature at $-\theta_0$ equals that at θ_0. It is always true that $J''(\theta_0) \ge 0$ (see the third term in (17)), and y_t persistently exciting of order $(4, \hat{n}_a)$ with respect to g implies $J''(\theta_0) > 0$ (positive definite).

Thus $J''(0) < 0$, making $\dot\theta = 0$ an unstable stationary point. (In all current forms of the CMA, $\dot\theta$ can never be allowed to equal zero, for this will in fact "freeze" a gradient descent algorithm.)

The two remaining solutions $\dot\theta = \pm\vartheta_0$ are *stable* stationary points. $[J''(-\theta_0) = J''(\theta_0) > 0]$. Under the assumed conditions, $\dot\theta = \pm\theta_0$ are the *only* stable stationary points of the complex or the real CMA (cases 1 and 2). Since in practice $\hat{H}_{xy}(d)$ is divided through by $\dot a_0$, the sign ambiguity in $\dot\theta^*$ becomes a sign ambiguity in $\hat{H}_{xy}(d)$; a sign change in a signal is normally negligible. It is interesting to note that this sign ambiguity corresponds to the phase ambiguity in the complex-channel case [3].

Further Extensions

Theorem 10 immediately extends to the case which includes interference canceling ($\hat n_b \geq n_b > 0$). Similar global convergence results may hold for the model-complete cases in which $\hat n_c > 0$; for example, the techniques in [1] might go through. It has been shown [9] that either noise ($\nu_t \not\equiv 0$) or model-incomplete interference ($\hat{H}_{uy} \neq H_{uy}\ \forall\hat\theta$) will cause bias in the parameters. For case 3, theorem 10 can be extended to prove global convergence to a biased solution in which the bias can be simply approximated as in [6].

Convergence results for specific algorithms can be obtained using the analytical approach described in [1]. In the model incomplete case, the general model (5) should converge to a local minimum of the error surface $J(\hat\theta)$; the number of sub-optimal local minima can be large in the model-incomplete case.

We expect that if $y_t = H_{xy}(d)m_x e^{j\phi_t}$ and ϕ_t is randomly distributed with almost any non-discrete distribution, then R_y^q will be nonsingular with probability one for any number n of parameters. It seems, however, that for good numerical conditioning, further restrictions are necessary on the modulating signal. For example, it might be appropriate to require y_t to possess at least n distinct frequencies of high spectral power, analogous to the situation in least-squares system identification [4]. Further work is necessary to specify precisely the modulation characteristics which maximize the equalization accuracy.

The error minimized in the model-complete case is exactly described by a fourth-order Taylor series (cf. equation (17)). A fourth-order type of gradient-descent algorithm, analogous to Newton's method for least squares problems, should yield the fastest-converging algorithms. To this end, note that all solutions to the ensuing third-order "vector polynomial" for the gradient can be expressed in closed form.

Conclusions

Some convergence properties of the Constant Modulus Algorithm (CMA) for channel equalization were described. Subject to mild restrictions on the modulation signal, the unbiased versions of the CMA can only converge to plus or minus the true solution in the model-complete case.

References

[1] K. J. Astrom and T. Soderstrom, "Uniqueness of the Maximum Likelihood Estimates of the Parameters of an ARMA Model," *IEEE Trans. Automat. Contr.*, vol. AC-19, No. 6, pp. 769–773, Dec. 1974.

[2] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, New York, 1981.

[3] J. R. Treichler and B. C. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. on Acoust., Speech, and Signal Proc.*, vol. ASSP-31, pp. 459–472, April 1983.

[4] L. Ljung and T. L. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge MA, 1983.

[5] M. G. Larimore and J. R. Treichler, "Convergence Behavior of the Constant Modulus Algorithm," *ICASSP-83: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 1.4, Boston MA, March 19-21, 1984.

[6] J. R. Treichler and M. G. Larimore, "A Real-Arithmetic Implementation of the Constant Modulus Algorithm," *ICASSP-84: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 3.2, San Diego CA, March 19-21, 1984.

[7] J. R. Treichler, "Algorithms that Restore Signal Properties," *ICASSP-84: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 21.4, San Diego CA, March 19-21, 1984.

[8] J. R. Treichler and M. G. Larimore, "New Processing Techniques Based on the Constant Modulus Adaptive Algorithm," to appear.

[9] J. O. Smith and B. Friedlander, "Analysis and Extensions of the Constant Modulus Algorithm," Tech. Rep. 6502-01, Systems Control Tech., Oct. 1984.

[10] J. O. Smith and B. Friedlander, "Simulation Results for Extensions of the Constant Modulus Algorithm," Tech. Rep. 6502-02, Systems Control Tech., Oct. 1984.

[11] J. O. Smith and B. Friedlander, "Extensions of the Constant Modulus Algorithm," Asilomar, Nov. 1984.

[12] J. R. Treichler, "Capture Properties of the Constant Modulus Algorithm," Asilomar, Nov. 1984.

[13] J. O. Smith and B. Friedlander, "Adaptive Interpolated Time-Delay Estimation," Asilomar, 1983. Full version to appear in the *IEEE Trans. on Aerospace*, May 1985.

APPENDIX B

EXTENSIONS OF THE CONSTANT MODULUS ALGORITHM

Extensions of the Constant Modulus Algorithm

Julius O. Smith
Benjamin Friedlander

Systems Control Technology Inc.
1801 Page Mill Rd., Palo Alto CA, 94303

Abstract

The Constant-Modulus Algorithm (CMA) computes and applies an adaptive channel equalizer for constant-amplitude signals such as frequency- and phase-modulation. This paper presents several extensions to the CMA including IIR equalization, a real-signal version having properties as good as the complex version, use of the Gauss-Newton method in place of gradient descent, interference rejection, and more.

1. Introduction

The Constant Modulus Algorithm (CMA) was introduced by J. R. Treichler et. al. [4,7,8,9,12] as a method for adaptive equalization of certain types of communications signals. The CMA calibrates a linear channel equalizer by seeking to make the equalized signal have constant modulus. Such a technique can be used with frequency-modulation (FM) and phase-modulation (PM) communications systems, when the amplitude envelope is constant or known in the absence of distortion. The CMA is one example of adaptive channel equalization based on restoring invariant properties of the undistorted signal [6].

In [4] the received signal is assumed available as a complex signal of the form $e^{j\phi_t}$. In this case, the modulus is exactly constant when channel distortions are absent. In [8,12], a real-only version of the CMA is developed. Most recently, the CMA has been extended to known-amplitude (as opposed to constant-amplitude), and to multichannel equalization. In all cases, the equalizer is formed by applying a fixed-step gradient-descent algorithm similar to the Widrow LMS algorithm [1].

The purpose of this paper is to present several extensions to the basic CMA:

- The channel equalizer is an arbitrary rational filter (poles and zeros), or infinite-impulse-response (IIR) filter, rather than being constrained to a finite-impulse-response (FIR) all-zero filter as in earlier works. Use of an all-pole equalizer, for example, allows an exact inverse to FIR channel distortions such as multipath.
- A new type of real-only algorithm is derived. The new real-only version is based on a more desirable error definition which is equivalent to the complex case when the instantaneous carrier period is used.
- The Recursive Gauss-Newton method (RGN) [3,6] replaces the gradient descent method for obtaining the equalizer coefficients. This modification can improve the convergence rate and asymptotic accuracy of the equalization.
- A new type of RGN algorithm is presented [5] which allows a trade-off between properties of the more recent recursive forms and the original "offline" or "batch" forms.
- A technique for removing partially-known interference is incorporated into the algorithm.

In section 2, the basic problem formulation is presented, and section 3 derives the algorithm.

2. Problem Formulation

Signal Model

Let y_t denote the received signal for $t = 1, 2, \ldots, T$. In the complex case, y_t is assumed to be of the form

$$y_t = H_{xy}(d)x_t + H_{uy}(d)u_t + H_{xy}(d)\nu_t \qquad (1)$$

where $x_t = m_{x_t}e^{j\phi_t}$ is the transmitted signal (ϕ_t is the real-valued information-bearing signal), ν_t is additive noise, u_t is an interference signal (assumed at least partially known), and $H_{xy}(d)$ and $H_{uy}(d)$ are the linear time-invariant channel filters associated with x_t and u_t respectively:

$$\begin{aligned} H_{xy}(d) &\triangleq \frac{C(d)}{A(d)} \qquad H_{uy}(d) \triangleq \frac{B(d)}{A(d)} \\ A(d) &\triangleq a_0 + a_1 d + a_2 d^2 + \cdots + a_{n_a} d^{n_a} \\ B(d) &\triangleq \quad b_1 d + b_2 d^2 + \cdots + b_{n_b} d^{n_b} \\ C(d) &\triangleq 1 + c_1 d + c_2 d^2 + \cdots + c_{n_c} d^{n_c} \end{aligned} \qquad (2)$$

where $\{a_i, b_i, c_i\}$ are real. The unit-delay operator d is defined by the relation $d^n x_t = x_{t-n}$ for an arbitrary signal x_t. We assume the modulus m_{x_t} of the transmitted signal x_t is known for all t.

The novel features introduced so far are the provision for $H_{uy}(d)$ and u_t to model partially-known interference, and

This work was supported by the U.S. Air Force (AFSC), Rome Air Development Center, under Contract No. F30602-84-C-0016.

the polynomial $C(d)$ in the channel transfer function $H_{zy}(d)$ for modeling FIR distortion types.

In the case of real signals, the transmitted signal is assumed to be of the form $z_t = m_{z_t} \cos(\phi_t)$ and u_t and ν_t are real interference and additive noise, respectively. We also assume $\phi_t = \omega_c t + \psi_t$ where $|\partial \psi_t / \partial t| \ll \omega_c$. In this case the known-modulus property can be replaced by the *known-envelope* property. However, to retain a least-squares problem, we use instead an *approximately known average power* property. Thus, in the real-signal case,

$$\frac{1}{P_t} \sum_{p=0}^{P_t - 1} z_{t-p}^2 \approx \sigma_{z_t}^2 \tag{3}$$

where P_t is the "instantaneous period" of z_t, and $\sigma_{z_t}^2$ is defined as the "instantaneous power," or "average power" of z_t. For simplicity, we later use equality in (3). The perturbations in (3) due to discrete P_t and modulation by ϕ_t can be compensated if desired by waveform interpolation and redefinition of continuous P_t to absorb the carrier-cycle support distortion under integration.

The Error Criterion

Below we define three error criteria for the CMA. The three error criteria correspond to the case of complex z_t [4], real z_t, and a simplified version for real z_t [8,12]. These will be referred to as cases 1,2, and 3, respectively.

Case 1: Error for Complex Signals

In the complex case, the CMA minimizes

$$J_1(\hat{\theta}; T) \triangleq \frac{1}{T} \sum_{t=1}^{T} w_t \hat{e}_{1,t}^2(\hat{\theta}) \tag{4}$$

with respect to the equalizer parameter-vector $\hat{\theta}$, where

$$\begin{aligned}
\hat{e}_{1,t}(\hat{\theta}) &\triangleq \left| \hat{z}_t(\hat{\theta}) \right|^2 - m_{z_t}^2 \\
\hat{z}_t(\hat{\theta}) &\triangleq \hat{H}_{zy}^{-1}(d) \Big[y_t - \hat{H}_{uy}(d) u_t \Big] \\
\hat{H}_{zy}(d) &\triangleq \frac{\hat{C}(d)}{\hat{A}(d)} \qquad \hat{H}_{uy}(d) \triangleq \frac{\hat{B}(d)}{\hat{A}(d)} \\
\hat{A}(d) &\triangleq \hat{a}_0 + \hat{a}_1 d + \hat{a}_2 d^2 + \cdots + \hat{a}_{n_a} d^{n_a} \\
\hat{B}(d) &\triangleq \qquad \hat{b}_1 d + \hat{b}_2 d^2 + \cdots + \hat{b}_{n_b} d^{n_b} \\
\hat{C}(d) &\triangleq 1 + \hat{c}_1 d + \hat{c}_2 d^2 + \cdots + \hat{c}_{n_c} d^{n_c} \\
\hat{\theta} &\triangleq \Big[\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{n_a}, \hat{b}_1, \ldots, \hat{b}_{n_b}, \hat{c}_1, \ldots, \hat{c}_{n_c} \Big]
\end{aligned} \tag{5}$$

and w_t is a non-negative weighting function. The signals $\hat{z}_t, y_t, u_t$ are complex, while the equalizer parameters $\hat{a}_i, \hat{b}_i$, and $\hat{c}_i$ are real.

We show in [11] that if $\nu_t \equiv 0$, $n_b = n_c = 0$, $\hat{n}_a \geq n_a$, $\hat{n}_b \geq n_b$, and z_t is *persistently exciting* in a novel sense, then the only parameters $\hat{\theta}$ which locally minimize $J_1(\hat{\theta}; T)$ are $\hat{\theta} = \pm\theta_0$ (where θ_0 is extended with zeros to the size of $\hat{\theta}$ if necessary). This is valuable to know since it means that the only minimizers of $J_1(\hat{\theta}; T)$ are global minimizers. In such a case, gradient descent methods can be used with confidence in the so-called "model-complete" case (i.e., when the channel model can exactly describe the true channel distortion). Furthermore, the persistence-of-excitation condition holds for "almost all" information-laden modulation signals ϕ_t.

Case 2: Error for Real Signals

We define the real-only case as minimizing

$$J_2(\hat{\theta}; T) \triangleq \frac{1}{T} \sum_{t=1}^{T} w_t \hat{e}_{2,t}^2(\hat{\theta}) \tag{6}$$

where

$$\hat{e}_{2,t}(\hat{\theta}) \triangleq \frac{1}{P_t} \sum_{p=0}^{P_t - 1} \hat{z}_{t-p}^2(\hat{\theta}) - \sigma_{z_t}^2 \tag{7}$$

and all other quantities are formally the same as in (5) with $\hat{z}_t, y_t, u_t$ real. Thus the main difference between the complex and real-only cases is that the "instantaneous error" $\hat{e}_{1,t}$ is replaced by an error $\hat{e}_{2,t}$ which is an average over the "instantaneous period" P_t. This definition for the real-signal case is *equivalent* in principle to the complex-signal case when the imaginary part is constructed at the receiver from the real part.

While (7) may seem like the natural error to use in the real-only case, it requires an estimate of the instantaneous period P_t which depends on the modulation signal ϕ_t. The counterpart in the complex CMA is to define the imaginary part as a quarter-cycle delay of the real part, where the quarter-cycle delay is one-fourth of the current instantaneous period. Note that measurement of the instantaneous period requires using a carrier frequency much higher than the signal bandwidth.

An alternative to tracking the instantaneous period is to fix P_t to a value which spans many cycles of the carrier under all modulation conditions. The counterpart in the complex CMA is to use the Hilbert transform to create the imaginary part of z_t (as is typically done in practice). Averaging over many carrier cycles has the advantage of not requiring an estimate of instantaneous period, but the disadvantage of limiting the potential convergence rate of the equalizer parameters $\hat{\theta}$.

Case 3: Simplified Error for Real Signals

A simplified real-only CMA [8,12] minimizes

$$J_3(\hat{\theta}; T) \triangleq \frac{1}{T} \sum_{t=1}^{T} w_t \hat{e}_{3,t}^2(\hat{\theta}) \tag{8}$$

where

$$\hat{e}_{3,t}(\hat{\theta}) \triangleq \hat{z}_t^2(\hat{\theta}) - \sigma_{z_t}^2 \tag{9}$$

While this is not obviously an appropriate error criterion for non-constant signals z_t, it can be shown [11] that (9) has asymptotic properties very similar to those of (7). Intuitively, the reason is that the periodic fluctuations of the error in (9) as the instantaneous carrier frequency average out. The error simplification introduces *asymptotic bias* (often approximable by a simple scale factor [8,11]) and a slightly different *asymptotic parameter variance* relative to the optimal versions (4) and (6) [11].

3. The Algorithm

Prior treatments of the CMA have been based on gradient descent [4,7,8,9,12]. Gradient descent can be seen as a linearization method, while Newton descent moves each iteration to the bottom of a quadratic local approximation. In Newton descent, close to the minimum point, convergence is quadratic while gradient descent converges linearly [3]. Thus, Newton's method converges faster near a local minimum (assuming the Hessian is nonzero).

The Gauss-Newton method [3] is a robust approximate Newton's method. It is preferable over Newton's method when time-recursive algorithms are desired. In this section, we derive the Gauss-Newton method as applied to the CMA.

The CMA cost function can be expressed as

$$J(\hat{\theta}; T) \triangleq \frac{1}{2}\frac{1}{T}\sum_{t=1}^{T} w_t \hat{e}_t^2(\hat{\theta}) \tag{10}$$

where w_t is a non-negative real-valued weighting function, and

$$\hat{e}_t(\hat{\theta}) = \frac{1}{2}\begin{cases} \left|\hat{z}_t(\hat{\theta})\right|^2 - m_{z_t}^2, & \text{Case 1: Complex Signals} \\ \frac{1}{P_t}\sum_{p=0}^{P_t-1} \hat{z}_{t-p}^2(\hat{\theta}) - \sigma_{z_t}^2, & \text{Case 2: Real Signals} \\ \hat{z}_t^2(\hat{\theta}) - \sigma_{z_t}^2, & \text{Case 3: Simplified Real} \end{cases} \tag{11}$$

The factors of 1/2 in (10) and (11) have been introduced to simplify later expressions. Let

$$\begin{aligned} \hat{e}_t'(\hat{\theta}) &\triangleq \frac{\partial \hat{e}_t(\theta)}{\partial \theta}(\hat{\theta}) \\ \hat{e}_t''(\hat{\theta}) &\triangleq \frac{\partial^2 \hat{e}_t(\theta)}{\partial \theta^2}(\hat{\theta}) \end{aligned} \tag{12}$$

The Gauss-Newton (GN) Method

Let $\hat{\theta}^*$ denote a local minimizer of $J(\hat{\theta}; T)$. When the error $\hat{e}_t$ can be driven to zero as $\hat{\theta}$ approaches $\hat{\theta}^*$, or $\hat{e}_t(\hat{\theta}^*)$ and $\hat{e}_t''(\hat{\theta}^*)$ are uncorrelated, the Hessian can be closely approximated by $\sum_t \hat{e}_t'\hat{e}_t'^T$. The *Gauss-Newton* [3] method is defined accordingly as follows. Given an initial parameter estimate $\hat{\theta}_0$, carry out the following iteration until convergence is achieved:

$$\begin{aligned} R(\hat{\theta}_i; T) &= \frac{1}{T}\sum_{t=1}^{T} w_t \hat{e}_t'(\hat{\theta}_i)\hat{e}_t'(\hat{\theta}_i)^T \approx J''(\hat{\theta}_i; T) \\ \hat{\theta}_{i+1} &= \hat{\theta}_i - R^{-1}(\hat{\theta}_i; T)J'(\hat{\theta}_i; T) \end{aligned} \tag{13}$$

The Hessian approximation serves two purposes. First, it eliminates the need for second-order differentiation—the gradient alone drives the parameter updates. Second, the matrix $R(\hat{\theta}_i; T)$ is more likely to be positive definite and invertible than is $J''(\hat{\theta}; T)$ [3].

Exact Recursive GN Updates

The Gauss-Newton algorithm (13) can made recursive, where the parameter estimate $\hat{\theta}$ is updated for each t. We now consider only one pass through the T data points. Therefore, the pass-number i in (13) is dropped for notational simplicity. The initial value $\hat{\theta}_i$ from the previous pass is denoted $\hat{\theta}_0$, and the final estimate $\hat{\theta}_{i+1}$ obtained at the end of pass i now becomes $\hat{\theta}_T$

Define

$$R_T \triangleq \frac{1}{w_T}\sum_{t=1}^{T} w_t \hat{e}_t'(\hat{\theta}_0)\hat{e}_t'(\hat{\theta}_0)^T \quad G_T \triangleq \frac{1}{w_T}\sum_{t=1}^{T} w_t \hat{e}_t'(\hat{\theta}_0)\hat{e}_t(\hat{\theta}_0) \tag{14}$$

Then

$$\begin{aligned} R_t &= \lambda_t R_{t-1} + \hat{e}_t'(\hat{\theta}_0)\hat{e}_t'(\hat{\theta}_0)^T \\ G_t &= \lambda_t G_{t-1} + \hat{e}_t'(\hat{\theta}_0)\hat{e}_t(\hat{\theta}_0) \end{aligned} \tag{15}$$

where $\lambda_t \triangleq w_{t-1}/w_t$ is known as the "forgetting factor." Typically λ_t is fixed between 0 and 1 to obtain a fixed exponential fading of past observations (with an approximate time constant of $1/(1-\lambda)$ samples). It can be shown [5] that for Δ any integer between 1 and t,

$$\begin{aligned} \hat{\theta}_t &= \hat{\theta}_0 - R_t^{-1}G_t \\ &= \hat{\theta}_{t-\Delta} - R_t^{-1}\frac{1}{w_t}\sum_{i=t-\Delta+1}^{t} w_i \hat{e}_i'(\hat{\theta}_0) \times \\ &\qquad \left[\hat{e}_i(\hat{\theta}_0) - \hat{e}_i'(\hat{\theta}_0)^T\left(\hat{\theta}_{t-\Delta} - \hat{\theta}_0\right)\right] \end{aligned} \tag{16}$$

This recursive form is exactly equivalent to the "off-line" counterpart (13).

Block-Recursive Gauss-Newton (BRGN)

The BRGN [5] is obtained by replacing the initial condition $\hat{\theta}_0$ by recently computed estimates $\hat{\theta}_t$ at periodic intervals. When $\hat{\theta}_0$ is replaced by $\hat{\theta}_t$, we effectively split the data into more than one "batch." The first batch consists of data up through time t, and it is used to produce the estimate $\hat{\vartheta}_t$. The second batch begins at time $t+1$, and it is initialized to the estimate produced by the first batch.

Let P_t denote the "batch size" in an off-line Gauss-Newton method implemented recursively via equation (16). Assume for simplicity that Δ divides P_t. Then (16) becomes

$$
\begin{aligned}
\hat{\theta}_t &= \hat{\vartheta}_{t-P_t} - R_t^{-1} G_t \\
&= \hat{\vartheta}_{t-\Delta} - R_t^{-1} \frac{1}{w_t} \sum_{i=t-\Delta+1}^{t} w_i \dot{\varepsilon}_i(\hat{\theta}_{t-P_t}) \quad \times \\
&\qquad \left[\hat{\varepsilon}_i(\hat{\theta}_{t-P_t}) + \dot{\varepsilon}_i(\hat{\theta}_{t-P_t})^T \left(\hat{\vartheta}_{t-\Delta} - \hat{\vartheta}_{t-P_t}\right)\right]
\end{aligned} \tag{17}
$$

This adaptive algorithm allows interpolation between the properties of an "off-line" or "batch" optimization and the truly time-recursive algorithms which are used extensively in the identification literature. The practical importance of the BRGN for the CMA is that it allows parameter estimates *per-period* in the *real-only case*. The carrier period can be estimated by BRGN *itself* or it can derived from the demodulated signal in a "decision-directed" mode.

<u>Recursive Gauss-Newton (RGN)</u>

By setting the block size P_t to 1 (forcing Δ to 1), the BRGN reduces to the *recursive Gauss-Newton method* (RGN) which is somewhat of a standard method for ARMAX system identification [6]:

$$
\begin{aligned}
\hat{\theta}_t &= \hat{\theta}_{t-1} - R_t^{-1} \dot{\varepsilon}_t(\hat{\theta}_{t-1}) \hat{\varepsilon}_t(\hat{\theta}_{t-1}) \\
R_t &= \lambda_t R_{t-1} + \dot{\varepsilon}_t(\hat{\theta}_{t-1}) \dot{\varepsilon}_t(\hat{\theta}_{t-1})^T
\end{aligned} \tag{18}
$$

Ljung [6] and others have shown that, under general conditions, the replacement of $\hat{\theta}_0$ by even the most recently available parameter estimates $\hat{\theta}_{t-1}$ does not alter the asymptotic convergence properties of the Gauss-Newton method.

To tailor GN, BRGN, or RGN to a specific application, the instantaneous error $\hat{\varepsilon}_t$ and its gradient $\dot{\varepsilon}_t$ with respect to the most recently computed parameters $\hat{\vartheta}$ are needed.

<u>The CMA Error and Gradient</u>

The "instantaneous gradient" of $\hat{\varepsilon}_t$ with respect to $\hat{\vartheta}$ is given by $\dot{\varepsilon}_t(\hat{\vartheta}) \triangleq g(\hat{z}_t \dot{z}_t)$ where

$$
g(\hat{z}_t \dot{z}_t) \triangleq \begin{cases} Re\{\hat{z}_t \dot{z}_t\}, & \text{Case 1: Complex Signals} \\ \dfrac{1}{P_t} \displaystyle\sum_{p=0}^{P_t-1} \hat{z}_{t-p} \dot{z}_{t-p}, & \text{Case 2: Real Signals} \\ \hat{z}_t \dot{z}_t, & \text{Case 3: Simplified Real} \end{cases} \tag{19}
$$

in all cases.

$$
\begin{aligned}
\hat{z}_t &\triangleq \hat{\varphi}_t^T \hat{\vartheta} \\
\hat{\varphi}_t &\triangleq [y_t, \ldots, y_{t-n_a}, -u_{t-1}, \ldots, -u_{t-n_b}, -\hat{z}_{t-1}, \ldots, -\hat{z}_{t-n_c}] \\
\hat{\theta} &\triangleq \left[\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{n_a}, \hat{b}_1, \ldots, \hat{b}_{n_b}, \hat{c}_1, \ldots, \hat{c}_{n_c}\right] \\
\dot{z}_t &= \hat{\varphi}_t^f \triangleq \hat{\varphi}_t - \hat{c}_1 \hat{\varphi}_{t-1}^f - \hat{c}_2 \hat{\varphi}_{t-2}^f - \cdots - \hat{c}_{n_c} \hat{\varphi}_{t-n_c}^f
\end{aligned} \tag{20}
$$

These quantities determine the error $\hat{\varepsilon}_t$ and its gradient $\dot{\varepsilon}_t$ needed to drive the Gauss-Newton algorithm.

Because the recursions for $\hat{z}_t$ and $\hat{\varphi}_t^f$ involve the application of the time-varying filter $1/\hat{C}_t(d)$, it is necessary that $\hat{C}_t(d)$ be stable at all times. Our strategy for stability projection is as follows: If any root of $\hat{C}_t(z)$ lies on or outside the unit circle, the roots are contracted uniformly by the factor $\rho = 0.95$ repeatedly until all roots are inside. This angle-invariant projection preserves the tuning of spectral resonances in the equalizer (e.g. multipath nulls).

In the RGN algorithm, $\hat{\varphi}_t^f = \hat{\varphi}_t/\hat{C}(d)$, where $\hat{C}$ is constantly being replaced by the latest available estimate. Since $\hat{z}_t$ is computed using $\hat{\vartheta}_{t-1}$, the first occurrence of $\hat{z}$ in $\hat{\varphi}^f$ depends also on $\hat{\vartheta}_{t-1}$. It has been observed in the analogous system identification context that convergence is accelerated by recomputing $\hat{z}_t$ using $\hat{\vartheta}_t$ before using it in the recursion for $\hat{\varphi}_t^f$.

<u>Algorithm Summary</u>

$$
\begin{aligned}
y_t^f &= y_t - \hat{c}_{t-1}(1) y_{t-1}^f - \cdots - \hat{c}_{t-1}(n_c) y_{t-n_c}^f \\
\hat{\varphi}_t &= [y_t, \ldots, y_{t-n_a}, -u_{t-1}, \ldots, -u_{t-n_b}, -\hat{z}_{t-1}, \ldots, -\hat{z}_{t-n_c}]^T \\
\hat{\varphi}_t^f &= [y_t^f, \ldots, y_{t-n_a}^f, -u_{t-1}^f, \ldots, -u_{t-n_b}^f, -\hat{z}_{t-1}^f, \ldots, -\hat{z}_{t-n_c}^f]^T \\
\hat{z}_t &= \hat{\varphi}_t^T \hat{\vartheta}_{t-1}, \qquad \dot{z}_t = \hat{\varphi}_t^f
\end{aligned}
$$

$$
\hat{\varepsilon}_t(\hat{\theta}) = \frac{1}{2} \begin{cases} \left|\hat{z}_t(\hat{\vartheta})\right|^2 - m_{2,t}^2, & \text{Case 1: Complex Signals} \\ \dfrac{1}{P_t} \displaystyle\sum_{p=0}^{P_t-1} \hat{z}_{t-p}^2(\hat{\vartheta}) - \sigma_{z,t}^2, & \text{Case 2: Real Signals} \\ \hat{z}_t^2(\hat{\vartheta}) - \sigma_{z,t}^2, & \text{Case 3: Simplified Real} \end{cases}
$$

$$
\dot{\varepsilon}_t = \begin{cases} Re\{\hat{z}_t \dot{z}_t\}, & \text{Case 1: Complex Signals} \\ \dfrac{1}{P_t} \displaystyle\sum_{p=0}^{P_t-1} \hat{z}_{t-p} \dot{z}_{t-p}, & \text{Case 2: Real Signals} \\ \hat{z}_t \dot{z}_t, & \text{Case 3: Simplified Real} \end{cases}
$$

$$
R_t = \lambda_t R_{t-1} + \dot{\varepsilon}_t \dot{\varepsilon}_t^T, \qquad \lambda_t = \frac{w_{t-1}}{w_t} \le 1
$$

$$
\begin{aligned}
\hat{\vartheta}_t &\triangleq \left[\hat{a}_t(0), \ldots, \hat{a}_t(n_a), \hat{b}_t(1), \ldots, \hat{b}_t(n_b), \hat{c}_t(1), \ldots, \hat{c}_t(n_c)\right]^T \\
&= \hat{\vartheta}_{t-1} - R_t^{-1} \dot{\varepsilon}_t \left[\hat{\varepsilon}_t + \dot{\varepsilon}_t\left(\hat{\vartheta}_{t-1} - \hat{\vartheta}_0\right)\right]
\end{aligned}
$$

$$\hat{C}_t(z) = \mathrm{Proj}(\hat{C}_t(z))$$

$$\begin{aligned}
\bar{z}_t &= \hat{\psi}_t^T \hat{\theta}_t, \qquad \bar{z}_t^f = \bar{z}_t - \hat{c}_t[1]\bar{z}_{t-1}^f - \cdots - \hat{c}_t[n_c]\bar{z}_{t-n_c}^f \\
u_t^f &= u_t - \hat{c}_t[1]u_{t-1}^f - \cdots - \hat{c}_t[n_c]u_{t-n_c}^f \\
\hat{z}_t^f &= \hat{z}_t - \hat{c}_t[1]\hat{z}_{t-1}^f - \cdots - \hat{c}_t[n_c]\hat{z}_{t-n_c}^f
\end{aligned} \tag{21}$$

Conclusions

The Constant Modulus Algorithm (CMA) for channel equalization has been extended in several ways. The extensions were aimed primarily at improving convergence rate and resistance to noise. The structural extensions include use of the recursive Gauss-Newton (RGN) method in place of gradient descent, a form of RGN which interpolates between batch and recursive forms, an interference canceling facility, and extension to moving-average channels.

We extended the CMA from a gradient descent to a Newton descent. This extends a first-order Talyor-series approximation of the error surface to a second-order one. However, the error minimized in this case is exactly described by a fourth-order Talyor series. A fourth-order type of descent algorithm should yield faster convergence. Note that all solutions to the third order "vector polynomial" for the gradient can be expressed in closed form.

In case 2, there is the possibility of adaptively tracking the instantaneous period of the carrier P_t. Period tracking is accomplished by adding P_t to the parameter vector θ and rederiving the error gradient. If the carrier frequency is large, the algorithm can easily track P_t. As a further refinement, the instantaneous period can be extended to non-integer values using the techniques discussed in [10].

In the case of multipath distortion, the channel model is of the form $1 + \gamma d^\tau$, where typically τ is many samples. It is unnecessary to compute all the parameters of a large-order inverse filter when tracking such a channel model. The zero coefficients can be eliminated from the model. A high-quality method for adaptively tracking the (interpolated) delay τ in this context is described in [10].

In digital modulation formats, the switching transients may be low-energy relative to the signal energy between switching instants. In order to avoid ill-conditioning due to this imbalance, equation (17) can be used to skip over intra-baud data. In other words, data acquisition can be arranged to occur only in a neighborhood of the switching times.

Due to the time-shift structure in the RGN algorithms we have derived, there exist so-called "fast" versions which require order n computations instead of n^2 as in the present versions. A point of departure from which such an extension is straightforward is given in [2].

References

[1] B. Widrow, et. al., "Adaptive Noise Canceiling: Principles and Applications," *Proc. IEEE*, vol. 3, pp. 1692-1716, Dec. 1975.

[2] M. Morf, L. Ljung and T. Kailath, "Fast Algorithms for Recursive Identification," Proc. IEEE Conf. on Decision and Control, Clearwater Beach, Florida, 1976.

[3] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, New York, 1981.

[4] J. R. Treichler and B. C. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. on Acoust., Speech, and Signal Proc.*, vol. ASSP-31, pp. 459-472, April 1983.

[5] J. O. Smith, "Techniques for Digital Filter Design and System Identification with Application to the Violin," Ph.D. Dissertation, Elec. Eng. Dept., Stanford University, June 1983.

[6] L. Ljung and T. L. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge MA, 1983.

[7] M. G. Larimore and J. R. Treichler, "Convergence Behavior of the Constant Modulus Algorithm," *ICASSP-83: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 1.4, Boston MA, March 19-21, 1984.

[8] J. R. Treichler and M. G. Larimore, "A Real-Arithmetic Implementation of the Constant Modulus Algorithm," *ICASSP-84: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 3.2, San Diego CA, March 19-21, 1984.

[9] J. R. Treichler, "Algorithms that Restore Signal Properties," *ICASSP-84: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 21.4, San Diego CA, March 19-21, 1984.

[10] J. O. Smith and B. Friedlander, "Adaptive Interpolated Time-Delay Estimation," Asilomar, 1983. Full version to appear in the *IEEE Trans. on Aerospace*, May 1985.

[11] J. O. Smith and B. Friedlander, "Analysis and Extensions of the Constant Modulus Algorithm," Tech. Rep. 6502-01, Systems Control Tech., Oct. 1984.

[12] J. R. Treichler and M. G. Larimore, "New Processing Techniques Based on the Constant Modulus Adaptive Algorithm," to appear.

APPENDIX C

WAVEGUIDE DIGITAL FILTERS

Waveguide Digital Filters

Julius O. Smith
Benjamin Friedlander

Systems Control Technology, Inc.
1801 Page Mill Road
Palo Alto, California, 94303

Abstract

The set of finite-order, linear, time-invariant filters is derived by sampling lossless propagation through a variable-impedance medium. This leads to a flexible class of *time-varying* filter structures, termed "Waveguide Filters" (WGF) in which signal power is decoupled from changes in the filter parameters. These structures are "balanced" in the sense that the decoupling between signal power and time-varying filter coefficients is maintained *for* each individual section in the structure. In addition, limit cycles and overflow oscillations are suppressed, even in the time-varying case, when implemented with "passive" arithmetic. We describe also a method for enforcing exact losslessness in the realization of an arbitrary digital filter in spite of the presence of round-off errors. Finally, the WGF structures can be interconnected in series or in parallel in a way which does not disturb the signal/coefficient decoupling or the power balance. Thus, the waveguide filters are very useful for modeling physical systems, and the exactness of their physical interpretation enhances their suitability for the time-varying case. All results are obtained for the multi-input/multi-output case.

1. Introduction

Digital filtering techniques have often been derived from classical or "analog" techniques [29]. Classical filter design has its roots in "network theory" for describing linear time-invariant systems accessed by means of "ports" [11]. Network theory itself is a body of mathematics built upon certain assumptions [7,13] which become true in the limit at low-frequencies according to Maxwell's equations for electromagnetic propagation [9]. Thus, the theory of filters grew originally out of the scalar theory of wave propagation.

Since the emergence of digital techniques, little attention has been paid to the close correspondence between filter computations and physical law. In signal processing applications, we normally approximate directly some desired transformation of the signal spectrum, and a true physical modeling is irrelevant.

The mainstream of filtering applications has involved time-invariant filters which approximate an ideal *amplitude response* such as low-pass, high-pass, band-pass, or band-reject characteristics, or which provide a desired *phase response* such as in equalizers for communications channels [29]. In the time-invariant case, the amplitude response and phase response completely determine a linear filter [29].

For time-varying filters, there is no longer a simple description in terms of amplitude and phase response. (A frequency response requires time-invariance.) In many cases, time-varying filters have been developed in an ad hoc manner, being regarded as "quasi-static" in most cases. Such extensions require the assumption that the filter coefficients vary slowly relative to the impulse-response duration of the filter. When the coefficients change too rapidly, unnatural artifacts can occur due to the incompatibility between the filter state (a function of all prior time in a recursive filter) and the new filter coefficients.

This Paper

This paper presents a class of recursive digital filters designed specifically to have the best possible behavior under time-varying conditions in the presence of round-off error. We call them "Waveguide Filters" (WGF), because they can be interpreted as networks of intersecting waveguides. The WGF structures are closely related to the "wave digital filters" developed principally by Fettweiss [17,28], the lattice filter structures arising in geoscience and speech modeling [45,31], and the "normalized ladder filter" discussed by Gray [30,39]. Waveguide filters have the following characteristics:

- The correspondence to physical wave-propagation systems is exact even though time is discrete. No bilinear transformation is necessary to connect digital quantities with physical quantities as is usual in the wave digital filter (WDF) context [17]. This allows a priori choice of filter structure to obtain precise models for physical processes.

- The instantaneous power at each internal filter section is invariant with respect to filter coefficient variation.

- Generalized versions of the "Normalized Ladder," "One-Multiplier Lattice," and other ladder/lattice filters are derived, all having invariant instantaneous power in the time-varying case.

- The structures can be coupled at a junction, cascaded, looped, or branched, to any degree of network complexity, and the desirable properties such as stability and power decoupling are retained.

- A synthesis procedure exists for computing all-pole or pole-zero sections.

- There is an identification method for determining the coefficients of the structure from measured input/output data. Similarly, there is a "linear prediction" modeling technique for these structures which provides ARMA models for time series.

- No overflow oscillations can occur, even in the time-varying case.

- No limit cycles (also called "granularity oscillations") can occur if one of many "passive" numerical round-off strategies is employed, even in the time-varying case. In the simplest case, the passive round-off strategy reduces to magnitude truncation (or truncation toward zero).

- As in the scalar lattice filter and WDF cases, sensitivity of coefficient quantization can be minimized by properly scaling the network to deliver "maximum power transfer" at frequencies where low sensitivity is required [50].

- A perfectly lossless digital realization can be implemented using a number system presented in this paper.

- The desirable structural properties are derived for multi-input, multi-output (MIMO) transfer-function matrices.

The derivation of the WGF is made exceedingly simple by using three simple principles of wave propagation in an ideal linear medium. To our knowledge, these principles have not been invoked before to derive digital filters. In this respect, we feel this paper has significant tutorial value. It is a new point of view.

2. *Related Prior Work*

This section reviews some of the most closely related work on digital filter structures. These include the orthogonal-polynomial filters of Szegö, the "wave digital filters" of Fettweis, and the "orthogonal filters" of Dewilde. Naturally, there are many more related lines of development, in view of the first law of signal processing.* These represent only the major recent areas closest to our point of view.

2.1. Wave Digital Filters

The wave digital filter (WDF) approach of Fettweis [16,17,20,28] comes closest to the point of view taken in this paper. Fettweis obtains a similar class of structures by use of the classical notion of *wave variables* [13].

* The first law of signal processing is "Everything is equivalent to everything else."

For example, if v and i denote the voltage and current at a terminal of an N-port network, the wave variables are defined by $x = v + Ri$ and $y = v - Ri$, where R is an arbitrary "reference impedance." These wave variables are logically equivalent to the left-going and right-going "pressure traveling waves" considered in this paper, and R plays the role of characteristic impedance in the associated transmission line. A generalization of wave variables to the form $x = \alpha v + \beta i$, $y = \gamma v + \delta i$ and a characterization of the specialization necessary to ensure realizability is given in [32].

Fettweis describes how to directly model resistors, capacitors, inductors, transformers, gyrators, and circulators using the WDF approach, and he describes the necessary rules for connecting ports together [17]. The modeling of a capacitor, for example, is accomplished by scaling the reference impedance R until the capacitor "reflectance" is exactly a unit-sample delay. (The model is parametrized in frequency so that the wave variables are really phasors.) An inductor also maps to a unit-sample delay but with a sign-change relative to a capacitor. A complete circuit is built out of basic elements by means of "*adaptors*" [28] which play the role of the junctions or scattering layers described in this paper; the adaptor accomplishes interconnection of ports at different reference impedances.

The WDF modeling of inductors and capacitors is limited because the continuous-time frequency variable is mapped to the discrete-time frequency variable via the bilinear transform. If the points $z = 1$ and $z = -1$ in the complex plane are identified with zero and infinite continuous-time frequencies, respectively, then only one more mapping frequency (say ψ) can be chosen. Thus, the bilinear transformation provides exact modeling only at the three frequencies 0, ψ, and ∞.

The WDF formulation models a system of differential equations at three frequencies, while our approach exactly models wave propagation in lossless media having spatially discrete changes in characteristic impedance. Consequently, in our formulation, a wave variable may be a "voltage" or "current" or a linear combination of the two without incurring realizability problems [32]. This is a considerable conceptual simplification for applications to physical modeling.

A general result in this paper is that overflow oscillations and limit cycles can be suppressed in *all forms* of scattering-type filter structures simply by using extended numerical precision in each scattering section, saving quantization (toward zero) for the final outgoing waves. The basic principles involved apparently appeared first in the WDF context [26,39].

2.2. Ladder and Lattice Filters

For some time it has been known that lattice and ladder filtering structures are superior to the so-called direct form in several ways. These include reduced sensitivity to coefficient quantization, less dependency of round-off noise on the filter frequency response, ease of stability checking, reduced probability of limit cycles or overflow oscillations, and section-wise orthogonality in the linear prediction context. For a discussion of ladder and lattice filters in adaptive estimation, see [45].

Lattice structures have been in use for decades in directly modeling layered scattering media. The mapping of underground striations in rock density, for example, is a basic diagnostic tool in oil exploration. The interface between two subterranean layers of rock of different densities produces a scattering layer because the characteristic impedance of the medium with respect to sound propagation changes across such a boundary.

Another example of the use of lattice structures for physical modeling is the "acoustic tube" models developed for speech analysis and synthesis. In this case, the vocal tract is modeled as a cascade of coaxial cylindrical tubes with varying cross-sectional areas and equal length. The change in area from one tube section to the next provides a change in the characteristic impedance of the air column for sound propagation, and so a series of equally spaced scattering layers is obtained.

Apparently, the filter structures developed in the above applications are only as general as a single chain of scattering layers with one input and one output, and the input and output sections are terminated in a non-extendable way. Little if any work has explored branching and intersecting chains of scattering layers. In the case of speech, the use of a separate acoustic tube branching off from the vocal tract to model the nasal tract would obviously be very natural. Apart from branching, it is not possible to continue the structures in common use from the output section to a larger section. This is because the typical arrangement is to assume a perfectly reflecting termination at the output. Doing this allows manipulation of the delays in the scattering network to place them more conveniently and combine in pairs such that the required signal sampling rate is reduced by a factor of 2. We have found that the cascade scattering chains, which dominate the recent literature, can be immediately extended to general acyclic trees with the same basic properties.

Our formulation is more general than even the acyclic-tree extension of prevalent lattice filters in that arbitrary networks can be constructed. Also, there does not seem to be an existing treatment of MIMO systems from the acoustic waveguide point of view, nor the generalization which allows transmission zeros to be an integral part of the waveguide (without having to add external "taps" for forming a linear combination of the contents of each waveguide section).

A particularly important antecedent to the WGF in the speech processing literature is the normalized ladder filter (NLF) developed by Gray and Markel [23,30,39]. Gray considered only the single-input, single-output (SISO) all-pole case. (Zeros are obtained in the NLF using "taps," which leads outside the class of structures considered here.) Their approach was based on orthonormal-polynomial expansion [1,2,6] which is closely connected with *linear prediction* theory [31]. They showed the following to be true:

- The NLF is optimal in the sense that each internal node has unity power gain. This means, for example, that the response to a unit impulse cannot overflow anywhere within a stable NLF filter. Also, if the input signal is white noise with unit variance, the variance of the signal at each internal node is exactly unity [30].

- The NLF is stable in the case of time-varying filter parameters [30] as long as the "reflection coefficients" $k_i(t)$ are always less than or equal to some $K<1$ in magnitude. ($|k_i(t)| < 1$ is not sufficient for bounded-input, bounded-output (BIBO) stability unless the input signal energy is finite.) It was derived incidentally that the total energy entering the ladder eventually "exits" through the particular delay element at the entrance to the ladder.

- The NLF has superior roundoff noise properties, especially when poles are clustered close together and/or close to the unit circle [30].
- The NLF is *free of zero-input overflow oscillations* [39].
- The NLF is free of zero-input limit cycles [39] in magnitude-truncation arithmetic.

The NLF is obtainable by transformations of a special case of the WGF structures derived here. The most significant difference is in the distribution of delay elements. We will show that delay distribution in the standard NLF is not obtainable from a WGF unless the waveguide is terminated by a pure reflection. This means, for example, that an NLF cannot be connected to another NLF to build a larger waveguide system with finite loading from one stage to the next. Also, the delay distribution chosen for the NLF is such that creating a loop with NLF's yields a degenerate (non-computable) structure because a delay-free loop appears. Another limitation of the NLF is that the concept of instantaneous power becomes artificial for individual sections (although Gray defines a non-physical but similar quantity in [30, eq. (2)]).

A disadvantage of the NLF is that it requires four multiplications per pole of the filter transfer function.* The one-multiplier lattice filter, on the other hand [31] requires only one multiplication per pole. Because of the choice of delay distribution in standard lattice-filter theory, only the NLF has been shown to be power-preserving in some sense. In contrast, we will show that even our counterpart to the one-multiplier lattice can be made normalized with respect to time-varying

* *As a side result, we show that one of these four multiplications can be eliminated.*

coefficients. Conceptually, this is achieved by compensating the amplitude of stored signal samples. The resulting normalized one-multiplier lattice section is computationally less expensive than the NLF.

The reason that most standard ladder and lattice structures (all but the NLF) cannot be power-normalized in the time-varying case is that the unnatural distribution of delays adopted makes *passivity* of a section nontrivial to show. This paper describes how power-normalization, perfect energy conservation, and complete suppression of limit cycles and overflow oscillations can be guaranteed for MIMO analogues of *all* ladder and lattice filter structures, with extensions to branching structures and general terminations.

For the case of reflectively terminated, time-varying, MIMO, acyclic trees, (which specialize to ordinary lattice/ladder structures in the SISO single-branch case), we derive efficient equivalent structures in which the delays are moved and combined to yield computational savings without loss of the desired power-invariance/numerical properties.

2.3. Synthesis and Approximation

The synthesis procedure we use for the WGF is based on the Schur algorithm which recursively computes a solution to the Nevanlinna-Pick problem [40,37,43]. The Nevanlinna-Pick problem is to interpolate a rational "Schur function*" through n complex values at n points in the closed unit disk in the complex plane. The Schur algorithm has also been called the "Nevanlinna recursion scheme" [43]. In other contexts, a special case of the the Schur algorithm, which computes only all-pole digital filters, has been called the "Durbin" [8] or "Levinson" [3] algorithm [34,40,42,38,31]. The complete Schur algorithm constructs a cascade WGF realization of a digital filter containing both poles and zeros.

* A *Schur function* $S(z)$ is defined as a complex function analytic and of modulus not exceeding unity in $|z| \leq 1$

The estimation problem has been addressed by DeWilde [40,42]. In this context, the Schur algorithm provides an ARMA estimation technique in which the pole estimates are optimal in the mean square sense for the given fixed zeros which are chosen a priori.

3. Traveling Waves and Lossless Scattering

For concreteness of discussion, we will focus on *pressure* and *flow waves* in a so-called *acoustic tube*. We could just as easily think of the electric and magnetic components of light, voltage and current in a transmission line, or force and transverse velocity on a vibrating string. An analysis of the acoustic tube is discussed by Markel and Gray [31] and Flanagan [19] in the context of vocal-tract modeling. Further details on the acoustics of sound in tubes can be found in Morse [4]. The term "waveguide" will be used interchangeably with "acoustic tube."

A derivation of traveling waves from the basic *wave equation* is presented in the appendix. The result is that in a cylindrical acoustic tube, longitudinal* pressure and flow waves propagate back and forth with speed c. Let x denote distance along the tube axis and let t denote time in seconds. Then the instantaneous pressure $P(x,t)$ and flow $U(x,t)$ is given by the sum of the left-going and right-going traveling-wave components:

$$P(x,t) = P^{+}(x,t) + P^{-}(x,t) \tag{1a}$$

$$U(x,t) = U^{+}(x,t) + U^{-}(x,t) \tag{1b}$$

* We assume the tube radius is much smaller than the wavelength of sound in the tube, so that pressure and flow are constant over any cross-section of the tube normal to the axis. In other words, waves do not propagate up and down but only left and right. For more details on the assumptions involved in acoustic tube models, see Flanagan [19].

3.1. Three Fundamental Constraints

The behavior of waves traveling unidirectionally in a lossless medium is governed by three laws: (1) the pressure is proportional to flow, (2) the pressure is a continuous function of position, and (3) the flow variable (e.g. mass or charge) is neither created nor destroyed in the medium.

Characteristic Impedance

An ideal linear propagation medium is completely determined by its characteristic impedance† $Z(x,t)$. The ***characteristic impedance*** is defined the constant of proportionality between pressure and flow in a unidirectional traveling wave:

$$P^+ = ZU^+ \tag{2a}$$

$$P^- = -ZU^- \tag{2b}$$

When the arguments (x,t) are omitted, it is understood that all quantities are written for some constant time t and position x. The minus sign for the left-going

† For an acoustic tube, the characteristic impedance is given by $Z = \sqrt{P_0\gamma_c\rho}/S = \rho c/S$, where ρ is the density (mass per unit volume) of air in the tube, c is the speed of propagation, P_0 is ambient pressure, γ_c is the ratio of the specific heat of air at constant pressure to that at constant volume, and S is the cross-sectional area of the tube. In a vibrating string, $Z = \sqrt{T\rho} = \rho c$, where ρ is string density (mass per unit length) and T is the tension of the string. In an electric transmission line, $Z = \sqrt{L/C} = Lc$ where L and C are the inductance and capacitance, respectively, per unit length along the transmission line. In free space, $Z = \sqrt{\mu_0/\epsilon_0} = \mu_0 c$, where μ_0 and ϵ_0 are the permeability and permittivity, respectively, of free space.

wave P^- accounts for the fact that flows in opposite directions subtract while pressure waves passing through each other add.

We will consider initially a more general situation in which $Z = Z(d)$ is a q by q complex matrix function of the complex variable (or unit-delay operator) d. For stability of propagation in the waveguide, we require that $Z(d)$ be analytic for $|d| \leq 1$. The results also extend to the case of vector $\mathbf{d}^T = [d_1, \ldots, d_K]$, but we will treat only one complex argument d for notational simplicity. The pressure and flow variables are q by m matrix complex functions of d. However, keep in mind that the physical analogy we are pursuing is for the case of real scalar Z, P, and U.

For lossless propagation in the scalar case, the characteristic impedance Z must be real. In the matrix-delay-operator case, lossless propagation will now be characterized by the requirement that Z be *para-Hermitian,* i.e.,

$$Z_*(d) = Z(d) \tag{3}$$

where

$$Z_*(d) \triangleq \overline{Z(1/\overline{d})}^T \tag{4}$$

denotes the *para-Hermitian conjugate* of $Z(d)$ [13,40], $(\cdot)^T$ denotes transposition, and $\overline{(\cdot)}$ denotes complex conjugation. For $d = e^{j\theta}$, $Z_*(e^{j\theta})$ coincides with the Hermitian transpose of $Z(e^{j\theta})$. The para-Hermitian conjugate is the unique analytic continuation (when it exists) of the Hermitian transpose $Z_*(e^{j\theta}) = \overline{Z(e^{j\theta})}^T$ from the unit circle into the complex plane. Thus, a lossless medium in our framework is defined as one in which the characteristic impedance is para-Hermitian. The extension to vector $\mathbf{d}$ is obtained by regarding $Z(\mathbf{d})$ as K functions of scalar complex variables d_i. Note that in the scalar case, Z para-Hermitian implies $Z = \overline{Z}$ which implies Z is real. Henceforth, we assume Z denotes a para-Hermitian characteristic impedance. For non-para-Hermitian Z, (2) should be modified to read $P^- = -Z_*U^-$ [13], and a passive medium is one in which $Z + Z_*$ is positive semi-definite.

It is worthwhile to interpret the various levels of extension we are considering for the characteristic impedance Z. When Z is real and scalar, we obtain exactly the ideal behavior of one dimensional traveling waves in a lossless medium. Extending to q by q matrix characteristic impedances facilitates development of multi-input, multi-output (MIMO) systems which have the desired numerical and power-invariance properties. The extension to analytic matrix functions of a complex variable provides a generalized scattering medium whose reflectance and transmission coefficients are themselves rational transfer function matrices. This provides for nesting of the WGF structures. The complex argument d of the characteristic impedance is interpreted as a unit-delay operator, and the meaning of the characteristic impedance is attached to its Laurent series expansion with respect to the unit circle in the d-plane. Additional complex variables d_i in the arguments to the characteristic impedance allow the generalized scattering layer to perform filtering in several domains such as time and space. Since the characteristic impedance is assumed stable and para-Hermitian, all delay-operator impedance matrices must be nonrecursive and zero-phase. Therefore, computability, stability, and nonlinear oscillation problems do not arise in the case of multiple domains.

Pressure Continuity and Medium Conservation

We will be interested in the situation wherein the characteristic impedance changes abruptly from one value to another, say from Z_1 to Z_2. The impedance discontinuity can be a sudden change along x in the acoustic tube, or it can be a change introduced at some time t (as needed for time-varying filters). First we consider changes with respect to x. Given the traveling waves impinging on the junction between Z_1 and Z_2, we seek formulas for the traveling waves leaving the junction. To solve this problem, we need two laws in addition to (2) for an ideal wave medium:

1) Pressure cannot change instantaneously across the junction (5a)

2) The sum of flows meeting at the junction is zero (5b)

These constraints are often called "Kirchoff's node equations" in the context of circuit theory. For changes in characteristic impedance with respect to time, (5) is not applicable. Time-varying characteristic impedances will be implemented as waveguide *transformers*, and will be used to obtain power-invariant lossless digital filters in the time-varying case.

The continuity and conservation constraints (5) together with the characteristic impedance constraints (2) determine what pressure and flow waves emerge from a junction between waveguide media of differing characteristic impedance, given the incoming waves.

Consider the case of N waveguides meeting at a common junction. Kirchoff's laws state that there can be only one resultant pressure P_J at the junction, and the sum of flows entering and leaving the junction must total to zero. Thus, we have the constraints

$$P_1 = P_2 = \cdots = P_N = P_J \tag{6a}$$

$$U_1 + U_2 + \cdots + U_N = 0 \tag{6b}$$

where

$$\begin{aligned} P_i &= P_i^+ + P_i^- \qquad & P_i^+ &= Z_i U_i^+ \\ U_i &= U_i^+ + U_i^- \qquad & P_i^- &= -Z_i U_i^+ \end{aligned} \tag{7}$$

Z_i = Characteristic impedance of the ith waveguide (q by q)

$\Gamma_i = Z_i^{-1}$ = Characteristic admittance of the ith waveguide (q by q)

P_i^+ = Incoming pressure wave along the ith waveguide (q by m)
U_i^+ = Incoming flow wave along the ith waveguide (q by m) (8)

P_i^- = Outgoing pressure wave along the ith waveguide (q by m)
U_i^- = Outgoing flow wave along the ith waveguide (q by m)
P_i = Instantaneous pressure wave in ith waveguide just outside junction (q by m)
U_i = Instantaneous pressure wave in ith waveguide just outside junction (q by m)
P_J = Resultant pressure in the junction (q by m)

(9)

3.2. Reflection at a Junction

Given a set of incoming traveling pressure waves, P_i^+, $i = 1, \ldots, N$, the constraints (2,6) determine the outgoing waves as follows. As before, Z_i and hence $\Gamma_i = Z_i^{-1}$ are para-Hermitian positive definite. Substituting (1b) and (2) into (6b) yields the resultant junction pressure:

$$P_J = 2\left(\sum_{i=1}^{N} \Gamma_i\right)^{-1} \sum_{i=1}^{N} \Gamma_i P_i^+ \tag{10}$$

Let

$$\Gamma_J \triangleq \sum_{i=1}^{N} \Gamma_i \qquad Z_J \triangleq \Gamma_J^{-1} \qquad U_J \triangleq 2\sum_{i=1}^{N} \Gamma_i P_i^+ \tag{11}$$

define the *junction admittance, junction impedance*, and *junction flow*, respectively. (In the extension to non-para-Hermitian Z_i, U_J becomes $U_J^+ = \sum(\Gamma_i + \Gamma_{i*})P_i^+$.) Relation (10) then reads $P_J = Z_J U_J$, or,

Junction Pressure = Junction Impedance $\times$ Junction Flow

Since $\Gamma_i P_i^+ = U_i^+$, we have $U_J = 2\sum_{i=1}^{N} U_i^+ \triangleq 2U^+ \Rightarrow |U_J| = |U^+| + |U^-|$, where $|\cdot|$ denotes elementwise complex modulus. That is, the junction flow can be interpreted as the magnitude sum of the incoming and outgoing current waves.

Now, given incoming traveling waves P_i^+, U_i^+ and the characteristic impedance Z_i of each branch terminating at the junction, we easily find the outgoing waves P_i^-, U_i^- to be

$$P_i^- = P_J - P_i^+ \tag{12a}$$
$$U_i^- = \Gamma_i P_i^- \tag{12b}$$

Equations (12) specify the scattering at the junction of N intersecting "wave guides," given the incoming waves P_i^+ (or U_i^+) and the branch characteristic impedances Z_i.

In view of relations (2), we can consider only left-going and right-going pressure waves, since the flow waves can be readily computed from the characteristic impedance of the propagation medium. At this point, we could instead choose wave variables of the form $x_i = P_i + Z_i U_i$ and $y_i = P_i - Z_i U_i$ and proceed along the lines of classical wave filtering [13]. However, such a path is less fundamental in the present development because we are considering only discrete-time filters.

We have treated only a *parallel junction* of waveguides. A *dual* set of equations is obtained by considering a *series junction*. However, pressure waves intersecting in a parallel junction are equivalent to flow waves intersecting at a series junction. When using flow waves as the primary variables, (12b) can be written

$$U_i^- = U_i^+ - \Gamma_i P_J \tag{13a}$$

$$P_J = 2Z_J \sum_{i=1}^{N} U_i^+ \tag{13b}$$

The series pressure junction is obtained by taking the dual of (13). That is, replace U_i by P_i and Γ_i by Z_i to obtain

$$P_i^- = P_i^+ - Z_i U_J^s \tag{14a}$$

$$U_J^s = 2\Gamma_J^s \sum_{i=1}^{N} P_i^+ \tag{14b}$$

$$\Gamma_J^s = {Z_J^s}^{-1} \tag{14c}$$

$$Z_J^s = \sum_{i=1}^{N} Z_i \tag{14d}$$

The junction impedance for a series junction is the sum of the branch impedances, while for a parallel junction, it is the parallel combination of the branch impedances (inverse of the sum of admittances).

Equation (12a) is a *computationally efficient* way to implement an N-port scattering junction. In the case $N = 2$, the well-known one-multiplier lattice filter section (minus its unit delay) is obtained immediately from (12a). More generally, an N-way intersection requires N multiplies and $N-1$ additions to obtain P_J, and one addition for each outgoing wave, for a total of N multiplies and $2N-1$ additions. The dual junction (14) also requires N multiplies and $2N-1$ additions. In the next section, a method for trading one multiplier for another $N-1$ additions [28] is described.

3.3. α-parameters

One parametrization of all passive N-junctions is the set of N branch impedances with positive-definite para-Hermitian parts (cf. §5.1). This section describes another parametrization, analogous to that used in the WDF context [28].

One parametrization of all passive N-junctions is the set of N branch impedances with positive-definite para-Hermitian parts. This section describes another parametrization, analogous to that used in the WDF context [28].

Define

$$\alpha_i = 2Z_J\Gamma_i \tag{15}$$

which is twice the junction impedance times the ith branch admittance. (In the non-para-Hermitian case, $\alpha_i = Z_J(\Gamma_i + \Gamma_{i*})$.) Then the junction pressure can be written as a linear combination of the incoming pressure waves in terms of the α_i as

$$P_J = \sum_{i=1}^{N} \alpha_i P_i^+ \tag{16}$$

Since $\sum_{i=1}^{N} \Gamma_i \triangleq \Gamma_J$,

$$\sum_{i=1}^{N} \alpha_i = 2I_q \tag{17}$$

where I_q is the q by q identity matrix.

In matrix form, (12a) can be written

$$\begin{bmatrix} P_1^- \\ P_2^- \\ \vdots \\ P_N^- \end{bmatrix} = \begin{bmatrix} \alpha_1 - I_q & \alpha_2 & \dots & \alpha_N \\ \alpha_1 & \alpha_2 - I_q & \dots & \alpha_N \\ \vdots & \vdots & & \vdots \\ \alpha_1 & \alpha_2 & \cdots & \alpha_N - I_q \end{bmatrix} \begin{bmatrix} P_1^+ \\ P_2^+ \\ \vdots \\ P_N^+ \end{bmatrix} \tag{18}$$

or

$$\underline{P}^- = \Sigma \underline{P}^+ \tag{19}$$

where

$$\Sigma \triangleq A - I_{Nq}, \qquad A \triangleq \begin{bmatrix} I_q \\ I_q \\ \vdots \\ I_q \end{bmatrix} [\alpha_1 \, \alpha_2 \dots \alpha_N] \tag{20}$$

The matrix Σ is called the *scattering matrix* of the junction. Since $P_i^- = (\alpha_i - I_q)P_i^+$ when $P_j^+ = 0$ for all $j \neq i$, we define the *reflection coefficient* at the ith port by

$$\rho_i \triangleq \alpha_i - I_q \tag{21}$$

Equations (12a,16,17) combine to give

$$P_i^- = P_i^+ + \sum_{\substack{j=1 \\ j \neq i}}^{N} \alpha_j (P_j^+ - P_i^+) \tag{22}$$

Thus, α_j can be interpreted as the fraction of the pressure differential between branches j and i which is reflected back along the ith branch with P_i^+, for any i. Use of this expression saves one matrix multiply but entails $3N-2$ matrix additions. If one multiply is worth $N-1$ additions or more, then (22) is less expensive to implement than (12a).

3.4. Pure Delays

Up to now we have been concerned only with the scattering of traveling waves at a single impedance-discontinuity junction. We now allow for many such junctions interconnected by waveguides for lossless, reflectionless propagation. Physically, an interconnection between junctions is a length of material at a single characteristic impedance. It is implemented digitally using *bi-directional delay lines*.

Consider the interconnection of two N-port junctions. Between the two junctions is a section of pure waveguide which is a lossless medium having characteristic impedance Z_{12}. Let c denote the speed of propagation in this waveguide section, and suppose the distance between the junctions is L. Then the propagation time from one junction to the other and back is $T_s = 2L/c$. Consider a pressure wave impulse $P^+(x-ct) = \delta(x-ct)$ traveling from junction 1 to junction 2 starting at time $t=0$. At time $T_s/2$ it reaches junction 2, and a reflection $P^-(x+ct) = \rho_{12}\delta[x+c(t-T_s)]$ starts out to the left from junction 2, heading back to junction 1. A fragment of the pulse is also sent out along all waveguides connected to junction 2, according to the relative branch impedances. At time $t = T_s$, the reflected pulse reaches junction 1, and scatters away again.

A section of waveguide joining two junctions by a propagation delay $T_s/2$ is called a *unit delay*. If the speed of propagation is everywhere the same, then all unit delays are of the same length $L = cT_s/2$.

An important observation to make at this point is that the impulse response at a particular junction of a network of junctions interconnected by waveguides of length $cT_s/2$ is nonzero only at times which are integer multiples of T_s. Thus, such a network is a simulation of a digital linear system with sampling rate $F_s = 1/T_s$.

Consider a linear chain of junctions. The input is defined as the pressure entering at the far left (junction 0), and the output is defined as the pressure emerging from the far right (junction M). Again, this structure is an exact simulation of a digital system with sampling rate $F_s = 1/T_s$ provided M is even. If M is odd, we obtain samples separated by T_s seconds at the output, but there is a time shift of $T_s/2$ relative to the sample instants at the input.

Now consider the more elaborate interconnection of junctions in an arbitrary network. Unless every path from the input to the output crosses an even number of unit delays, the impulse response of the network will be nonzero at multiples of $T_s/2$, yet the minimum delay in any feedback loop is T_s seconds. Thus, it is not possible to simulate a general digital system having sampling rate $2/T_s$ even though such a sampling rate is required to compute the response.

Two cases arise:

(1) *Half-rate structures* which require an even number of branches on every path from the input to the output, plus possibly a single odd section which causes a half-sample shift of the output relative to the input. In the latter case, we do not allow the resulting structure to be placed in a feedback loop. These restrictions leave us with a general class of linear digital filter structures.

(2) *Full-rate structures* in which the most general transfer function between two junctions is of the form

$$H(z) = \frac{b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}}{1 + 0 + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}}. \tag{23}$$

in the scalar case, where $n_a \geq 2$ and $n_b \geq 1$ are integers. In the full-rate case we must settle for the class of rational filters in which $a_1 = 0$. The corresponding digital filter design methods must support this constraint. Therefore, techniques based on *linear programming* seem natural in this context [36]. In the estimation context, the methods described in [49] can be adapted to this problem.

4. Cascade Waveguides

We now specialize discussion to the case of cascade waveguide sections. The junction between two waveguides of differing characteristic impedance will create a scattering junction. The stretch of pure waveguide material between scattering junctions will provide delay lines for the propagation signal. From these structures all digital filters can be built in such a way that they behave nicely with respect to time-varying parameters and numerical roundoff/overflow. Note that almost all special properties in the cascade case carry over to arbitrary acyclic trees.

4.1. The Two-Port Junction

If there are only two waveguides meeting at a junction, we obtain the classical "scattering theory" in which an incoming wave is split into a "reflected" and "transmitted" part. From (15), the α-parameters are

$$\begin{aligned} \alpha_1 &= 2(Z_1^{-1} + Z_2^{-1})^{-1} Z_1^{-1} = 2(\Gamma_1 + \Gamma_2)^{-1}\Gamma_1 = 2(I_q + Z_1\Gamma_2) \\ \alpha_2 &= 2(Z_1^{-1} + Z_2^{-1})^{-1} Z_2^{-1} = 2(\Gamma_1 + \Gamma_2)^{-1}\Gamma_2 = 2(I_q + Z_2\Gamma_1) \end{aligned} \tag{24}$$

From (22), the reflected pressure waves are

$$\begin{aligned} P_1^- &= \alpha_2\left(P_2^+ - P_1^+\right) + P_1^+ = (\alpha_1 - I_q)P_1^+ + \alpha_2 P_2^+ \\ P_2^- &= \alpha_1\left(P_1^+ - P_2^+\right) + P_2^+ = \alpha_1 P_1^+ + (\alpha_2 - I_q)P_2^+ \end{aligned} \tag{25}$$

If $P_2^+ = 0$, then the incidence of P_1^+ produces a reflected wave $P_1^- = (\alpha_1 - I_q)P_1^+$. Thus, we define the following *reflection coefficients:*

$$\begin{aligned} \rho_1 &\triangleq \alpha_1 - I_q = (Z_2 - Z_1)(Z_2 + Z_1)^{-1} = (\Gamma_1 + \Gamma_2)^{-1}(\Gamma_1 - \Gamma_2) \\ \rho_2 &\triangleq \alpha_2 - I_q = -\rho_1 \end{aligned} \tag{26}$$

It is now apparent that if the reflection coefficient at port 1 is $\rho_1 = \rho$, then at port 2 the reflection coefficient is $-\rho$. Another point of view is that inverting the impedance-step ratio from $Z_1{:}Z_2$ to $Z_2{:}Z_1$ merely changes the sign of the reflection coefficient. It is easy to show that for scalar Z_i, exchanging pressure waves for flow waves also toggles the sign of all reflection coefficients in the network. Thus, left is the dual of right as pressure is the dual of flow(and parallel is the dual of series).

In the matrix-impedance case ($q > 1$), however, replacement of pressure by flow changes ρ to $-(\Gamma_1 Z_2 - I_q)(\Gamma_1 Z_2 + I_q)^{-1}$ which equals $-\rho$ only if Z_1 commutes with Z_2. Two Hermitian matrices commute if they have the same eigenvectors, i.e., their "principal axes of dilation" are aligned. There exists a unitary transformation of any Hermitian matrix which commutes with any other Hermitian matrix; that is, a Hermitian matrix can be "rotated" until it commutes with any other Hermitian matrix. For waveguides with impedance matrices so aligned (all Z_i have the same eigenvectors on the unit circle), junction reversal is equivalent to wave-variable exchange; either causes a sign reversal in all reflection coefficients.

4.2. The Scattering Matrix

In block-matrix notation, the junction output is given by

$$\begin{bmatrix} P_1^- \\ P_2^- \end{bmatrix} = \begin{bmatrix} \alpha_1 - I_q & \alpha_1 \\ \alpha_2 & \alpha_2 - I_q \end{bmatrix} \begin{bmatrix} P_1^+ \\ P_2^+ \end{bmatrix} = \begin{bmatrix} \rho_1 & I_q + \rho_1 \\ I_q - \rho_1 & -\rho_1 \end{bmatrix} \begin{bmatrix} P_1^+ \\ P_2^+ \end{bmatrix} \triangleq \Sigma \underline{P}^+ \tag{27}$$

or

$$\underline{P}^- = \Sigma \underline{P}^+ \tag{28}$$

This is called the *scattering formulation*, and Σ is called the *scattering matrix*. It is a special case of the N-junction scattering matrix defined in (18).

4.3. The Chain-Scattering Matrix

In the two-port case only, we can also define the *chain-scattering matrix* via

$$\begin{bmatrix} P_1^+ \\ P_1^- \end{bmatrix} = \Theta \begin{bmatrix} P_2^- \\ P_2^+ \end{bmatrix} \tag{29}$$

where

$$\Theta \triangleq \begin{bmatrix} \Theta_{11} & \Theta_{21} \\ \Theta_{12} & \Theta_{22} \end{bmatrix} \tag{30}$$

While the scattering matrix computes outgoing waves from incoming waves, the chain-scattering matrix computes the left-going and right-going waves in section 1 given the left-going and right-going waves in section 2. The relation between the scattering matrix Σ and the chain-scattering matrix Θ is given by

$$\begin{aligned} \Theta_{11} &= \Sigma_{21}^{-1} & \Theta_{12} &= -\Sigma_{21}^{-1}\Sigma_{22} \\ \Theta_{21} &= \Sigma_{11}\Sigma_{21}^{-1} & \Theta_{22} &= \Sigma - \Sigma_{11}\Sigma_{21}^{-1}\Sigma_{22} \end{aligned} \tag{31}$$

and

$$\begin{aligned} \Sigma_{11} &= \Theta_{21}\Theta_{11}^{-1} & \Sigma \quad &= \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} \\ \Sigma_{21} &= \Theta_{11}^{-1} & \Sigma_{22} &= -\Theta_{11}\Theta_{12} \end{aligned} \tag{32}$$

4.4. One-Multiplier Forms

Equations (25) and (26) combine to give

$$\begin{aligned} P_1^- &= P_2^+ + \Delta P \\ P_2^- &= P_1^+ + \Delta P \end{aligned} \tag{33}$$

where

$$\begin{aligned} \Delta P &\triangleq \rho\left(P_1^+ - P_2^+\right) \\ \rho &\triangleq \rho_1 = 1 - \alpha_1 \end{aligned} \tag{34}$$

Thus, only one matrix multiplication is necessary to compute the reflected waves from the incoming waves. In the scalar case, this reduces to the so-called *one-multiplier lattice section* [31] (minus the unit-sample delay ordinarily associated with each section). It is well-known that any rational digital filter can be built using one-multiplier lattice sections [31]. In fixed-point implementations, the only source of error would typically be in the computation of ΔP.

To ensure the absence of limit cycles and overflow oscillations, the additions in (33) must be performed before rounding, and the final rounding to obtain P_1^- and P_2^- must be *norm reducing*. (In the scalar case, magnitude truncation is sufficient.) The added expense for postponing round-off until the final outgoing waves are computed is typically negligible, requiring only logic to determine the desired direction of truncation from the signs of P_i^+ and ΔP, and the low-order product of ΔP. In other words, the double precision computations required are only conceptual because the low-order half of the incoming waves is zero.

Another one-multiplier form is obtained by organizing (25) as

$$\begin{aligned} P_1^- &= P_1^+ + \alpha(P_2^+ - P_1^+) \\ P_2^- &= P_1^- + (P_1^+ - P_2^+) = P_J - P_2^+ \end{aligned} \tag{35}$$

where $\alpha \triangleq \alpha_1$. As in the previous case, only one multiply and three adds are required per section. An advantage to (35) is that the computation of P_2^- is noiseless (assuming roundoff only after multiplication). This can be used to simplify hardware for suppressing limit cycles.

In the scalar case, the single section parameter ρ of (35) must lie between -1 and 1, while in (34), the single section parameter α must lie between 0 and 2. Otherwise, the junction is not passive. The practical implication of non-passive junctions is potential filter instability in the presence of feedback.

5. Signal Power in Lossless Waveguides

This section presents basic definitions of signal energy and power in a waveguide. These concepts provide the necessary handles on filter stability, passivity with respect to numerical round-off and overflow, and energy modulation due to time-varying parameters.

5.1. Instantaneous Propagating Power

The *instantaneous power* in a waveguide containing instantaneous pressure P and flow U is defined as the m by m product of pressure and flow:

$$\mathcal{P} = P_{*}U \tag{36}$$

The *total instantaneous power* is defined as the *trace* of the instantaneous power:

$$\mathcal{P}_T \triangleq \mathrm{Tr}\{\mathcal{P}\} = \mathrm{Tr}\{P_{*}U\} \tag{37}$$

The total instantaneous power is a complex scalar measure of power flow. It can be interpreted in a manner similar to the complex power in scalar transmission-line theory in which sinusoidal phasors are propagated in either direction. The instantaneous power can be expanded into four terms as follows:

$$\begin{aligned}
\mathcal{P} = P_{*}U &= (P_{*}^{+} + P_{*}^{-})(U^{+} + U^{-}) = (U_{*}^{+} - U_{*}^{-})Z_{*}(U^{+} + U^{-}) \\
&= P_{*}^{+}U^{+} + P_{*}^{-}U^{-} + P_{*}^{+}U^{-} + P_{*}^{-}U^{+} \\
&= P_{*}^{+}\Gamma P^{+} - P_{*}^{-}\Gamma P^{-} + P_{*}^{+}\Gamma P^{-} - P_{*}^{-}\Gamma P^{+} \\
&= U_{*}^{+}ZU^{+} - U_{*}^{-}ZU^{-} - U_{*}^{+}ZU^{-} + U_{*}^{-}ZU^{+}
\end{aligned} \tag{38}$$

The *right-going* and *left-going* power are defined, respectively, by

$$\begin{aligned}
\mathcal{P}^{+} &= P_{*}^{+}U^{+} = U_{*}^{+}ZU^{+} = P_{*}^{+}\Gamma P^{+} \\
\mathcal{P}^{-} &= P_{*}^{-}U^{-} = -U_{*}^{-}ZU^{-} = -P_{*}^{-}\Gamma P^{-}
\end{aligned} \tag{39}$$

Since Z is para-Hermitian, P^+ and P^- are Hermitian forms, and can be expressed as

$$\begin{aligned} P^+ &= \sum_{i=1}^{m} \lambda_i^+ \underline{v}_i^+ \underline{v}_{i*}^+ \\ P^- &= \sum_{i=1}^{m} \lambda_i^- \underline{v}_i^- \underline{v}_{i*}^- \end{aligned} \tag{40}$$

where $\underline{v}_i^+$ is the ith eigenvector of P^+, and λ_i^+ is its ith (real) eigenvalue. The m-vectors $\underline{v}_i^+$ can be chosen orthonormal. Similar remarks apply to the eigenvalues and eigenvectors of P^-. It can be shown that the waveguide is passive if and only if $\lambda_i^+, \lambda_i^- \geq 0$. Consequently, we will assume in the sequel that

$$\lambda_i^+ > 0, \qquad \lambda_i^- > 0, \qquad i = 1, 2, \ldots, m \tag{41}$$

This implies P^+ and P^- are positive definite Hermitian matrices. The orthonormal vectors $\underline{v}_i^+$ and $\underline{v}_i^-$ (which are vector analytic functions of d), indicate "directions" along which power flows in the m-dimensional manifold determined by U_i (or P_i) and Z.

In the non-para-Hermitian case, the medium is passive iff

$$Z(d) + Z_*(d) \geq 0, \ \forall |\, d\,| \leq 1 \tag{42}$$

That is, the para-Hermitian part of the characteristic impedance of a passive medium must be positive definite in the unit circle. It can be shown using the maximum modulus theorem [5] that (42) holds if $Z + Z_* \geq 0$ for $|\, d\,| = 1$.

We define the *cross power* by

$$P^\times = U_*^+ Z U^- \tag{43}$$

The instantaneous power (38) can now be written as

$$P = (P^+ - P^-) + (P^\times - P_*^\times) \tag{44}$$

which we interpret as a sum of a net traveling-power term $\mathcal{P}^+ - \mathcal{P}^-$ plus the skew-para-Hermitian part of the cross-power. In the real scalar case, $\mathcal{P}^\times \equiv \mathcal{P}_*^\times$ and the cross power is zero.

Since the eigenvalues of a Hermitian matrix are purely real, we define the difference between the right-going and left-going power $\mathcal{P}^+ - \mathcal{P}^-$ as the *active power*. Similarly, since the eigenvalues of a skew-Hermitian matrix are purely imaginary, we define the skew-para-Hermitian part of the cross-power $\mathcal{P}^\times - \mathcal{P}_*^\times$ as the *reactive power*. These definitions parallel those of scalar transmission-line theory. The total power in each case is defined by the corresponding trace.

5.2. Power at a Junction

For the N-way waveguide junction, the constraints (2,6) yield

$$\mathcal{P}_J \triangleq \sum_{i=1}^{N} P_{i*} U_i = \sum_{i=1}^{N} P_J U_i = P_J \sum_{i=1}^{N} U_i = 0 \tag{45}$$

Thus, the N-way junction is *lossless;* no net power, active or reactive, flows into or away from the junction.

5.3. Quantization Effects

While the ideal waveguide junction is lossless, finite wordlength effects, can make exactly lossless networks unrealizable. In fixed-point arithmetic, the product of two numbers requires more bits (in general) for exact representation than either of the multiplicands. If there is a feedback loop around a junction, the number of bits needed to represent exactly a circulating signal grows without bound. Therefore, some sort of round-off rule must be included in a finite-precision implementation of a WGF. The guaranteed absence of limit cycles and overflow oscillations is tantamount to ensuring that all finite-wordlength effects result in power *absorption* at each junction, and never power creation.

The sum of power flows entering a junction is given by

$$
\begin{aligned}
\mathcal{P}_J = \underline{P}_* \underline{U} &\triangleq \sum_{i=1}^{N} P_{i*} U_i \\
&\triangleq \sum_{i=1}^{N} \left(P_{i*}^+ + P_{i*}^-\right)\left(U_i^+ + U_i^-\right) \\
&= \sum_{i=1}^{N} \left(P_{i*}^+ + P_{i*}^-\right)\Gamma_i\left(P_i^+ - P_i^-\right)
\end{aligned}
\tag{46}
$$

where $\Gamma_i = Z_i^{-1}$ is the characteristic admittance of the ith waveguide. Define

$$
\begin{aligned}
\langle \underline{x}, \underline{y} \rangle_{\underline{\Gamma}} &\triangleq \sum_{i=1}^{N} x_{i*} \Gamma_i y_i \\
\| \underline{x} \|_{\underline{\Gamma}} &\triangleq \sum_{i=1}^{N} x_{i*} \Gamma_i x_i = \sqrt{\langle \underline{x}, \underline{x} \rangle_{\underline{\Gamma}}}
\end{aligned}
\tag{47}
$$

Then by (46),

$$
\begin{aligned}
\mathcal{P}_J &= \left\langle \underline{P}^+ + \underline{P}^-, \underline{P}^+ - \underline{P}^- \right\rangle_{\underline{\Gamma}} \\
&= \| \underline{P}^+ \|_{\underline{\Gamma}}^2 - \| \underline{P}^- \|_{\underline{\Gamma}}^2 + \left\langle \underline{P}^-, \underline{P}^+ \right\rangle_{\underline{\Gamma}} - \left\langle \underline{P}^+, \underline{P}^- \right\rangle_{\underline{\Gamma}}
\end{aligned}
\tag{48}
$$

The junction is passive if

$$
\mathcal{P}_J + \mathcal{P}_{J*} = 2\left[\| \underline{P}^+ \|_{\underline{\Gamma}}^2 - \| \underline{P}^- \|_{\underline{\Gamma}}^2\right] \geq 0 \tag{49}
$$

or, $\| \underline{P}^+ \|_{\underline{\Gamma}} \geq \| \underline{P}^- \|_{\underline{\Gamma}}$.

Since $\Gamma_i > 0$, the quadratic form $x_* \Gamma_i x$ forms an elliptic norm for x. By the norm equivalence theorem, condition (49) can be written

$$\left\| \mathcal{P}^- \right\| \leq \left\| \mathcal{P}^+ \right\| \tag{50}$$

where the norm is arbitrary [18]. Let $\hat{P}_i^- = P_i^- - \epsilon_i$ denote the quantized value of P_i^-. Then a sufficient condition for the absence of limit cycles and overflow oscillations in an N-port junction is

$$\left\| \hat{P}_i^- \right\| \leq \left\| P_i^- \right\| \tag{51}$$

Since the norm is arbitrary, two convenient choices are the L^1 norm (maximum abolute column sum) and the L^∞ norm (maximum abolute row sum) [18]. Alternately, a *sufficient condition for the absence of overflow oscillations and limit cycles in networks built from N-port waveguide junctions is that magnitude truncation be used on each element of the final q by m outgoing wave variables P_i^-.*

5.4. The Normalized Ladder Section

We can normalize the pressure and flow variables by the Hermitian square root of the characteristic impedance to obtain propagation waves in units of root power:

$$\begin{aligned} \tilde{P}_i^+ &\triangleq Z_i^{-\frac{1}{2}} P_i^+ \qquad \tilde{P}_i^- \triangleq Z_i^{-\frac{1}{2}} P_i^- \\ \tilde{U}_i^+ &\triangleq Z_i^{\frac{1}{2}} U_i^+ \qquad \tilde{U}_i^- \triangleq Z_i^{\frac{1}{2}} U_i^- \end{aligned} \tag{52}$$

where

$$Z_{i*}^{\frac{1}{2}} = Z_i^{\frac{1}{2}} \tag{53}$$

is the unique para-Hermitian square root of Z_i. The *para-Hermitian square root* of Z_i is defined as the analytic continuation of the Hermitian square root of $Z_i(e^{j\theta})$. Uniqueness is inherited from uniqueness of both the Hermitian square root and the process of analytic continuation.

In the non-para-Hermitian case, we normalize the traveling waves by

$$\begin{aligned} \tilde{P}_i^+ &\triangleq R_i^{-\frac{1}{2}} P_i^+ \qquad \tilde{P}_i^- \triangleq R_i^{-\frac{1}{2}} P_i^- \\ \tilde{U}_i^+ &\triangleq R_i^{\frac{1}{2}} U_i^+ \qquad \tilde{U}_i^- \triangleq R_i^{\frac{1}{2}} U_i^- \end{aligned} \tag{54}$$

where

$$\begin{aligned} R_i &\triangleq \frac{1}{2}(Z_i + Z_{i*}) \\ R_{i*}^{\frac{1}{2}} &= R_i^{\frac{1}{2}} \end{aligned} \tag{55}$$

That is, $R_i^{\frac{1}{2}}$ is the para-Hermitian square root of the para-Hermitian part R_i of the ith branch impedance Z_i [13].

By restricting all waveguides to normalized waves, we obtain the WGF generalization of the *normalized ladder form* (NLF). As a result of this normalization, the *stored power in each WGF section* is unaffected by time variation of the characteristic impedance in each waveguide. This means that the signal power is decoupled from time variation in the filter coefficients.

6. Conclusions

We have derived a general framework for recursive digital filtering which has many desirable properties (stated in the introduction). This architecture for digital filtering is most valuable in the case of time-varying recursive networks in which it is desired to eliminate limit cycles and overflow oscillations. The added complexity relative to the best pre-existing recursive filter architectures is negligible. Therefore, these structures are likely to find plentiful use in advanced time-varying filtering applications.

7. Appendix—The Wave Equation

For an ideal waveguide, we have the following *wave equation*.

$$P_{tt}(x,t) = c^2 P_{xx}(x,t), \tag{56}$$

where $P(x,t)$ denotes longitudinal pressure displacement in the tube at the point x along the tube at time t in seconds. If the length of the tube is L, then x is taken to lie between 0 and L. The partial derivative notation used above is defined by

$$P_{xy} \triangleq \frac{\partial}{\partial x}\left(\frac{\partial P}{\partial y}\right). \tag{57}$$

The constant c is given by $c = \sqrt{T/\rho}$ where T is the tube "tension," and ρ is the mass per unit length of the tube. An elegant derivation of the wave equation is given by Morse [4].

The general *traveling-wave* solution to (56) is given by

$$P(x,t) = P^{+}(x - ct) + P^{-}(x + ct). \tag{58}$$

This solution form is interpreted as the sum of two fixed wave-shapes traveling in opposite directions along the tube. The specific waveshapes are determined by the initial pressure $P(x,0)$ and flow $U(x,0)$ throughout the tube.

[1] G. Szegö, "Ein Grenzwertsatz uber die Toeplitzschen Determinanten einer Reellen Positiven Funktion," *Math. Ann.*, vol. 76, pp. 490–503, 1915.

[2] G. Szegö, *Orthogonal Polynomials*, Amer. Math. Soc., Colloq. Publ. no. 23, New York, 1939.

[3] N. Levinson, "The Weiner RMS (Root-Mean-Square) Error Criterion in Filter Design and Prediction," *J. Math. Physics*, vol. 25, pp. 261–278, 1947.

[4] P. M. Morse, *Vibration and Sound*, published by the American Institute of Physics for the Acoustical Society of America, 1976 (1st ed. 1936, 2nd ed. 1948).

[5] Z. Nehari, *Conformal Mapping*, Dover, New York, 1952.

[6] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications*, University of California Press, Berkeley and Los Angeles CA, 1958.

[7] M. E. Van Valkenburg, *Introduction to Modern Network Synthesis*, John Wiley and Sons, Inc., New York, 1960.

[8] J. Durbin, "The Fitting of Time Series Models," *Rev. L'Institut Intl. de Statistique*, vol. 28, pp. 233–243, 1960.

[9] R. Plonsey and R. E. Collin, *Principles and Applications of Electromagnetic Fields*, McGraw-Hill, New York, 1961.

[10] *Banach Spaces of Analytic Functions*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1962.

[11] V. Belevitch, "Summary of the History of Circuit Theory," *Proc. IRE*, vol. 50, no. 5, pp. 848–855, May 1962.

[12] N. I. Achieser, *The Classical Moment Problem*, Oliver and Boyd, Edinburgh, 1965.

[13] V. Belevitch, *Classical Network Theory*, Holden Day, San Francisco, CA, 1968.

[14] B. Noble, *Applied Linear Algebra*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1969, 1977.

[15] V. M. Adamjan, D. Z. Arov, and M. G. Krein, "Analytic Properties of Schmidt Pairs for a Hankel Operator and the Generalized Schur-Takagi Problem," *Math. USSR Sbornik*, vol. 15, pp. 31–73, 1971.

[16] A. Fettweis, "Digital Filters Related to Classical Structures," *AEU: Archive für Elektronik und Übertragungstechnik*, vol. 25, pp. 79-89, Feb. 1071.

[17] A. Fettweis, "Some Principles of Designing Digital Filters Imitating Classical Filter Structures," *IEEE Trans. on Circ. Theory*, vol. CT-18, pp. 314–316, March 1071.

[18] J. M. Ortega, *Numerical Analysis*, Academic Press, New York, 1972.

[19] J. L. Flanagan, *Speech Analysis, Synthesis, and Perception*, Springer-Verlag, New York, 1972.

[20] A. Fettweis, "Pseudopassivity, Sensitivity, and Stability of Wave Digital Filters," *IEEE Trans. on Circ. Theory*, vol. CT-19, pp. 668–673, Nov. 1972.

[21] A. Sedlmeyer and A. Fettweis, "Digital Filters with True Ladder Configuration," *Int. J. Circuit Theory Appl.*, vol. 1, pp. 5–10, 1973.

[22] A. Fettweis, "Reciprocity, Inter-Reciprocity, and Transposition in Wave Digital Filters," *Int. J. Circuit Theory Appl.*, vol. 1, pp. 323–337, 1973.

[23] A. H. Gray and J. D. Markel, "Digital Lattice and Ladder Filter Synthesis," *IEEE Trans. on Audio Electroacoust.*, vol. AU-21, pp. 491–500, Dec. 1973.

[24] A. Fettweis, "Wave Digital Latice Filters," *Int. J. Circuit Theory Appl.*, vol. 2, pp. 203–211, 1974.

[25] A. Fettweis, "Wave Digital Filters with Reduced Number of Delays," *Int. J. Circuit Theory Appl.*, vol. 2, pp. 319–330, 1974.

[26] K. Meerkötter and W. Wegener, "A New Second-Order Digital Filter without Parasitic Oscillations," *AEU: Archive für Elektronik und Übertragungstechnik*, vol. 29, pp. 312–314, Feb. 1975.

[27] A. Fettweis and K. Meerkötter, "Suppression of Parasitic Oscillations in Wave Digital Filters," *IEEE Trans. Circ. and Sys.*, vol. CAS–22, No. 3, pp. 239–246, Mar. 1975.

[28] A. Fettweis and K. Meerkötter, "On Adaptors for Wave Digital Filters," *IEEE Trans. on Acoust., Speech, and Signal Proc.*, vol. ASSP–23, pp. 516–525, Dec. 1975.

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

[30] A. H. Gray and J. D. Markel, "A Normalized Digital Filter Structure," *IEEE Trans. on Acoust., Speech, and Signal Proc.*, vol. ASSP–23, pp. 268–277, June 1975.

[31] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, Springer-Verlag, New York, 1976.

[32] S. S. Lawson, "On a Generalization of the Wave Digital Filter Concept," *Int. J. Circuit Theory Appl.*, vol. 6, pp. 107–120, 1978.

[33] P. H. Delsarte, Y. V. Genin, and Y. Kamp, "Orthogonal Polynomial Matrices on the Unit Circle," *IEEE Trans. Circ. and Sys.*, vol. CAS–25, pp. 145–160, 1978.

[34] P. DeWilde, A. C. Vieira, and T. Kailath, "On a Generalized Szegö-Levinson Realization Algorithm for Optimal Linear Predictors Based on a Network Synthesis Approach," *IEEE Trans. Circ. and Sys.*, vol. CAS–25, No 9, pp. 663–675, September 1978.

[35] L. V. Ahlfors, *Complex Analysis*, McGraw-Hill, New York, 1979.

[36] Digital Signal Processing Committee, ed., *Programs for Digital Signal Processing*, IEEE Press, New York, 1979.

[37] P. H. Delsarte, Y. V. Genin, and Y. Kamp, "The Nevanlinna-Pick Problem for Matrix-Valued Functions," *Soc. for Indust. and Appl. Math.*, vol. 36, no. 1, pp. 47–61, Feb. 1979.

[38] A. Bultheel and P. Dewilde, "On the Relation between Padé Approximation Algorithms and Levinson/Schur Recursive Methods," *Sig. Proc.: Th. and Appl.*, M. Kunt and F. de Coulon, eds., North-Holland, (EURASIP), pp. 517-523, 1980.

[39] A. H. Gray, "Passive Cascaded Lattice Digital Filters," *IEEE Trans. Circ. and Sys.*, vol. CAS-27, No. 5, pp. 337-344, May 1980.

[40] P. DeWilde and H. Dym, "Schur Recursions, Error Formulas, and Convergence of Rational Estimators for Stationary Stochastic Estimators," *IEEE Trans. on Info. Theory*, vol. IT-27, pp. 446-461, July 1981.

[41] A. Fettweis, "Principles of Complex Wave Digital Filters," *Int. J. Circuit Theory Appl.*, vol. 9, pp. 119-134, 1981.

[42] P. DeWilde and H. Dym, "Lossless Chain Scattering Matrices and Optimum Linear Prediction: The Vector Case," *Int. J. Circuit Theory Appl.*, vol. 9, pp. 135-175, 1981.

[43] P. H. Delsarte, Y. V. Genin, and Y. Kamp, "On the Role of the Nevanlinna-Pick Problem in Circuit and System Theory," *Int. J. Circuit Theory Appl.*, vol. 9, no. 2, pp. 177-187, April 1981.

[44] P. DeWilde, *Orthogonal Filters*, class notes, EE702, Stanford University, Stanford CA, Fall 1982.

[45] B. Friedlander, "Lattice Filters for Adaptive Processing," *Proc. IEEE*, vol. 70, pp. 829-867, Aug. 1982.

[46] B. Friedlander, "Efficient Computation of the Covariance Sequence of an Autoregressive Process," *IEEE Trans. Automat. Contr.*, vol. AC-28, No. 1, pp. 97-99, Jan. 1983.

[47] B. Friedlander, "Efficient Computation of the Covariance Sequence of an Autoregressive Process," *IEEE Trans. Automat. Contr.*, vol. AC-28, No. 1, pp. 97-99, Jan. 1983.

[48] A. Fettweis, "Digital Circuits and Systems," *IEEE Trans. Circ. and Sys.*, vol. CAS-31, No. 1, pp. 31-48, Jan. 1984.

[49] B. Friedlander and J. O. Smith, "Analysis and Performance Evaluation of an Adaptive Notch Filter," *IEEE Trans. on Info. Theory*, vol. IT-30, pp. 283–295, March 1984.

[50] P. P. Vaidyanathan and Sanjit K. Mitra, "Low Passband Sensitivity Digital Filters: A Generalized Viewpoint and Synthesis Procedures," *Proc. IEEE*, vol. 72, pp. 404–423, April 1984.

[51] W. A. Schneider, "The Common Depth Point Stack," *Proc. IEEE*, vol. 72, pp. 1238-1254, Oct. 1984.

[52] P. P. Vaidyanathan, "The Doubly Terminated Lossless Digital Two-Pair in Digital Filtering," *IEEE Trans. Circ. and Sys.*, vol. CAS–32, No. 2, pp. 197–200, Feb. 1985.

[53] P. P. Vaidyanathan and S. K. Mitra, "Passivity Properties of Low-Sensitivity Digital Filter Structures," *IEEE Trans. Circ. and Sys.*, vol. CAS–32, No. 3, pp. 217–224, March 1985.

[54] J. O. Smith, "*An Empirical Study of Limit Cycles* in Various Filter Structures and Number Systems," in preparation.

APPENDIX D

ADAPTIVE EQUALIZATION OF RAPIDLY TIME-VARYING MULTIPATH CHANNELS

Adaptive Equalization of Rapidly Time-Varying Multipath Channels

Julius O. Smith
Benjamin Friedlander

Systems Control Technology, Inc.
1801 Page Mill Road
Palo Alto, California, 94303

Abstract

An adaptive equalizer is proposed for eliminating distortion due to multipath propagation in high-speed digital radio systems. The equalizer is aimed at the case of very fast channel fluctuation, where "fast" is defined relative to the impulse-response duration of the inverse of the instantaneous multipath transfer function. The method is of the decision-feedback type where the demodulated symbols are used to construct an estimate of multipath-induced intersymbol interference. The reconstructed baseband waveform is then delayed and weighted according to the current multipath parameters, and this simulated echo is subtracted from the incoming baseband waveform. The distinguishing feature of our approach is that an explicit model of multipath parameters replaces the tranvsersal equalizer studied previously.

1. Introduction

Microwave systems using digital modulation have developed rapidly since their introduction in the early 1970's. This rapid growth is due in part to the tendency toward digital encoding of all types of data, better transmission quality in digital formats (especially in an interference environment), ease of digital switching, and the rapidly falling costs of digital electronics.

Multipath propagation is often the primary source of error in high-speed digital radio, just as it is in the more familiar FM modulation systems. However, the effect of multipath on microwave-frequency communication differs from its effect on analog FM, and different compensation techniques naturally arise. For narrowband FM radio, outage is a function of thermal noise or flat-fade margin [12,2]. For digital transmission, frequency-selective fading causes severe amplitude and delay distortion which yield errors in excess of flat-fade margin predictions [5,6]. As information rates can be on the order of 100MHz, even small levels of delay distortion can cause intersymbol interference. For example, a 6ns multipath delay can cause complete digital eye closure in a 22Mbps 4-QAM system operating at 6GHz, even though the received signal power falls only 10dB [12].

Techniques for minimizing multipath distortion include frequency diversity, spatial diversity, and equalization. The first two involve transmission on multiple frequencies or use of multiple antennas; when multipath degrades one frequency band or antenna signal, the receiver switches or fades to another frequency/antenna. Spatial diversity alone can reduce multipath outage by a factor of 6 to 12 [4]. Amplitude equalization alone gives approximately a factor of 6 improvement [4]. When spatial diversity techniques are combined with amplitude equalizers, the improvement is surprisingly a factor of 100 to 800, because maximum power combiners convert in-band fade notches to slope, and slope is what is corrected by many amplitude equalizers [12,7]. For a minimum-phase multipath fade, slope equalizers often substantially correct the delay distortion as well as the amplitude fading. However, for nonminimum-phase fades, a slope equalizer attempts to double the phase-delay rather than reduce it.

Adaptive transversal equalizers have been employed recently to equalize both phase and amplitude [9]. Typically five taps are used, where each tap multiplies the received signal over the time of one symbol. Thus, a linear combination of five symbol periods is adaptively optimized. The two error criteria currently in use are the "zero forcing" (ZF) and "least mean square" (LMS) errors [12]. The LMS technique is generally regarded as superior in performance, but more complex to implement. An important advantage of the transversal equalizer is that it is equally effective against nonminimum-phase as well as minimum-phase fades. In one study [10], outage reduction using spatial diversity with transversal equalization was a factor of 3.3 better than that achieved using the same spatial diversity with amplitude (notch) equalization [12].

A further improvement can be obtained using a so-called "decision-feedback" equalizer (DFE) [1,16,14,15,20,21]. Unlike the transversal equalizer, which approximately inverts the multipath channel transfer function, the DFE uses demodulated symbols to subtract out intersymbol interference (ISI) on later symbols. With added processing delay, ISI can be subtracted also from earlier symbols. An LMS version of the DFE is described in [1]. The DFE is theoretically superior to linear equalizers in the case of deep in-band notches [17,18,19]. The DFE approach is similar to echo cancellation [3] wherein the echoes to be canceled are constructed from the estimated symbol-stream and channel model.

An exact inverse filter for the multipath transfer function is recursive, and for minimum-phase multipath an adaptive inverse filter can be utilized [11,13]. Since the inverse filter must be constrained to a stable set, nonminimum-phase multipath transfer functions cannot be inverted by a recursive equalizer. However, under certain conditions the multipath delay can be accurately estimated by tracking with a stabilized inverse filter even in the non-minimum-phase case [13].

In practical high-speed digital radio systems (line of sight microwave links), we have the following characteristics [12]:

- Only two or three paths usually matter.
- Delays range up to approximately 11ns (multi-GHz radio).
- Distortion is due more to intersymbol interference caused by multipath delay than by the fading associated with multipath.
- Fade-notch depth $-20\log|1-|g_t||$ can vary as fast as 100dB per second, where g_t is the multipath secondary-path gain.
- Notch frequencies k/n_t, $k = 1, 2, \ldots$, can change as fast as 50MHz per second, where n_t is the multipath delay.
- Nonminimum-phase fades occur 30 to 40 percent of the time, i.e., $|g_t| > 1$.

Under these conditions, it is not very practical to attempt inversion of the multipath transfer function $H(d, t)$.

This paper outlines a multipath equalizer which is based on "decision-directed" simulation of the multipath channel baseband output using the demodulated data stream. The simulated channel output contains estimated echoes due to multipath which are subtracted from the incoming baseband signal. Our approach differs from the DFE's currently proposed in that the multipath is modeled explicitly rather than using a transversal equalizer to approximately model the effect of multipath on the baseband. The effect of hetrodyning the multipath channel output is absorbed into a complex multipath gain which is is a function of the multipath parameters and the frequency shift. All known features of the modulation format are thus exploited to constrain the online optimization.

2. Formulation

Assume that the modulation format of the transmitted signal is 4-QAM at radian frequency $\omega_c = 2\pi f_c$. Let $\omega(t)$ denote the instantaneous carrier frequency, $\omega_m = 2\pi f_m$ the modulation rate, $x(t)$ the baseband waveform (the convolution of a rate ω_m 4-QAM impulse train with a Nyquist shaping pulse), and $u(t) = x(t)e^{j\omega_c t}$ the transmitted waveform.

A model for a time-varying multipath channel is

$$H(d,t) = 1 + g_t d^{n_t} \tag{1}$$

where g_t is the gain of the secondary path at time t, n_t is the instantaneous multipath delay, and d is the unit-sample delay operator.

The received signal is then $v(t) = H(d,t)u(t) = u(t) + g_t u(t-n_t)$. The recovered baseband signal is

$$y(t) = v(t)e^{-j\omega_c t} = x(t) + \left(g_t e^{-j\omega_c n_t}\right)x(t-n_t) \triangleq x(t) + \alpha_t x(t-n_t) \tag{2}$$

Thus, the baseband signal appears as a multipath-distorted signal itself where the gain of the secondary path is now complex:

$$\alpha_t \triangleq g_t e^{-j\omega_c n_t} \tag{3}$$

This observation allows application of multipath cancellation techniques which are intended for the modulated carrier. (The modulated signal being in the GHz cannot easily be dealt with directly.)

The corresponding time-varying inverse filter appears as

$$\begin{aligned} \hat{x}(t) &= y(t) - \hat{\alpha}_t \hat{x}(t-\hat{n}_t) \\ &= \{x(t) + \alpha_t x(t-n_t)\} - \hat{\alpha}_t\{\hat{x}(t-\hat{n}_t)\} \\ &= x(t) + \alpha_t\{x(t-n_t) - \hat{x}(t-\hat{n}_t)\} \end{aligned} \tag{4}$$

The above inverse multipath filter is robust if $|\hat{\alpha}_t| \ll 1$ and the time variation of the multipath gain α_t is slow compared with $n_t/(1-|\alpha_t|)$. For reasons mentioned in the introduction, it is not very effective to attempt an exact linear inverse of the multipath channel.

3. Multipath Equalization

This section describes a decision-directed FIR equalizer which nonlinearly estimates delayed copies of the baseband signal $x(t)$. These multipath echoes are then subtracted out of the received signal $y(t)$ to give the equalizer output. A block diagram of the processing steps is shown in Fig. 1.

The input to the system is the received signal $y(t)$. Any number of echoes may be estimated, but for clarity we will consider only one. Thus, $y(t) = x(t) + \alpha_t x(t - n_t)$. The estimated echo $\hat{\alpha}_t[\tau_p(t)]\bar{x}\{t - \hat{n}_t[\tau_p(t)]; \tau_s(t)\}$ is subtracted to give

$$\hat{x}(t) = y(t) - \hat{\alpha}_t[\tau_p(t)]\bar{x}\{t - \hat{n}_t[\tau_p(t)]; \tau_s(t)\} \tag{5}$$

where $\tau_s(t)$ is the latest time on which the nonlinear estimate of $x(t)$ is based, and $\tau_p(t)$ is the latest time used in the parameter estimates $\hat{\alpha}_t, \hat{n}_t$. The equalized signal $\hat{x}(t)$ is now sampled synchronously to recover the estimated information sequence. The estimated amplitude is quantized to the nearest signalling amplitude. Clearly, acceptable results will be obtained only when this quantization step produces the true signalling amplitude with high probability.

Given the information sequence, a simulated baseband $\bar{x}_t$ is constructed using knowledge of the Nyquist shaping pulse. This nonlinearly enhanced estimate posseses all known characteristics of the modulation format. The parameters of the modulation format are estimated to the extent that they are unknown; for example, the carrier amplitude A and the precise phase of the switching transient would normally need to be estimated online. The basic idea is to minimize error in the k-step prediction of the baseband signal with respect to switching-time phase, carrier amplitude, and perhaps even the information sequence. The number k of steps in the prediction is determined by the amount of processing delay in the instantaneous carrier estimator, and $\tau_s(t) = t - k$. The mean square of the difference $\bar{x}_t - \hat{x}(t)$ is minimized with respect to these parameters.

The multipath parameters are estimated as follows. The simulated transmission $\bar{x}_t$ is passed through a simulated channel

$$\hat{H}(\hat{\alpha}_t[\tau_p(t)], \hat{n}_t[\tau_p(t)]) = 1 + \hat{\alpha}_t[\tau_p(t)]d^{\hat{n}_t[\tau_p(t)]} \tag{6}$$

to produce a synthetic version of the received signal

$$\bar{y}_t = \hat{H}(\hat{\alpha}_t[\tau_p(t)], \hat{n}_t[\tau_p(t)])\bar{x}_t \tag{7}$$

The channel parameters $\hat{\alpha}_t$ and $\hat{n}_t$ are adaptively adjusted to follow the multipath gain α_t and delay n_t. They also must be predicted ahead of the processing delay. An attractive criterion for minimization here is the difference between the synthetic signal $\bar{y}_t$ and the received signal $y(t)$.

4. Description of Simulation

- A demodulator turns the equalizer output into a symbol stream
- The symbol stream drives a simulated transmitter
- The synthesized transmitter signal is fed to the latest channel model
- The echoes from the channel model output are subtracted from the incoming received signal to produce the equalizer output.

Note that the subtraction time, demodulation time, synthesis time, and channel-model time must all add up to less than the multipath delay. Thus, the throughput of the whole chain should be on the order of a nanosecond, and must therefore be implemented in analog form (e.g. optical) in VLSI. The signal-format estimation and channel modeling proceed in parallel at slower rates. The signal format presumably is close to constant. The channel model changes at the rate atmospheric changes take place, which is sufficiently slow that digital computations can be considered.

Clearly, if the signal comes in recordable bursts, the delay from received signal to cancellor signal is no longer critical, and the multipath cancellation can be carried out offline.

```
         Received signal
                |
                |
               / \
              | + |<------- Estimated Echoes
               \ /
                |
                |-----> Equalized Signal
                |
          ------------
          | Demodulator |
          ------------
                |
                |-----> Information
                |
          ------------
          | Synthesizer |
          ------------
                |
                |
      --------------------
      | Secondary paths model |
      --------------------
                |
                |----------> Estimated Echoes
```

Fig. 1 — Block diagram of the proposed equalizer.

5. References

[1] P. Monsen, "Feedback Equalization for Fading Dispersive Channels," *IEEE Trans. on Info. Theory*, vol. IT-17, pp. 56–64, Jan. 1971.

[2] W. T. Barnett, "Multipath Propagation at 4, 6, and 11 GHz," *Bell Syst. Tech. J.*, vol. 51, pp. 321–361, Feb. 1972.

[3] S. B. Weinstein, "Echo Cancellation in the Telephone Network," *IEEE Comm. Magazine,*, vol. 15, Jan. 1977.

[4] T. S. Giuffrida, "Measurements of the Effects of Propagation on Digital Radio Systems Equipped with Space Diversity and Adaptive Equalization," *Proc. IEEE Int. Conf. on Commun.*, pp. 48.1.1–48.1.6, June 1979.

[5] C. W. Anderson et. al., "The Effect of Selective Fading on Digital Radio," *IEEE Trans. Comm.*, vol. COM-27, pp. 1870–1875, Dec. 1979.

[6] W. T. Barnett, "Measured Performance of High-Capacity Digital Radio," *Tel. Eng. and Mngmnt.*, vol. 83, pp. 88–89, Dec. 1, 1979.

[7] P. Hartmann and B. Bynum, "Adaptive Equalization for Digital Microwave Radio Systems," *Proc. IEEE Int. Conf. on Commun.*, pp. 8.5.1–8.5.6, June 1980.

[8] J. E. Mazo, "Analysis of Decision-Directed Equalizer Convergence," *Bell Syst. Tech. J.*, vol. 59, pp. 1857–1876, Dec. 1980.

[9] S. Qureshi, "Adaptive Equalization," *IEEE Comm. Magazine*, vol. 20, pp. 9–16, March 1982.

[10] K. Morita et. al., "Design Considerations for 4/5/6 L-DI Digital Radio Systems," *Rev. Elec. Commun. Labs.*, vol. 30, no. 5, pp. 846–858, 1982.

[11] B. Friedlander and J. O. Smith, "Multipath Delay Estimation," Tech. Rep. 5466-04, Systems Control Tech., Dec. 1983.

[12] C. A. Siller, "Mulitipath Propagation," *IEEE Comm. Magazine*, vol. 22, pp. 6–15, Feb. 1984.

[13] J. O. Smith and B. Friedlander, "Estimation of Multipath Delay," *ICASSP-84: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing*, Paper 15.9, San Diego CA, March 19-21, 1984. Full version to appear.

[14] M. T. Dudek and J. M. Robinson, "A Decision-Feedback Equalizer and Novel Carrier Recovery Circuit for Digital Radio Relay Systems operating at up to 5 Bit/Hz," *Proc. IEEE Int. Conf. on Commun.*, pp. 41.5.1–41.5.6, June 1980.

[15] M. Joindot et. al., "Baseband Adaptive Equalization for a 16-QAM System in the Presence of Multipath Propagation," *Proc. IEEE Int. Conf. on Commun.*, pp. 13.3.1–13.3.6, June 1981.

[16] H. Sari, "A Comparison of Equalization Techniques on 16-QAM Digital Radio Systems during Selective Fading," *GLOBLECOM '82*, pp. F3.5.1–3.5.6, Nov.-Dec. 1982.

[17] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, 1983.

[18] P. Monsen, "Adaptive Equalization of the Slow Fading Channel," *IEEE Trans. Comm.*, vol. COM-22, pp. 1064–1075, Aug. 1974.

[19] P. A. Bello and K. Pahlavan, "Performance of Adaptive Equalization for Staggered QPSK and QPR over Frequency-Selective LOS Microwave Channels," *Proc. IEEE Int. Conf. on Commun.*, pp. 3H.1.1–3H.1.6, June 1982.

[20] M. S. Mueller and J. Salz, "A Unified Theory of Data Equalization," *Bell Syst. Tech. J.*, vol. 60, pp. 2023–2038, Nov. 1981.

[21] M. Kavehrad, "Adaptive Decision Feedback Cancellation of ISI over Multipath Fading Radio Channels," *Proc. IEEE Int. Conf. on Commun.*, pp. C8.5.1–C8.5.3, June 1983.

APPENDIX E

TIME-VARYING AUTOREGRESSIVE MODELING OF A CLASS OF NONSTATIONARY SIGNALS

TIME-VARYING AUTOREGRESSIVE MODELING OF A CLASS OF NONSTATIONARY SIGNALS

1. INTRODUCTION

The need to estimate and process nonstationary signals arises in many applications. The nonstationary nature of the signals may be caused by the motion of the signal source or the receiver (e.g., in radar or sonar problems), by the properties of the medium in which the signal propagates, or by other physical phenomena. Often the signals are inherently nonstationary such as in the case of a modulated carrier in a voice communication system.

A commonly encountered type of nonstationary process consists of a narrowband or sinusoidal signal with a time-varying center frequency and (possibly stationary) measurement noise. Examples include: the radar return from an accelerating target, a frequency modulated carrier, and a variety of acoustic signals.

Much of the work on the estimation and modeling of noisy signals is based on the assumption that the signals may be considered to be stationary over the observation interval. Relatively little work has been done to take explicitly into account the effects of nonstationarity. A promising approach to this problem, which was recently explored by several authors, is to use time-varying parametric models to represent nonstationary signals. Linear models of the autoregressive (AR) and autoregressive moving average (ARMA) types were considered in [1]-[4]. The parameters of these models are assumed to be time-varying, and the functional dependence on time is assumed to be known up to a finite (preferably small) number of parameters.

In this paper we apply the time-varying parametric modeling approach to the class of nonstationary signals discussed above. The parametric model used is described in section 2. It is shown that the roots of the time-varying AR polynomial, evaluated at a particular time point, contain information about the instantaneous frequency of the signal, in a manner similar to that of the stationary case. In section 3 we briefly summarize an algorithm for estimating the time-varying AR parameters, which is similar to the modified

Yule-Walker (MYW) method [5]. Some examples illustrating the behavior of the algorithm are presented in section 4.

In the stationary case the MYW equations provide an efficient (both in the statistical and the computational sense) technique for estimating the AR parameters of narrowband processes. The structural similarity of the MYW equations in the stationary and nonstationary cases may lead one to believe that the accuracy aspects of the resulting estimates will also be similar. Unfortunately this is not the case. In the nonstationary case, estimation accuracy is degraded by the presence of noise much more than in the stationary case. The reasons for this behavior are briefly discussed in section 3.

2. THE TIME-VARYING AUTOREGRESSIVE MODEL

We will say that a zero-mean process x(t) is a time-varying autoregressive (TVAR) process of order p, if x(t) obey the recursion

$$x(t) = -\sum_{i=0}^{p} a_i(t-i)x(t-i) + e(t) \tag{1}$$

where e(t) is a stationary white noise process with zero-mean and variance σ^2 . The time-varying parameters $\{a_i(t), i=1,\ldots,p\}$ are assumed to be linear combinations of a set of basis time functions $\{f_k(t), k=0,\ldots,m\}$,

$$a_i(t) = \sum_{k=0}^{m} a_{ik} f_k(t) \ . \tag{2}$$

The TVAR model is, therefore, completely specified by a set of constant parameters $\{a_{ik},\ 1 < i < p,\ 0 < k < m\ ;\ \sigma^2\}$.

The choice of the basis functions $f_k(t)$ is an important part of the modeling process. A convenient choice which will be made here is

$$f_k(t) = \left(\frac{t}{T}\right)^k , \tag{3}$$

where T is a normalizing constant, equal to the number of data points in the observation interval. This basis set is by no means the best one. Further work is needed to develop a systematic procedure for designing optimal basis functions for specific classes of nonstationary signals.

By analogy to constant parameter AR models we can associate a spectral function with the autoregressive parameters. Let us define

$$S_x(\omega;t) \triangleq \sigma^2/|A(z;t)|^2, \qquad z=e^{j\omega} , \tag{4}$$

where

$$A(z;t) \triangleq 1+a_1(t-1)z^{-1}+\ldots+a_p(t-p)z^{-p} \ . \tag{5}$$

The roots of the polynomial A(z;t) can be evaluted for any given time point.

The trajectories of these roots provide useful information about the characteristics of the signal being modeled.

To gain some insight into the properties of the TVAR model consider the simple case of a sinusoidal signal with linearly time varying frequency,

$$x(t) = \sin 2\pi(f_0+\alpha t)t \ , \tag{6}$$

where f_0 is the initial frequency, and α is the frequency rate of change. The instantaneous frequency of this signal is given by,

$$f_i(t) = f_0 + 2\alpha t \ . \tag{7}$$

Using simple trigonometric identities it can be shown that x(t) obeys precisely the following recursion,

$$x(t) + b_1(t-1)x(t-1) + b_2(t-2)x(t-2) = 0, \tag{8a}$$

where

$$b_1(t-1) = -\frac{\sin 4\pi(f_i(t)+\alpha)}{\sin 2\pi(f_i(t)-\alpha)} , \tag{8b}$$

$$b_2(t-2) = \frac{\sin 2\pi(f_i(t)-\alpha)}{\sin 2\pi(f_i(t)-3\alpha)} . \tag{8c}$$

As the parameter α tends to zero, these coefficients tend to $b_1 = -2\cos 2\pi f_0$, $b_2 = 1$, which are the parameters of the AR representation of a constant frequency sinusoid. More generally, for sufficiently small values of α we have

$$b_1(t-1) \doteq 2\cos 2\pi[f_0+\alpha(t-1)], \ b_2(t-2) \doteq 1. \tag{9}$$

Thus, the roots of the polynomial A(z;t) will be (approximately) on the unit circle at angles $\pm 2\pi[f_0+\alpha(t-1)]$, corresponding to the instantaneous frequency at time (t-1). Evaluation of the spectral function $S_x(\omega;t)$ will, therefore, produce sharp peaks at the instantaneous sinusoidal frequencies.

This analysis can be generalized to multiple sinusoids as is discussed in [6]. Sinusoidal signals whose frequencies change as a nonlinear function of time can also be approximately represented by TVAR models. This is also true for signals which are not sinusoidal but have a non-zero instantaneous bandwidth. As an example consider the "narrowband" AR process with time-varying center frequency defined by

$$x(t)-2r\cos 2\pi f_i(t-1)x(t-1)+r^2x(t-2)=e(t) \tag{10}$$

where r is very close to unity.

3. THE MODIFIED YULE-WALKER ESTIMATOR

The practical problem considered here is one in which it is desired to fit a TVAR model to a finite number of noisy observations of a signal of the type discussed above. Let $\{y(t), t=1,\ldots,T\}$ be the observed process,

$$y(t) = x(t) + n(t) \ , \tag{11}$$

where $n(t)$ is a noise process uncorrelated with the signal and $x(t)$ is a TVAR process of order p (1). It is straightforward to show that

$$y(t) = -\sum_{i=1}^{p} a_i(t-i)y(t-i) + v(t) \ , \tag{12}$$

where

$$v(t) = e(t) + n(t) + \sum_{i=1}^{p} a_i(t-i)n(t-i) \ . \tag{13}$$

Assuming that $a_i(t)$ are given by (2), we can rewrite (12) as

$$y(t) = -\sum_{i=1}^{p} A_i' z(t-i) + v(t) \tag{14a}$$

where

$$z'(t) \triangleq (f_0(t),\ldots,f_m(t))y(t) \ , \tag{14b}$$

$$A_i \triangleq [a_{i0},\ldots,a_{im}]' \ , \tag{14c}$$

or in matrix notation (in the so-called autocorrelation form),

$$\begin{bmatrix} y(1) \\ \vdots \\ y(t) \\ 0 \\ \vdots \end{bmatrix} = - \begin{bmatrix} z'(1) & & 0 \\ \vdots & \ddots & z'(1) \\ z'(T) & & \vdots \\ 0 & \ddots & z'(T) \\ \vdots & & 0 \\ & & \vdots \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_p \end{bmatrix} + \begin{bmatrix} v(1) \\ \vdots \\ v(T) \\ \vdots \end{bmatrix}$$

$$Y = -Z\,A + V\ . \tag{15b}$$

Let $\overline{Z}$ be a shifted down and extended version of the matrix Z,

$$\overline{Z} = \underbrace{\begin{bmatrix} 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots\cdots & 0 \\ z'(1) & \ddots & \vdots \\ \vdots & & z'(1) \\ z'(T) & \ddots & \vdots \\ 0 & & z'(T) \\ \vdots & & 0 \end{bmatrix}}_{N(m+1)\ \text{Columns}} \begin{matrix} \left.\vphantom{\begin{matrix}0\\ \vdots \\ 0\end{matrix}}\right\} q \text{ rows} \\ \\ \\ \\ \\ \\ \end{matrix} \ . \tag{16}$$

The estimate of the TVAR parameter vector A is computed by solving the set of overdetermined MYW equations

$$[\overline{Z}'Z]\hat{A} = -\overline{Z}'Y\ , \tag{17}$$

where $Z'Z$ is a block-toeplitz matrix with N-p blocks of size (m+1)x(m+1). The estimate $\hat{A}$ can be written as

$$\hat{A} = -[Z'\overline{Z}\overline{Z}'Z]^{-1}Z'\overline{Z}\overline{Z}'Y\ , \tag{18}$$

although in practice we would solve equation (17) without explicit matrix inversion.

To evalute the performance of this estimator it is necessary to study the statistics of the estimation error. It is straightforward to show that

$$A-\hat{A} = [Z'\overline{ZZ'}Z]^{-1}Z'\overline{ZZ'}V \text{ , (estimation error).} \tag{19}$$

In the stationary case (with m=1) it can be shown that if v(t) is finitely correlated and if the delay parameter q (cf. (16)) is appropriately chosen, then the error $A-\hat{A}$ will tend to zero as the number of data points T tends to infinity. In other words, $\hat{A}$ in (18) is a consistent estimator [7]. In [1] it is stated that the estimator will also be consistent for certain time-varying ARMA processes.

The asymptotic error covariance matrix for the MYW estimates was presented in [8],[9] for the stationary case. No results of this type seem to be available for the nonstationary case. However, some preliminary work indicates that in the case of sinusoidal signals with time-varying frequencies, the estimation error of the MYW method may be large compared to the best achievable accuracy predicted by the Cramer-Rao bound. A heuristic explanation of this fact is as follows: for the MYW method to work well in the presence of noise it is necessary that the correlation function of the signal decay relatively slowly (cf. [8]). In the nonstationary case the entries of the matrix $[\overline{Z'Z}]$ can be interpreted as averaged correlation coefficients. It can be shown that for sinusoidal signals with time varying frequencies this "correlation function" decays very quickly, leading to inefficient parameter estimates. In other words, the MYW estimator behaves as if the signal was wideband, whereas the optimal estimator makes full use of its underlying narrowband structure. See [6] for more details.

4. SOME EXAMPLES

To illustrate the behavior of the TVAR parameter estimator we present a few examples. In all of these examples the polynomial basis function in (3) was used.

Example 1

The signal consists of three sinusoids with linearly time-varying frequencies

$$x(t)=\sin 2\pi(0.4-8\times10^{-4}t)t+\sin 2\pi(0.15+10^{-3}t)t + \sin 2\pi(0.1+8\times10^{-4}t)t \tag{20}$$

The model order was $p=6$, and the polynomial order $m=3$ (total of 24 parameters). The number of data points was $T=128$ and the signal was practically noise free (SNR=50 dB). Figure 1 depicts the trajectory of the angles of the roots of $A(z,t)$. Triangles depict true values and circles are the TVAR estimates. Note the good fit achieved by the model, except in the neighborhood of the frequency cross-over points.

Example 2:

Here the signal was a single sinusoid with linearly time-varying frequency

$$x(t) = \sin 2\pi(0.15 + 10^{-4}t)t \tag{21}$$

The model order was $p=2$, the polynomial order was $m=1$ and $T=512$. White noise was added to the signal to give a signal-to-noise ratio of 5 dB. The MYW equations were used with $q = N = 4$. Figure 2 depicts the result. A good model fit is observed throughout the observation interval.

Example 3:

The TVAR model can be used for signals with non-linearly varying frequencies. In this example we have a single sinusoid with sinusoidally varying frequency,

$$x(t) = \sin 2\pi(0.25+0.02 \sin(0.01t))t \ . \tag{22}$$

The model order was $p=2$, polynomial order $m=3$, $T=512$ and SNR = 0 dB. The MYW equations were used with $q=N=8$, and the result is depicted in figure 3. At this low SNR the model fit is not very good in some portions of the observation interval.

From these and many other examples we can make the following observations: (i) the choice of a model order p equal to or somewhat larger than twice the number of sinusoids yielded the best results. (ii) the choice of polynomial order m depends on the maximum rate of change of the frequencies. (iii) At low signal-to-noise ratios the estimator performed poorly, compared to the optimal processor.

5. CONCLUSIONS

Parametric models with time-varying coefficients provide a powerful approach to the representation and processing of nonstationary signals. In particular, the TVAR model was shown to be useful for representing a class of narrowband signals with time-varying center frequencies. The nonstationary equivalent of the modified Yule-Walker method was used to estimate the TVAR parameters. Some problems related to the accuracy of this estimator were briefly discussed. It is suggested that in the nonstationary case the performance of this estimator may be different than expected from its behavior in the stationary case.

REFERENCES

1. Y. Grenier, "Time-dependent ARMA Modeling of Nonstationary Signals," IEEE Trans. Acoustics Speech and Signal Processing, Vol. ASSP-31, No. 4, pp. 899-911, August 1983.

2. T. Subba Rao, "The Fitting of Non-Stationary Time-Series Models with Time-Dependent Parameters," J. of the Royal Statist. Soc. Series B, Vol. 32, no. 2, pp 312-322, 1970.

3. L.A. Liporace, "Linear Estimation of Non-Stationary Signals," JASA, Vol. 58, n° 6, pp 1288-1295, 1975.

4. M. Hall, A.V. Oppenheim, A. Willsky, "Time-Varying Parametric Modeling of Speech," IEEE Decision and Control Conf., New Orleans, pp 1085-1091, 1977.

5. J.A. Cadzow, "Spectral Esitmation: An Overdetermined Rational Model Equation Approach," Proc. IEEE, Vol. 70, No. 9, pp. 907-939, September 1982.

6. K.C. Sharman and B. Friedlander, "Time-Varying Autoregressive Modeling of Nonstationary Signals," in preparation.

7. B. Friedlander, "Instrumental Variable Methods for ARMA Spectral Estimation," IEEE Trans. Acoustics Speech and Signal Processing, Vol. ASSP-31, No. 2, pp 404-415, April 1983.

8. B. Friedlander and K.C. Sharman, "Performance Evaluation of the Modified Yule-Walker Equations," submitted for publication.

9. P. Stoica, T. Söderström and B. Friedlander, "Optimal Instrumental Variable Estimation of the AR Parameters of an ARMA Process," submitted for publication.

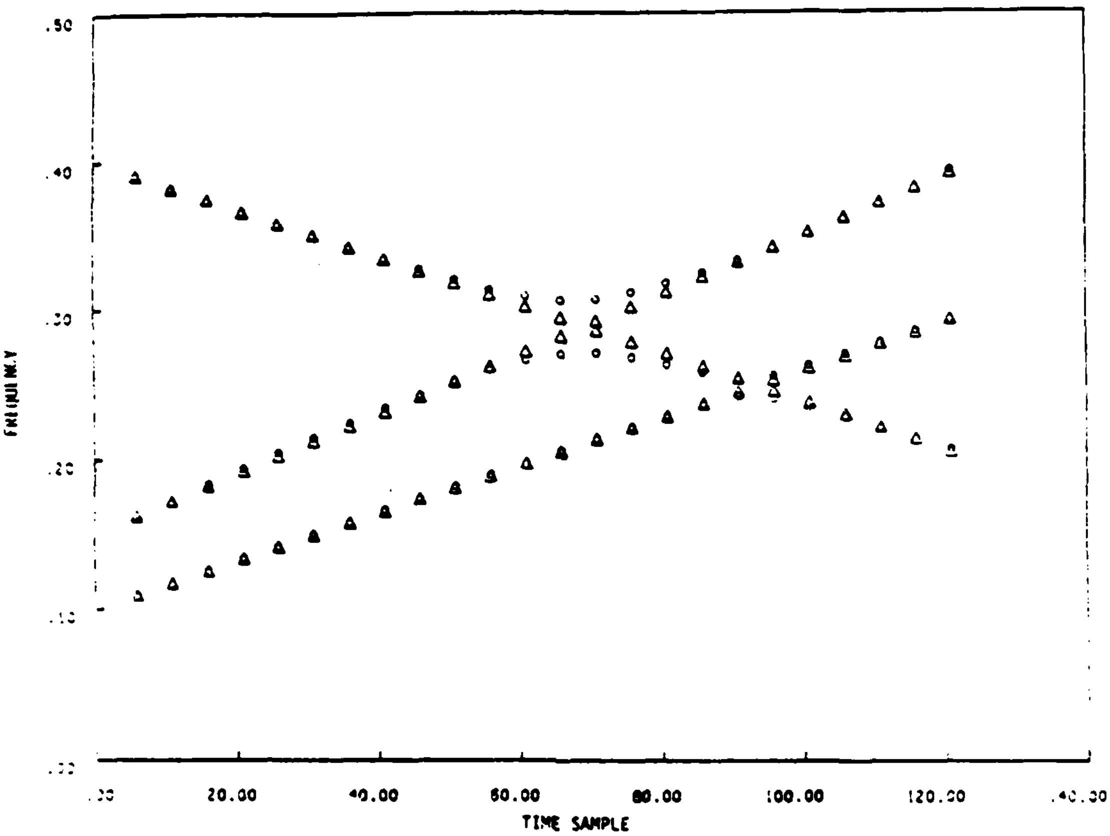


Figure 1 Frequency vs. Time, Example 1

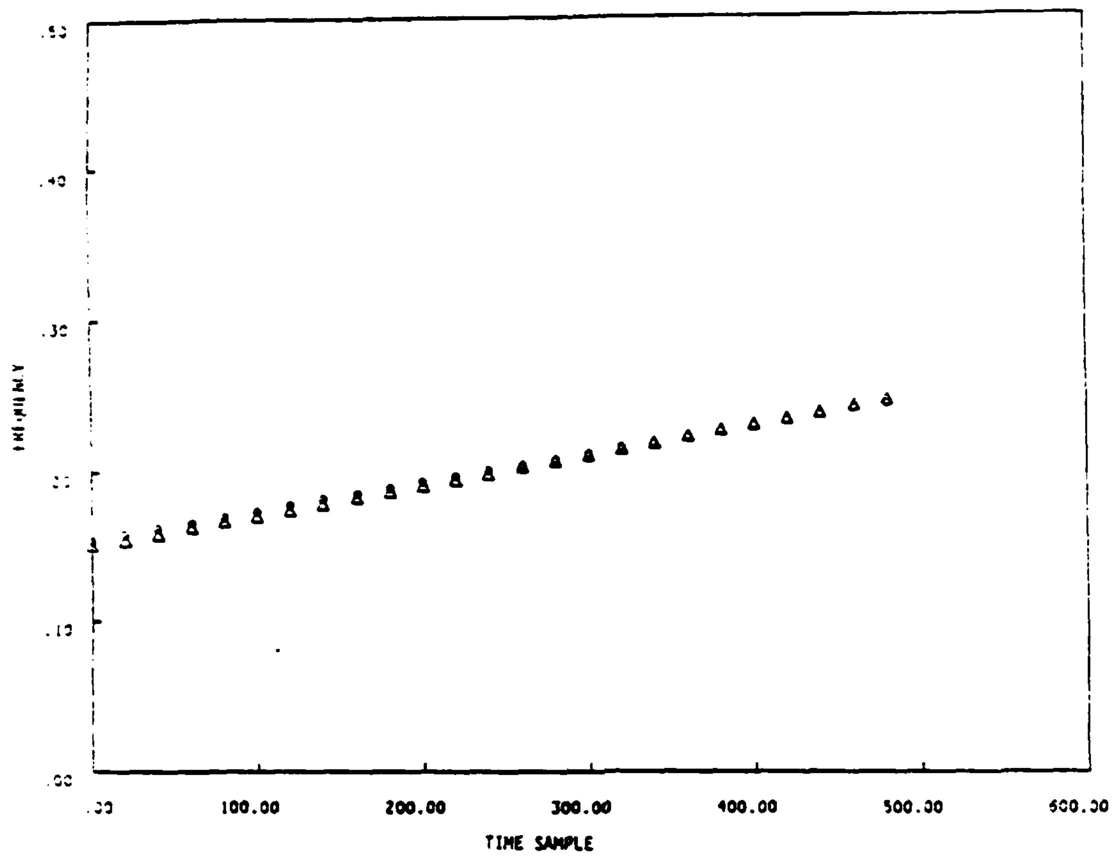


Figure 2 Frequency vs. Time, Example 2

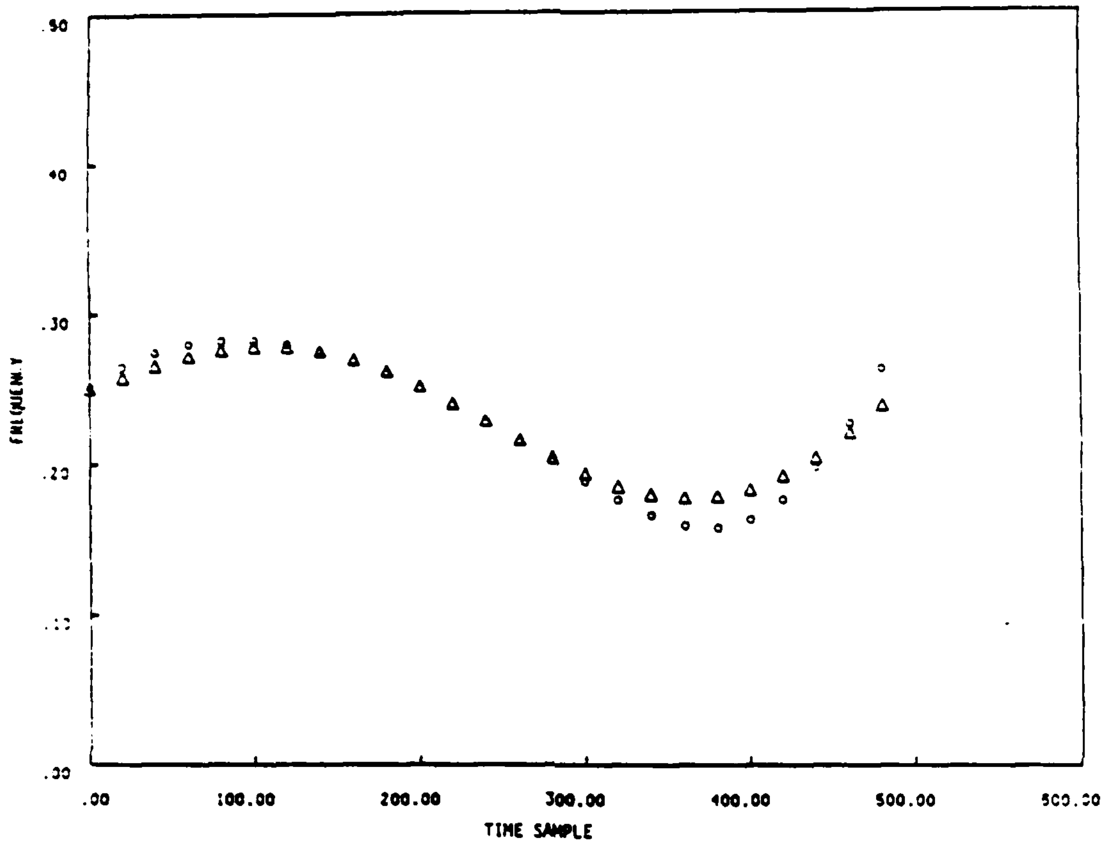


Figure 3 Frequency vs. Time, Example 3

APPENDIX F

SIMULATION RESULTS FOR EXTENSIONS OF THE CONSTANT MODULUS ALGORITHM

PREVIOUS PAGE IS BLANK

Simulation Results for Extensions of the Constant Modulus Algorithm

Julius O. Smith
Benjamin Friedlander

Systems Control Technology Inc.
1801 Page Mill Rd., Palo Alto CA, 94303

Abstract

The Constant-Modulus Algorithm (CMA) computes and applies an adaptive channel equalizer for constant-amplitude signals such as frequency- and phase-modulation. The report "Analysis and Extensions of the Constant Modulus Algorithm" by the authors describes several extensions to the CMA. Those which are examined via simulations here include: (1) Extension from FIR equalizers to IIR equalizers (poles as well as zeros allowed in the equalizer), and (2) Newton's method replaces gradient descent.

This work was supported by the U.S. Air Force (AFSC), Rome Air Development Center, under Contract No. F30602-84-C-0016.

1. Introduction

This is an attachment to the report "Analysis and Extensions of the Constant Modulus Algorithm" by the authors (hereafter referred to as the main report). The purpose here is to provide simulation results for the algorithms studied there. The results here are highly preliminary.

For convenience, let CMA-LC denote the LMS version of the complex CMA algorithm (gradient descent), and let CMA-RC denote the RLS version (Newton descent). The simplified real-only versions are denoted CMA-LR and CMA-RR, respectively.

The loss function minimized by the algorithm is

$$J(\hat{\theta}) = \frac{1}{2}\frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t^2(\hat{\theta}) \tag{1}$$

where

$$\hat{\varepsilon}_t(\hat{\theta}) = \frac{1}{2}\begin{cases} \left| \hat{x}_t(\hat{\theta}) \right|^2 - m_{x_t}^2, & \text{Complex Signals} \\ \hat{x}_t^2(\hat{\theta}) - \sigma_{x_t}^2, & \text{Simplified Real Case} \end{cases} \tag{2}$$

1.1. Gradient Descent versus Newton Descent

The versions CMA-LC and CMA-LR use gradient descent (essentially the LMS algorithm), and CMA-RC and CMA-RR use the recursive Gauss-Newton method (RGN). Prior treatments of the CMA have been based exclusively on gradient descent.

All four algorithms are implemented in their time-recursive forms, that is, they attempt to minimize $J(\hat{\theta})$ with respect the $\hat{\theta}$ recursively in time as T increases.

The LMS version is

$$\hat{\theta}_t = \hat{\theta}_{t-1} - \mu J'(\hat{\theta}_{t-1}) \tag{3}$$

where $J'(\hat{\theta}_{t-1})$ is the gradient estimate of $J(\hat{\theta})$ with respect to $\hat{\theta}$ at time t, evaluated at $\hat{\theta} = \hat{\theta}_{t-1}$, and μ is a step-size parameter which is set by the user.

The RGN version is given by

$$\hat{\theta}_t = \hat{\theta}_{t-1} - J''(\hat{\theta}_{t-1})^{-1} J'(\hat{\theta}_{t-1}) \tag{4}$$

where $J''(\hat{\theta})$ is the Hessian, or second-derivative matrix of $J(\hat{\theta})$ with respect to $\hat{\theta}$.

Before proceeding to the examples, we summarize the signal model and algorithm from the main report.

1.2. The Signal Model

Let y_t denote the received signal for $t = 1, 2, \ldots, T$. In the complex case, y_t is assumed to be of the form

$$y_t = H_{xy}(d)x_t + H_{uy}(d)u_t + H_{xy}(d)\nu_t \tag{5}$$

where

$$x_t = m_x e^{j\varphi_t} \tag{6}$$

is the transmitted signal (φ_t is the real-valued information-bearing signal), ν_t is additive noise, u_t is an interference signal (assumed at least partially known), and $H_{xy}(d)$ and $H_{uy}(d)$ are the linear time-invariant channel filters associated with x_t and u_t respectively:

$$\begin{aligned}
H_{xy}(d) &\triangleq \frac{C(d)}{A(d)} \\
H_{uy}(d) &\triangleq \frac{B(d)}{A(d)} \\
A(d) &\triangleq a_0 + a_1 d + a_2 d^2 + \cdots + a_{n_a} d^{n_a} \\
B(d) &\triangleq \phantom{a_0 +{}} b_1 d + b_2 d^2 + \cdots + b_{n_b} d^{n_b} \\
C(d) &\triangleq 1 + c_1 d + c_2 d^2 + \cdots + c_{n_c} d^{n_c}
\end{aligned} \tag{7}$$

where $\{a_i, b_i, c_i\}$ are real. The unit-delay operator d is defined by the relation

$$d^n x_t \triangleq x_{t-n} \tag{8}$$

for an arbitrary signal x_t. We assume the modulus m_x of the transmitted signal x_t is known for all t.

In the case of real signals, the transmitted signal is assumed to be of the form

$$x_t = m_x \cos(\varphi_t) \tag{9}$$

and u_t and ν_t are real interference and additive noise, respectively. We also assume $\varphi_t = \omega_c + \psi_t$ where

$$\left| \frac{\partial \psi_t}{\partial t} \right| \ll \omega_c \tag{10}$$

1.3. Algorithm Summary

Below, the notation $\hat{c}[i]$ denotes the ith element of the vector $\hat{c}$.

$$
\begin{aligned}
y_t^f &= y_t - \hat{c}_{t-1}[1]y_{t-1}^f - \cdots - \hat{c}_{t-1}[n_{\hat{c}}]y_{t-n_{\hat{c}}}^f \\
\hat{\varphi}_t &= [y_t, y_{t-1}, \ldots, y_{t-n_{\hat{a}}}, -u_{t-1}, \ldots, -u_{t-n_{\hat{b}}}, -\bar{x}_{t-1}, \ldots, -\bar{x}_{t-n_{\hat{c}}}]^T \\
\hat{\varphi}_t^f &= [y_t^f, y_{t-1}^f, \ldots, y_{t-n_{\hat{a}}}^f, -u_{t-1}^f, \ldots, -u_{t-n_{\hat{b}}}^f, -\bar{x}_{t-1}^f, \ldots, -\bar{x}_{t-n_{\hat{c}}}^f]^T \\
\hat{x}_t &= \hat{\varphi}_t^T \hat{\theta}_{t-1} \\
\hat{x}_t' &= \hat{\varphi}_t^f \\
\hat{e}_t' &= g(\hat{x}_t \hat{x}_t') = \begin{cases} Re\{\hat{x}_t \hat{x}_t'\}, & \text{Complex Signals} \\ \hat{x}_t \hat{x}_t', & \text{Simplified Real Case} \end{cases}
\end{aligned}
$$

$$
\hat{e}_t(\hat{\theta}) = \frac{1}{2}\begin{cases} \left|\hat{x}_t(\hat{\theta})\right|^2 - m_{x_t}^2, & \text{Complex Signals} \\ \hat{x}_t^2(\hat{\theta}) - \sigma_{x_t}^2, & \text{Simplified Real Case} \end{cases}
$$

$$
R_t = \begin{cases} \mu, & \text{LMS version (Gradient descent)} \\ \lambda_t R_{t-1} + \hat{e}_t' \hat{e}_t'^T, & \text{RML version (Newton descent)} \end{cases} \tag{11}
$$

$$
\begin{aligned}
\lambda_t &= \frac{w_{t-1}}{w_t} < 1 \\
\hat{\theta}_t &\triangleq \left[\hat{a}_t[0], \hat{a}_t[1], \ldots, \hat{a}_t[n_{\hat{a}}], \hat{b}_t[1], \ldots, \hat{b}_t[n_{\hat{b}}], \hat{c}_t[1], \ldots, \hat{c}_t[n_{\hat{c}}]\right]^T \\
\hat{\theta}_t &= \hat{\theta}_{t-1} - R_t^{-1} \hat{e}_t' \hat{e}_t \\
\hat{C}_t(z) &= \mathrm{Proj}\{\hat{C}_t(z)\} \\
\bar{x}_t &= \hat{\varphi}_t^T \hat{\theta}_t \\
\bar{x}_t^f &= \bar{x}_t - \hat{c}_t[1]\bar{x}_{t-1}^f - \cdots - \hat{c}_t[n_{\hat{c}}]\bar{x}_{t-n_{\hat{c}}}^f \\
u_t^f &= u_t - \hat{c}_t[1]u_{t-1}^f - \cdots - \hat{c}_t[n_{\hat{c}}]u_{t-n_{\hat{c}}}^f \\
\hat{x}_t^f &= \hat{x}_t - \hat{c}_t[1]\hat{x}_{t-1}^f - \cdots - \hat{c}_t[n_{\hat{c}}]\hat{x}_{t-n_{\hat{c}}}^f
\end{aligned}
$$

2. Series 1 — First-Order FIR Equalizer

In this series of examples, the channel is described by

$$H_{xy}(d) = \frac{1}{a_0 + a_1 d} \tag{12}$$

and the channel input is given by

$$x_t = m_x e^{j\varphi_t} \tag{13}$$

for $t = 0, 1, 2, \ldots, T-1$.

The CMA estimates a_0 and a_1. As shown in the main report, in this case the CMA is guaranteed to converge with probability one to the true solution $[a_0, a_1]$ as long as at least two parameters are estimated and the phase modulation φ_t is sufficiently rich.

2.1. Example 1-1 — Sinsusoidal FM, No Noise

The specific values used in this example are

$$\begin{aligned}
\varphi_t &= \omega_c t + \pi\beta \sin\omega_m \\
T &= 2048 \\
a_0 &= 1 \\
a_1 &= -0.99 \\
m_x &= 0.1 \\
\beta &= 2.375 \\
\omega_m &= 2\pi 0.10 \\
\omega_c &= 2\pi 0.11
\end{aligned}$$

Figure 1a shows the channel input signal x_t, and Fig. 1b shows the channel output signal y_t. The amplitude modulation due to the channel filtering is quite visible.

Figure 2a shows the angle modulation waveform, and Figure 2b shows the post-detection version, obtained by demodulating $y_t = H_{xy}(d)x_t$ to obtain $\hat{\varphi}_t$. Figure 3 gives the spectra of the curves in Fig. 2. The demodulated signals have some artificial discontinuities due to wrap-around.

Figure 4a shows a close-up of the inverted normalized loss function

$$-J(\hat{\theta}) = -\frac{1}{T}\sum_{t=1}^{T}\left[|\hat{x}_t|^2 - m_x^2\right]^2 \tag{14}$$

where

$$\hat{x}_t = \hat{a}_0 y_t + \hat{a}_1 y_{t-1} \tag{15}$$

and $\hat{a}_0 = 1$. The peak is zero at $\hat{a}_1 = -0.99$, indicating a lack of bias in the solution. Note, however, that the vertical axis covers only a small range (near the machine accuracy limit); thus, the error surface is quite flat near the solution, indicating a large asymptotic variance is expected in the estimate of $\hat{a}_1$. This point is clearer in Fig. 4b which displays the inverted loss function over a wider range of $\hat{a}_1$.

Figure 5a shows a close-up of the inverted normalized loss function (14) where this time $\hat{a}_1 = -0.99$ and $\hat{a}_0$ is varied. The peak is zero at $\hat{a}_0 = 1$ as it must be. Again, however, the error surface is very flat. Figure 5b shows a view from "farther back."

A very flat loss function means that *both* the gradient J' and Hessian J'' are close to zero. This means trouble for RGN which attempts to move in the direction $[J'']^{-1}J'$. Newton's method is well-known to be unstable where the curvature vanishes. Root finders, for example, typically test for this condition and take countermeasures when necessary. In our case, a reasonable modification is to add a small constant times the identity matrix to the Hessian estimate. This is accomplished at initialization of the algorithm.

Since the third derivative vanishes for an even function, and since a fourth-order Taylor series exactly represents the loss function (14), a descent method based on the fourth-derivative matrix (or tensor) would be expected to converge very rapidly.

Figure 6 is a 3D plot of the inverted error surface (14) versus $\hat{a}_0$ and $\hat{a}_1$. Figure 7 shows the same plot from a different angle. Note how flat the surface is near the middle. There almost appears to be a ring of equal cost. However, we know from analysis that there are only two local maxima in this function.

The LMS CMA

First we will run the standard LMS version of the CMA to establish a base for comparison. We found empirically that $\mu = 100$ was a good setting. Smaller μ gave slower convergence (by decreasing the slope magnitude of the essentially linear $\hat{a}_1$ trajectory). Larger μ was avoided because the estimate is already starting to oscillate.

The initial conditions for this and all later examples are $\hat{a}_0 = 1, \hat{a}_1 = 0$. All other initial state, when required is initialized to zero unless otherwise stated.

Figure 8 shows the estimate of $\hat{a}_1$ produced by CMA-LC for this example. Figure 9a shows only the first 300 samples of the parameter trajectory CMA-LC so that the details of the received signal y_t (Fig. 9b) and the modulus error $[|\hat{x}_t|^2 - m_x^2]/2$ (Fig. 9c) can be seen in relationship to the $\hat{a}_1$ trajectory. The final value of $\hat{a}_1$ at sample 2048 is $\hat{a}_1 = -0.9892$ which is very close to the correct value $= -0.99$.

Figure 10 shows the estimate of $\hat{a}_1$ for the real-signal algorithm CMA-LR, and Figure 11 gives the close-up view. We see that the asymptotic convergence is very slow. In this case, there is no theoretical assurance that the true solution is obtained asymptotically.

The RGN CMA

Figure 12 shows the estimate of $\hat{a}_1$ produced by CMA-RC for this example. Figure 13a shows only the first 300 samples of the parameter trajectory for CMA-RC so that the details of the received signal y_t (Fig. 13b) and the modulus error $[|\hat{x}_t|^2 - m_x^2]/2$ (Fig. 13c) can be seen in relationship to the $\hat{a}_1$ trajectory. The parameter $\hat{a}_1$ converges almost immediately to a constant ($\hat{a}_1 = -0.9929$), and the modulus error rapidly approaches a constant near 0.175. While $\hat{a}_1 = -0.9929$ is close to $= -0.99$, it is not as close as we expect. Since we know there can be no bias, it is clear that the CMA-RC algorithm is incorrectly implemented. That is, we have a "bug" in the simulations program. Unfortunately, time was not available to track it down. Aside from the bias, note the extremely rapid convergence as compared with the LMS version.

Figure 14 shows the estimate of $\hat{a}_1$ for the real-signal algorithm CMA-RR, and Figure 15 gives the close-up view.

2.2. Example 1-2 — Noise FM, No Noise

This example is the same as example 1.1 in every respect except for the modulation signal which is now

$$\varphi_t = H_l(d)\epsilon_t \tag{16}$$

where ϵ_t is unit-variance white Gaussian noise, and $H_l(d)$ is a 6th order elliptic function lowpass filter with cut-off at one-twentieth the sampling rate (a tenth-band filter). The remaining figures are exactly analogous to example 1.1.

3. Conclusions

The LMS version of the CMA seems to be working in both the complex and the real-signal cases. The RML version of the CMA seems to be possibly correct in the real-signal case, but definitely incorrect in the complex case.

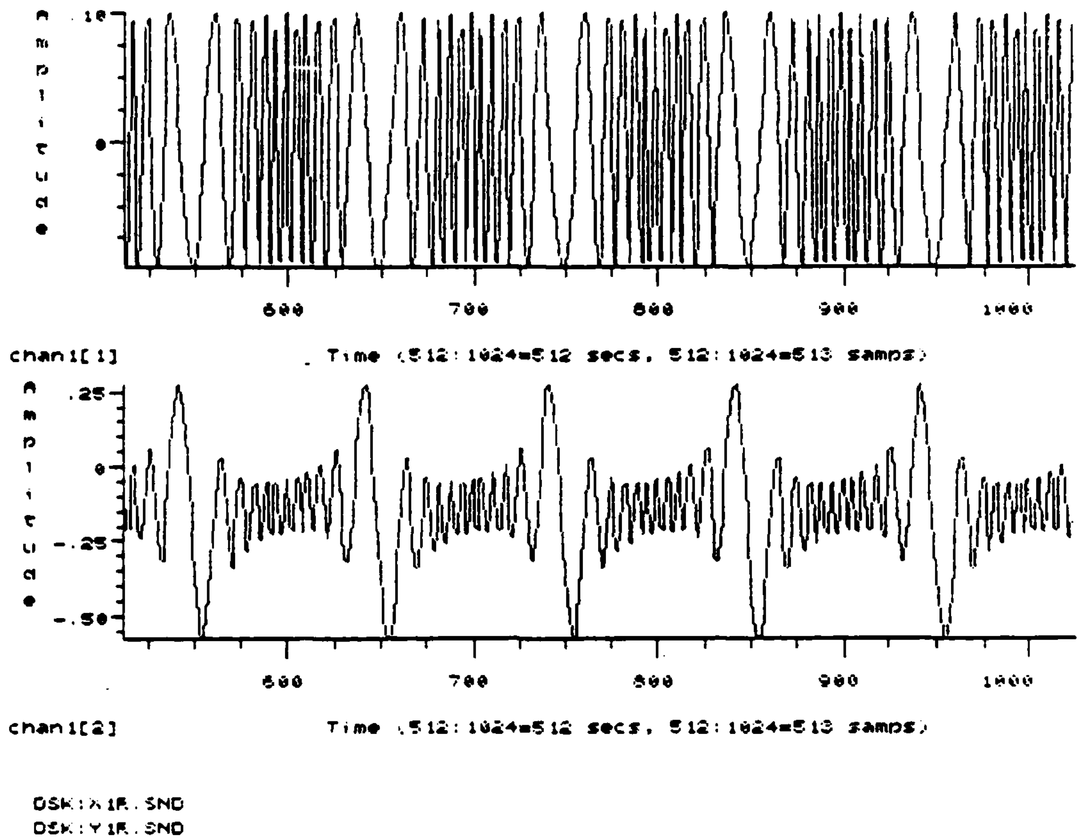

Amplitude
10
0
600
700
800
900
1000
chan1[1]
Time (512:1024=512 secs, 512:1024=513 samps)
Amplitude
.25
0
-.25
-.50
600
700
800
900
1000
chan1[2]
Time (512:1024=512 secs, 512:1024=513 samps)
DSK:X1R.SND
DSK:Y1R.SND

Figure 1a

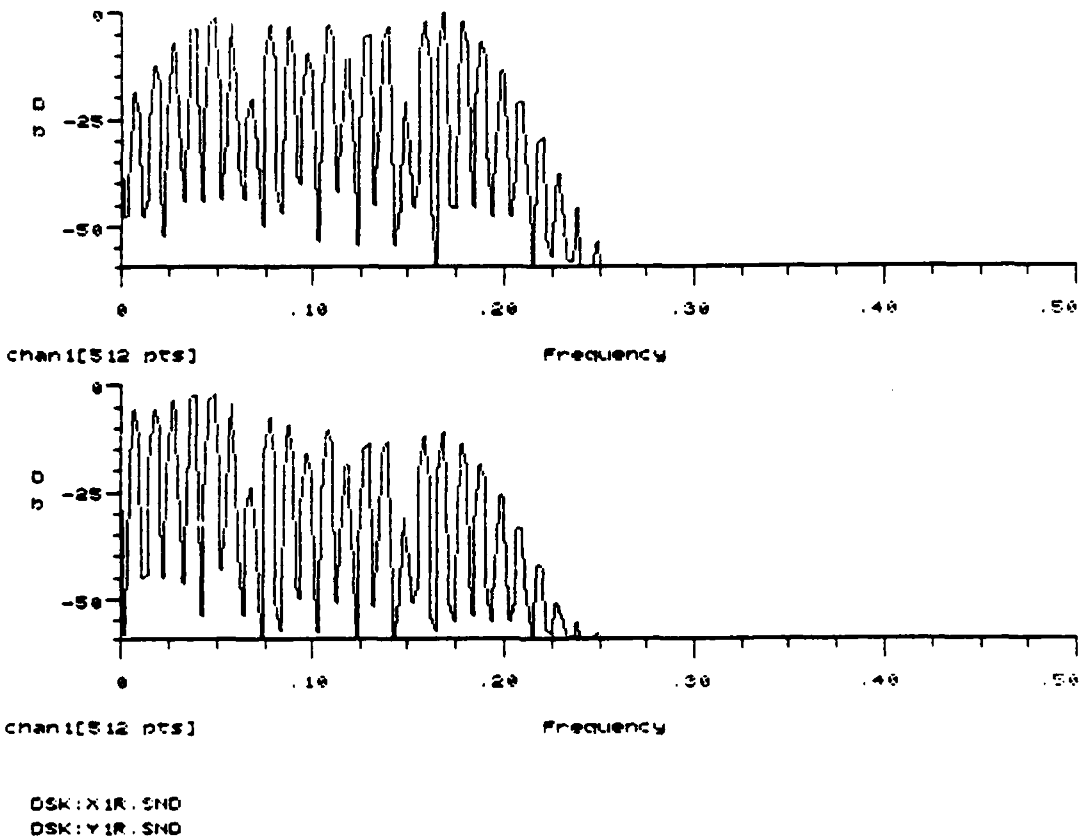


Figure 1b Spectra of Figure 1 Curves

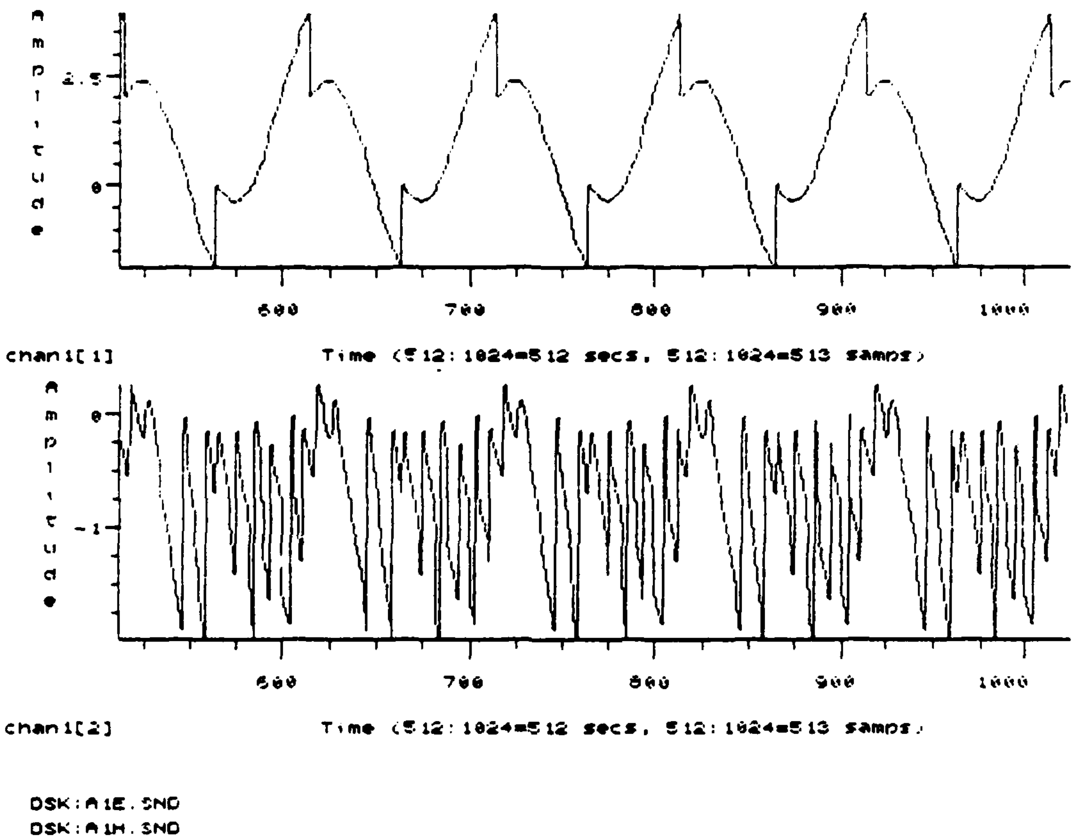

Amplitude
2.5
0
600
700
800
900
1000
chan1[1]
Time (512:1024=512 secs, 512:1024=513 samps)
Amplitude
0
-1
600
700
800
900
1000
chan1[2]
Time (512:1024=512 secs, 512:1024=513 samps)
DSK:A1E.SND
DSK:A1H.SND

Figure 2

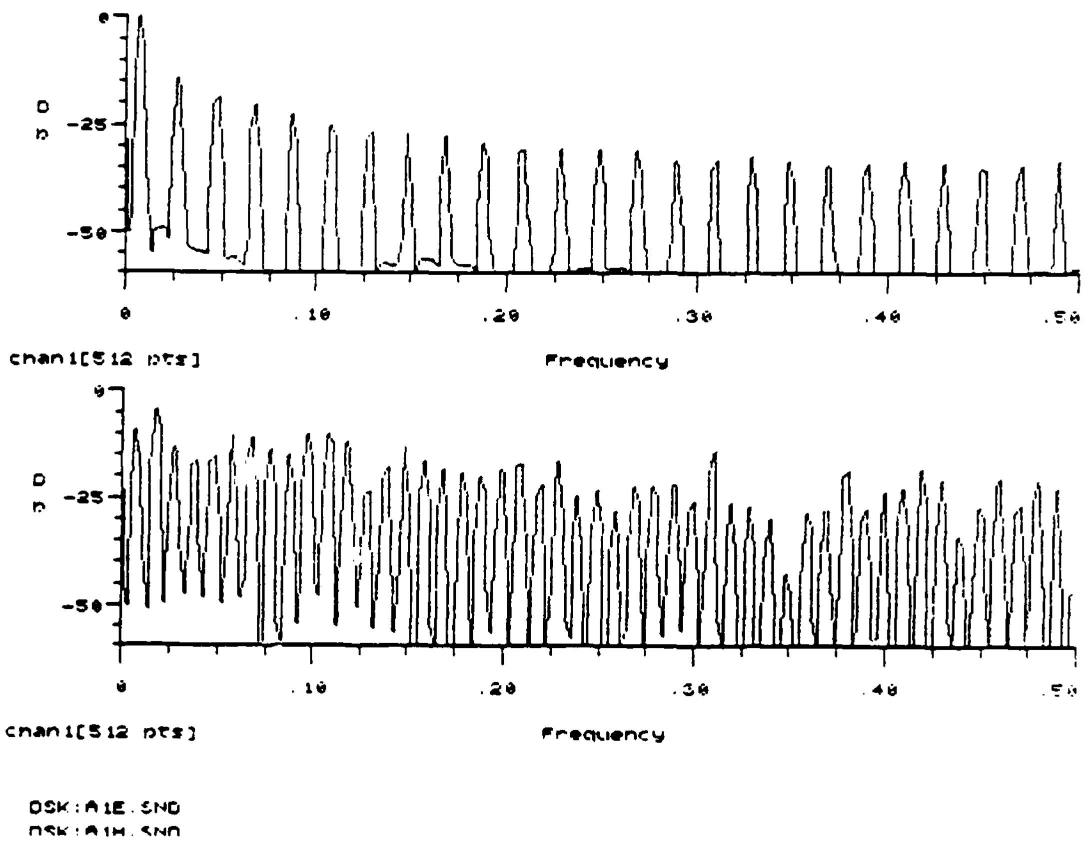
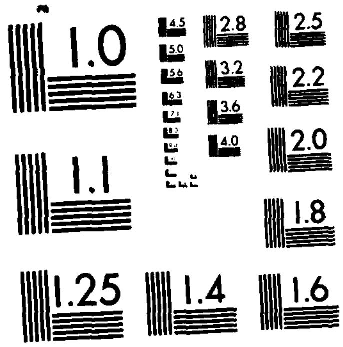

0
-25
-50
dB
0
.10
.20
.30
.40
.50
chan1[512 pts]
Frequency
0
-25
-50
dB
0
.10
.20
.30
.40
.50
chan1[512 pts]
Frequency
DSK:A1E.SND
DSK:A1H.SND

Figure 3

AD-A163 054 TIME DOMAIN ALGORITHMS(U) SYSTEMS CONTROL TECHNOLOGY INC PALO ALTO CA B FRIEDLANDER ET AL SEP 85 PA-6502-01 RADC-TR-85-167 F30602-84-C-0016 2/2

UNCLASSIFIED F/G 17/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

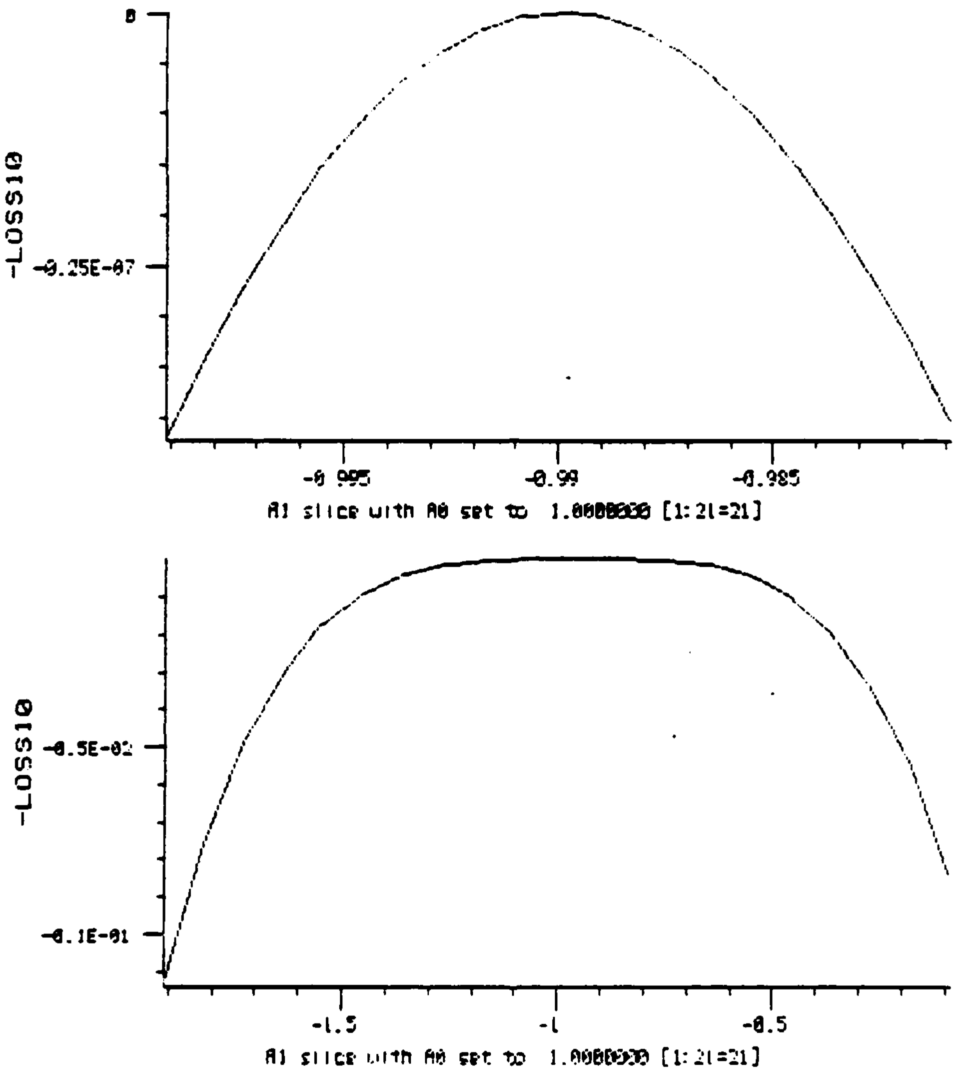
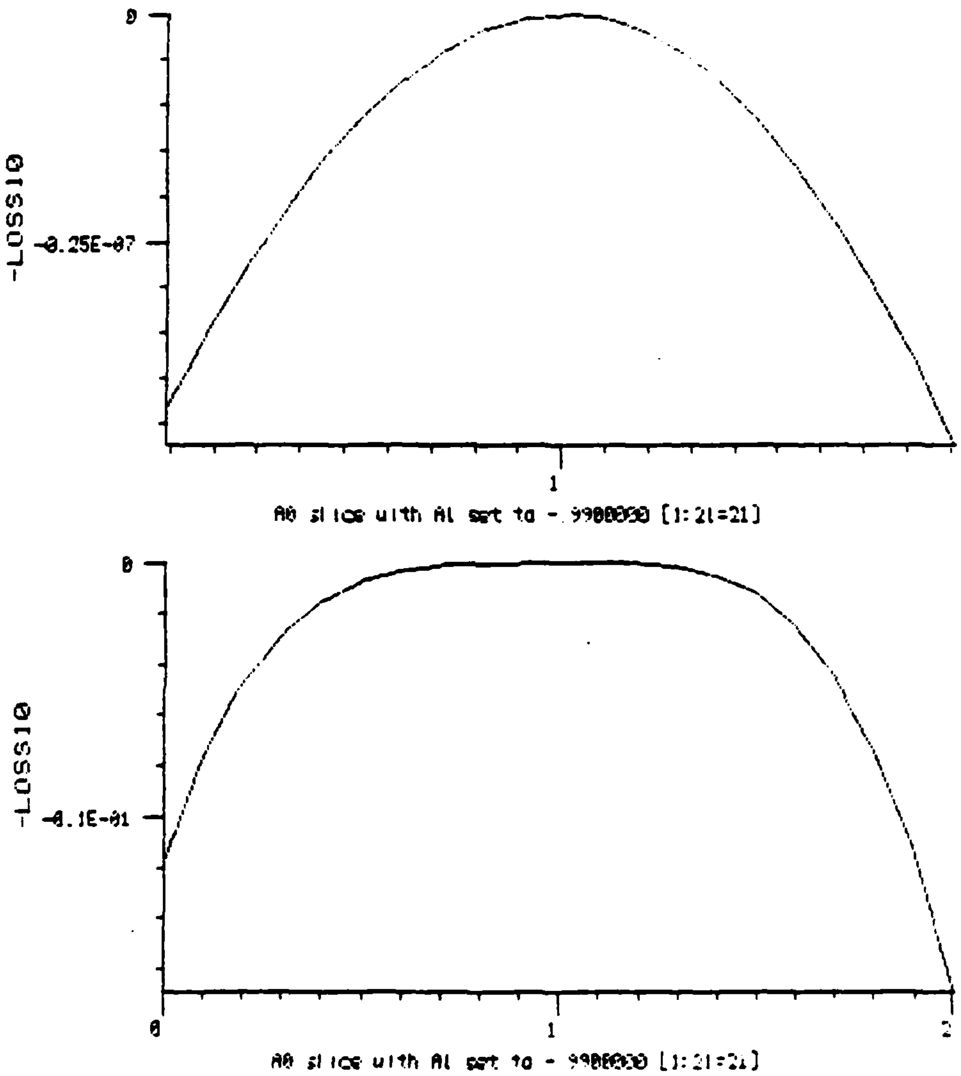

-LOSS10
0
-0.25E-07
-0.995
-0.99
-0.985
A1 slice with A0 set to 1.0000000 [1:21=21]
-LOSS10
-0.5E-02
-0.1E-01
-1.5
-1
-0.5
A1 slice with A0 set to 1.0000000 [1:21=21]

Figure 4



0
-0.25E-07
-LOSS10
1
A0 slice with A1 set to -.9900000 [1:21=21]
0
-0.1E-01
-LOSS10
0
1
2
A0 slice with A1 set to -.9900000 [1:21=21]

Figure 5

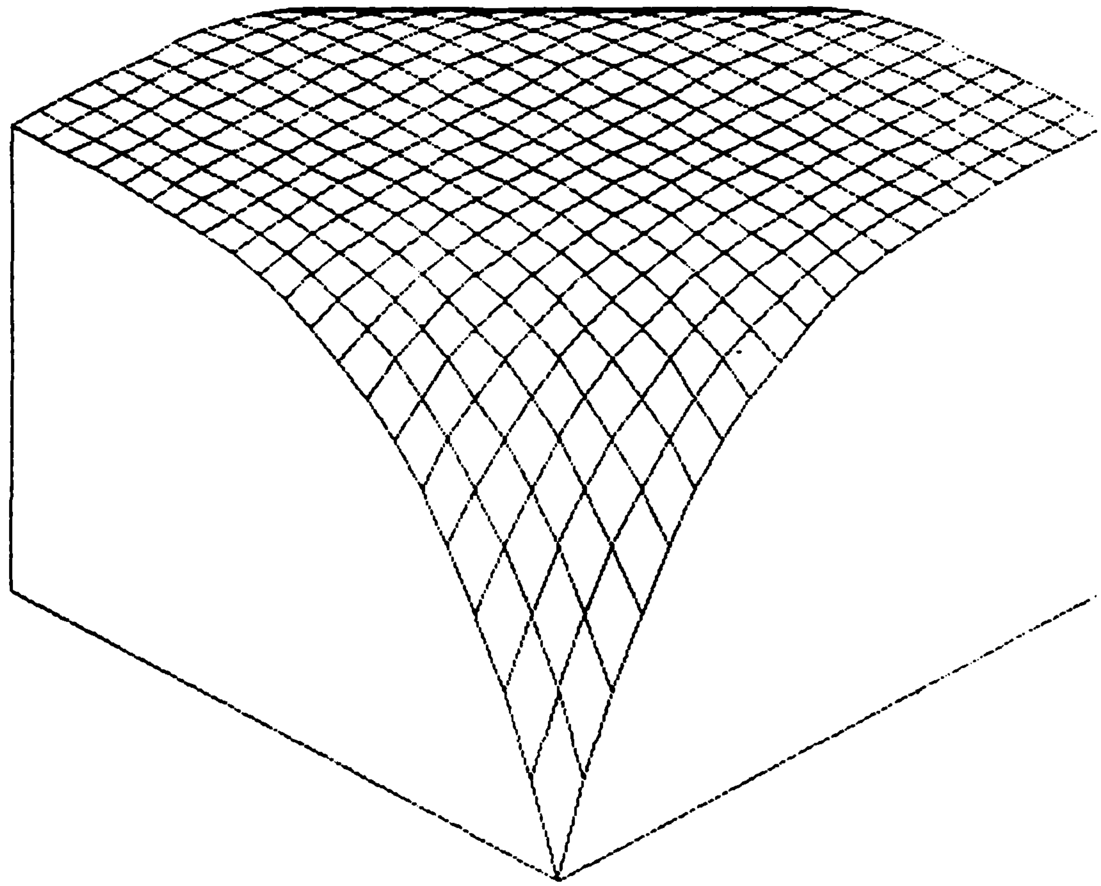


Figure 6

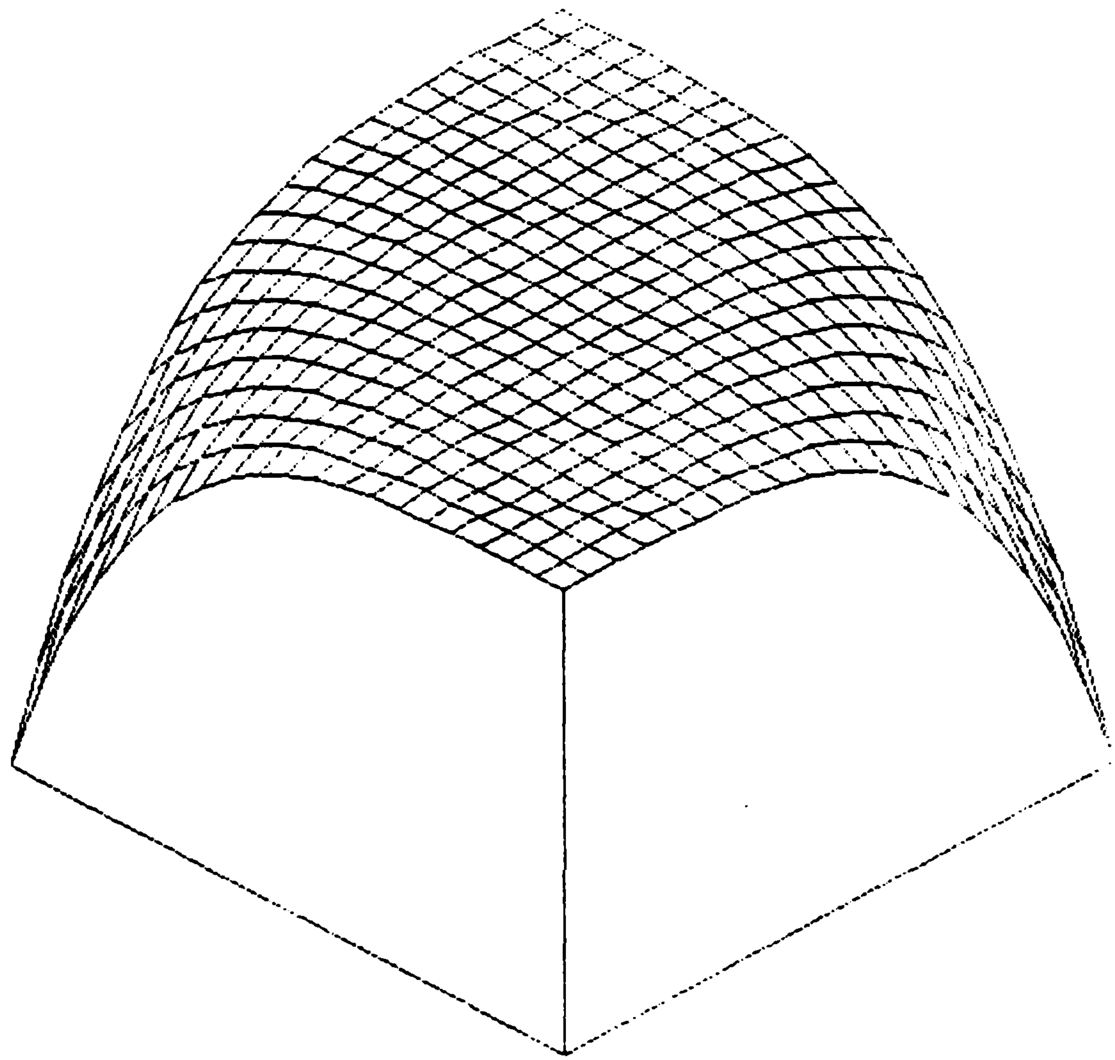


Figure 7

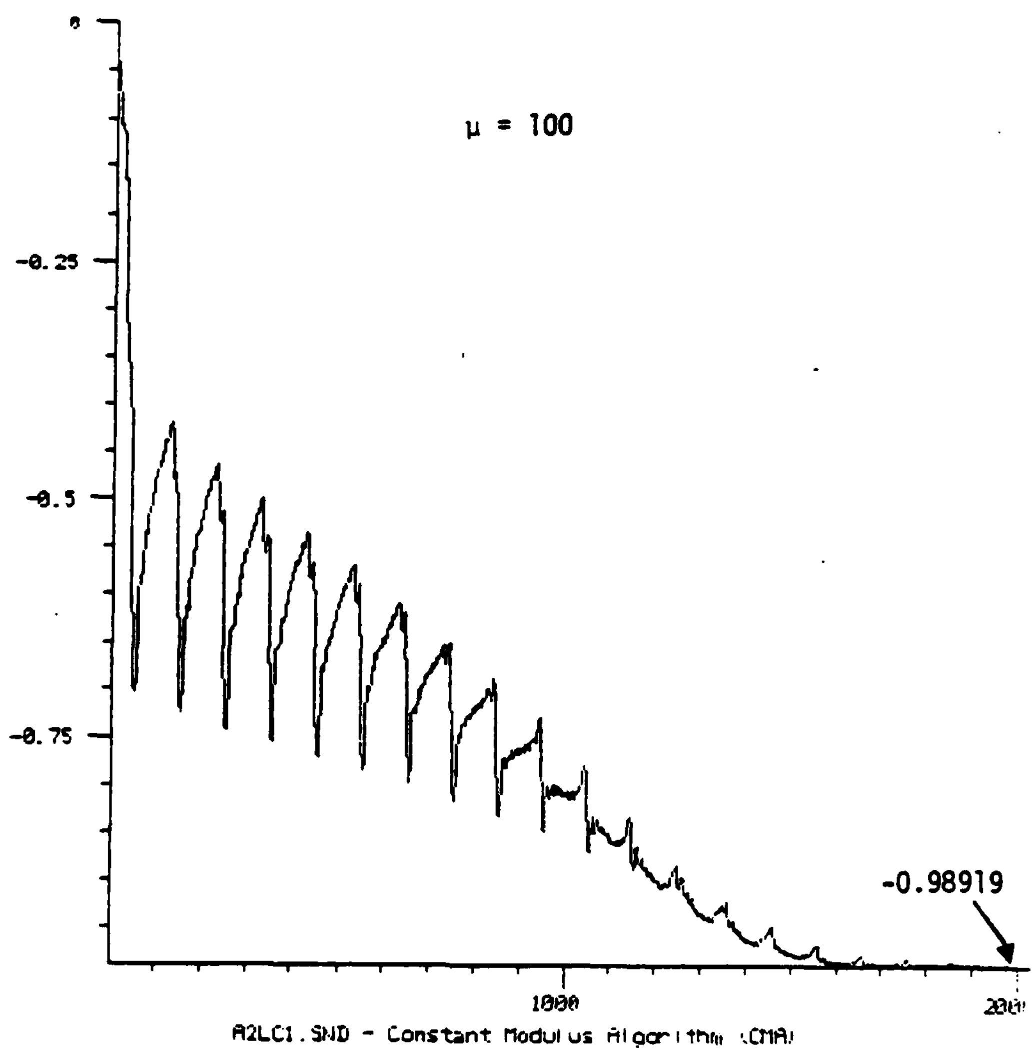
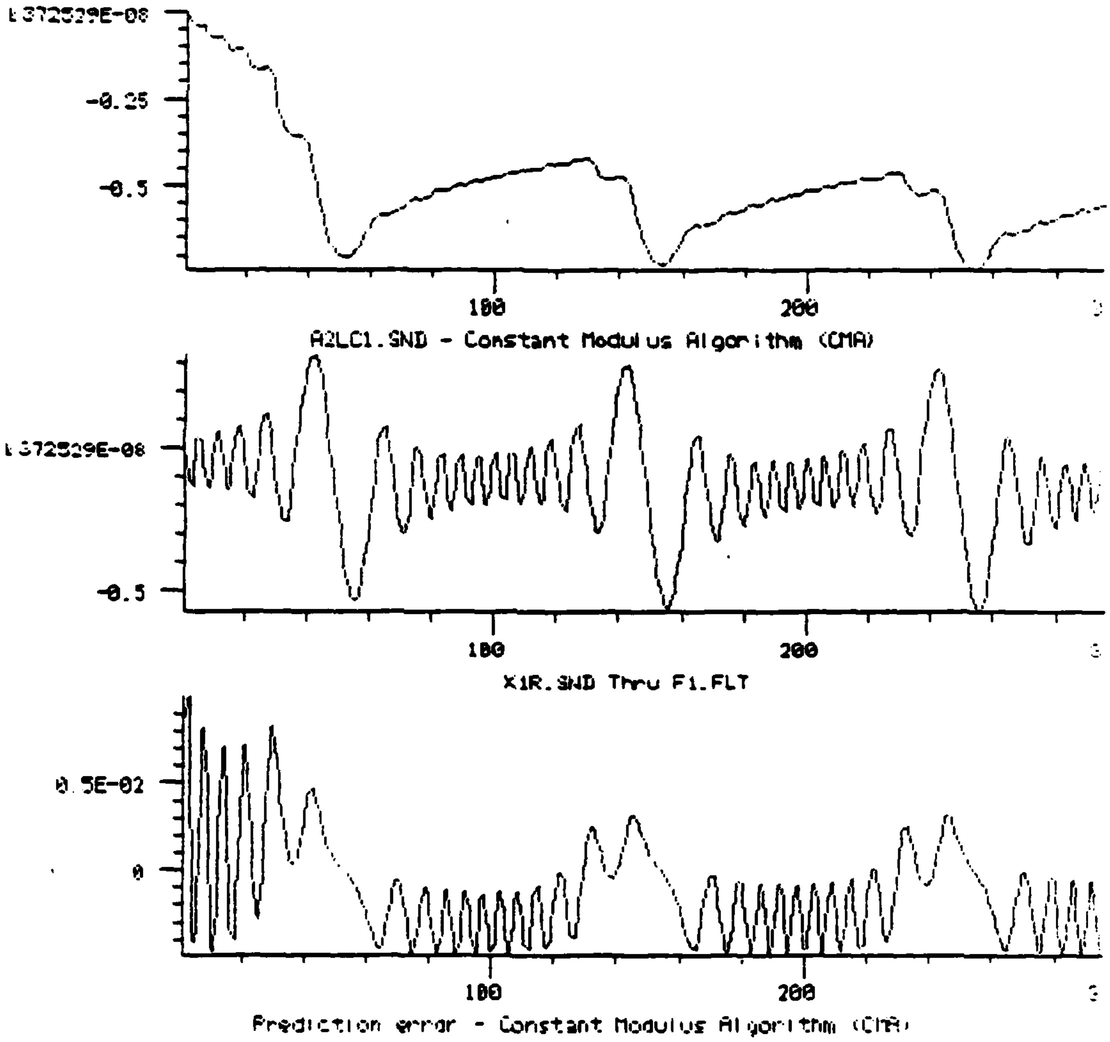

μ = 100
0
-0.25
-0.5
-0.75
-0.98919
1000
A2LC1.SND - Constant Modulus Algorithm (CMA)

Figure 8



1.372529E-08
-0.25
-0.5
100
200
A2LC1.SND - Constant Modulus Algorithm (CMA)
1.372529E-08
-0.5
100
200
X1R.SND Thru F1.FLT
0.5E-02
0
100
200
Prediction error - Constant Modulus Algorithm (CMA)

Figure 9

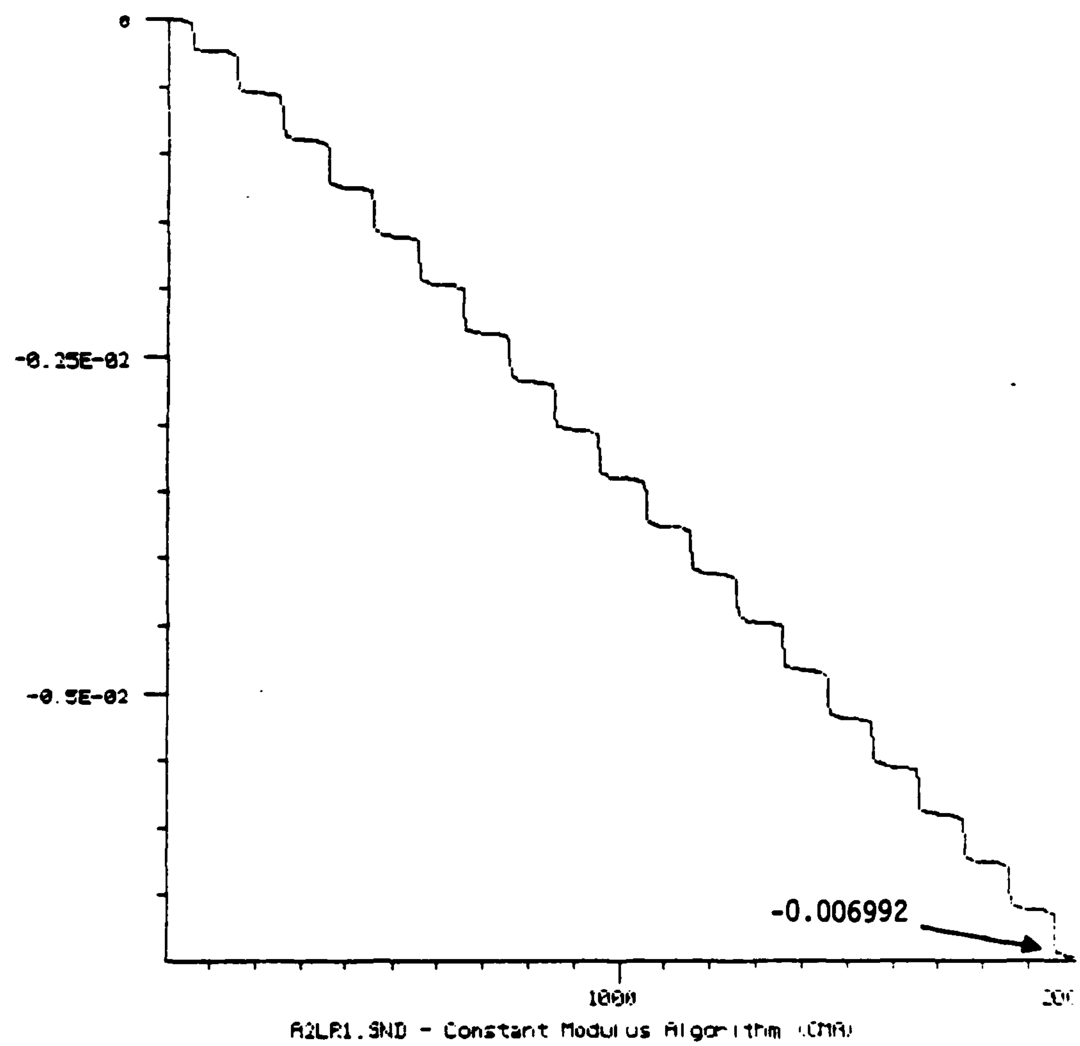

0
-0.25E-02
-0.5E-02
-0.006992
1000
A2LR1.SND - Constant Modulus Algorithm (CMA)

Figure 10

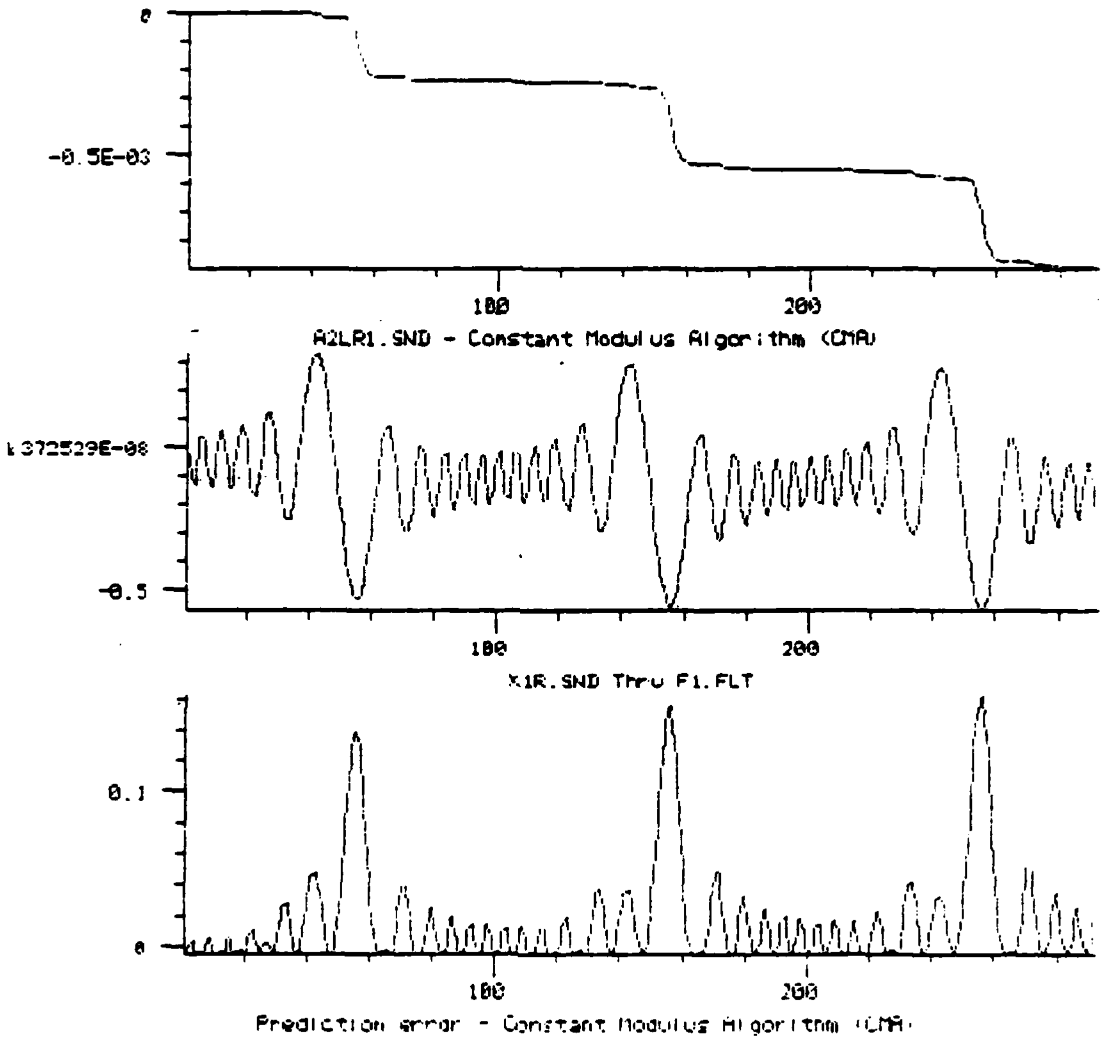
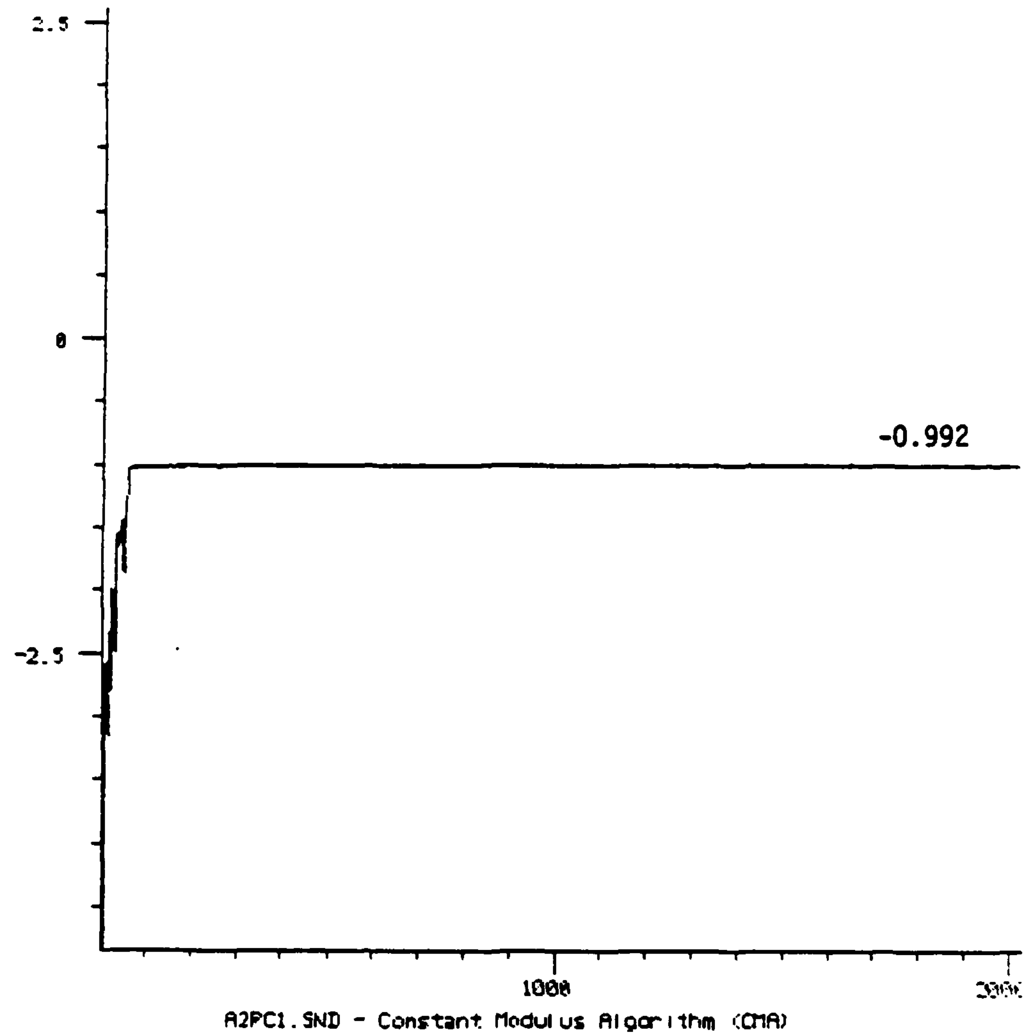

0
-0.5E-03
100
200
A2LR1.SND - Constant Modulus Algorithm (CMA)
-0.5
100
200
X1R.SND Thru F1.FLT
0.1
0
100
200
Prediction error - Constant Modulus Algorithm (CMA)

Figure 11



2.5
0
-2.5
-0.992
1000
2000
A2PC1.SND - Constant Modulus Algorithm (CMA)

Figure 12

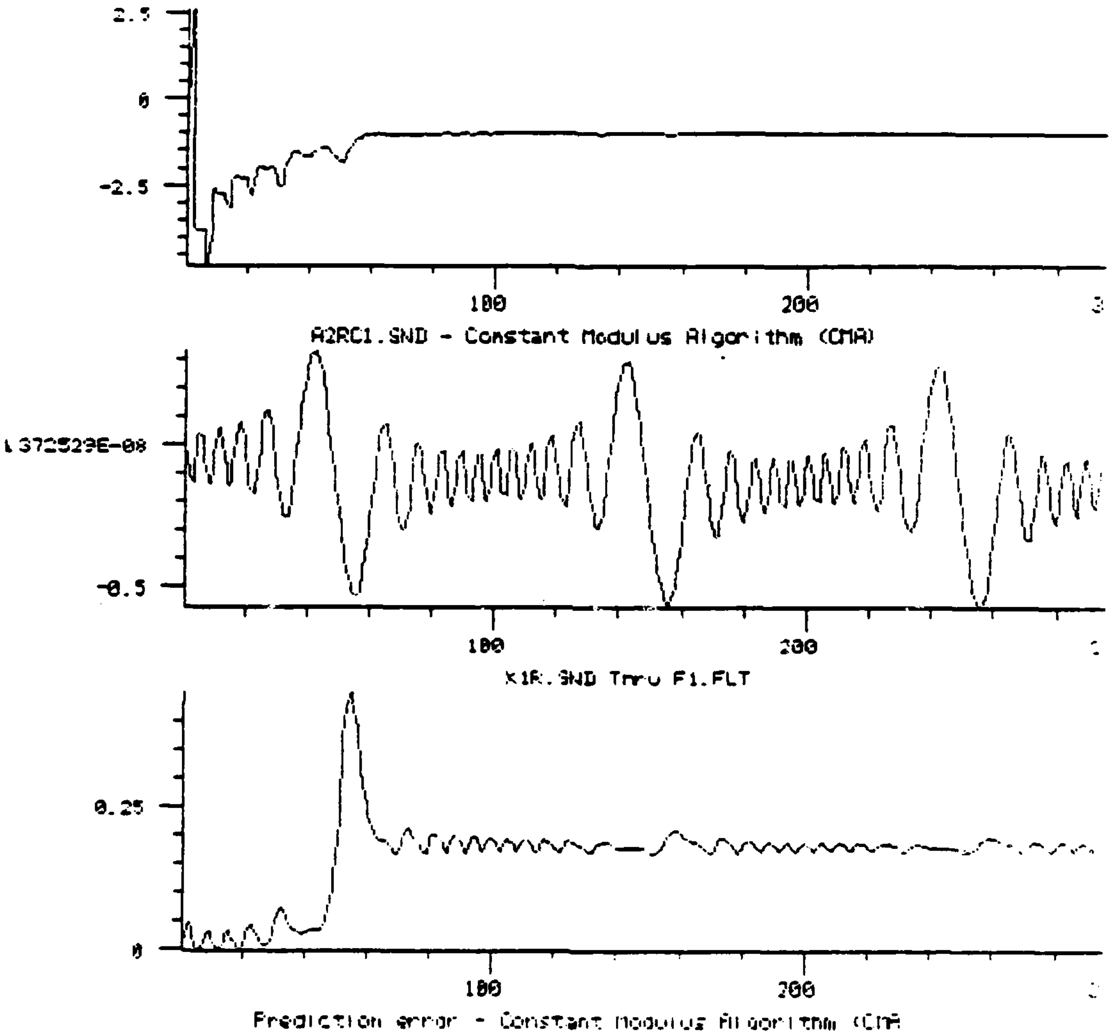

2.5
0
-2.5
100
200
A2RC1.SND - Constant Modulus Algorithm (CMA)
1.372529E-08
-0.5
100
200
X1R.SND Thru F1.FLT
0.25
0
100
200
Prediction error - Constant Modulus Algorithm (CMA

Figure 13

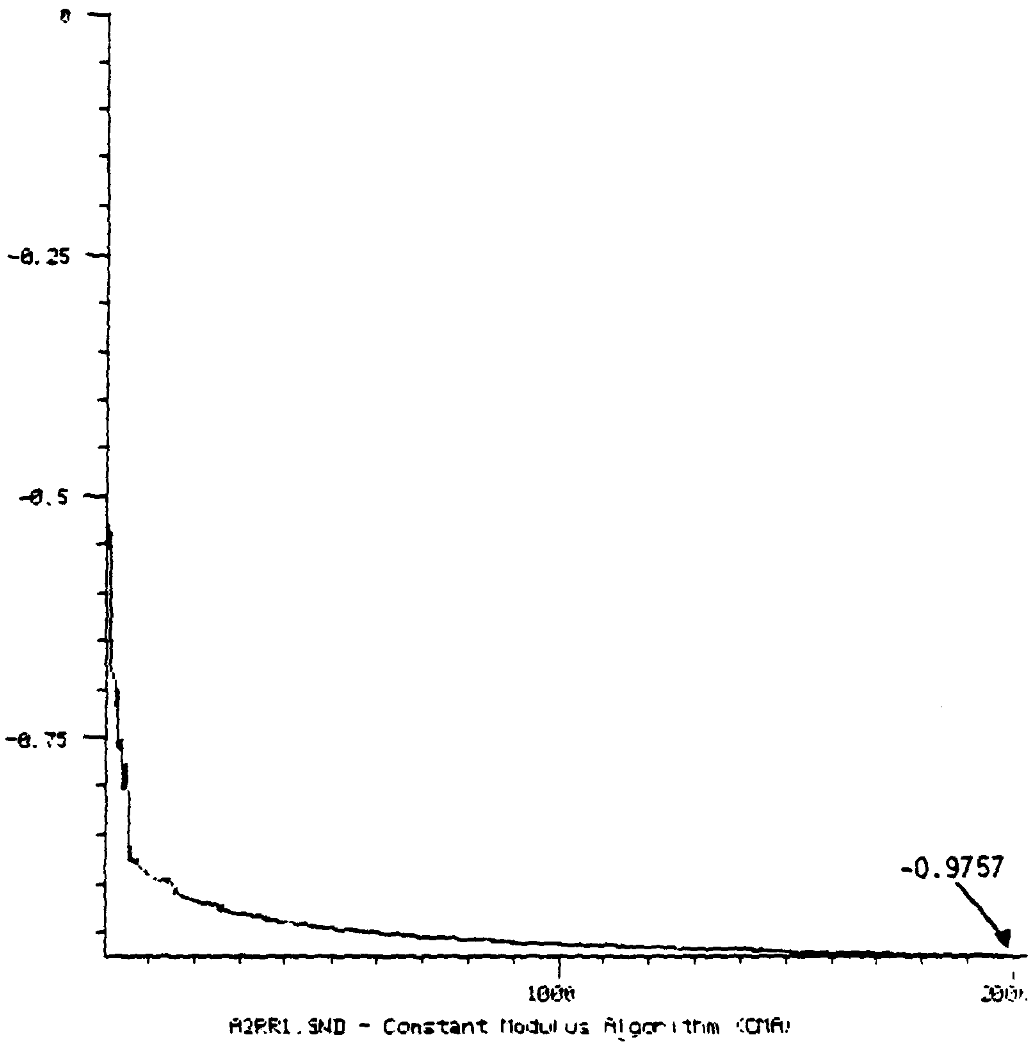
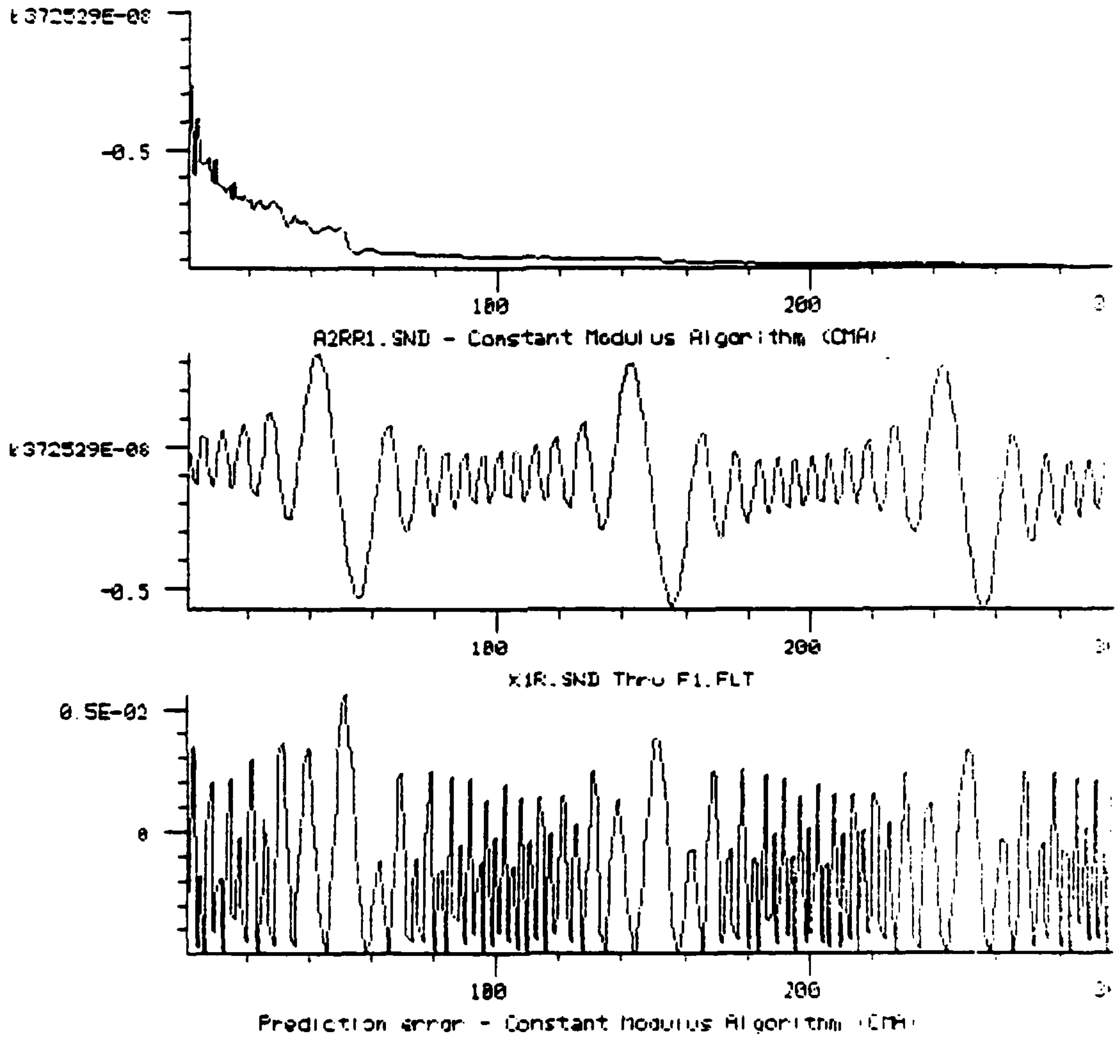

0
-0.25
-0.5
-0.75
-0.9757
1000
2000
A2RR1.SND - Constant Modulus Algorithm (CMA)

Figure 14



t 372529E-08
-0.5
100
200
A2RR1.SND - Constant Modulus Algorithm (CMA)
t 372529E-06
-0.5
100
200
X1R.SND Thru F1.FLT
0.5E-02
0
100
200
Prediction error - Constant Modulus Algorithm (CMA)

Figure 15

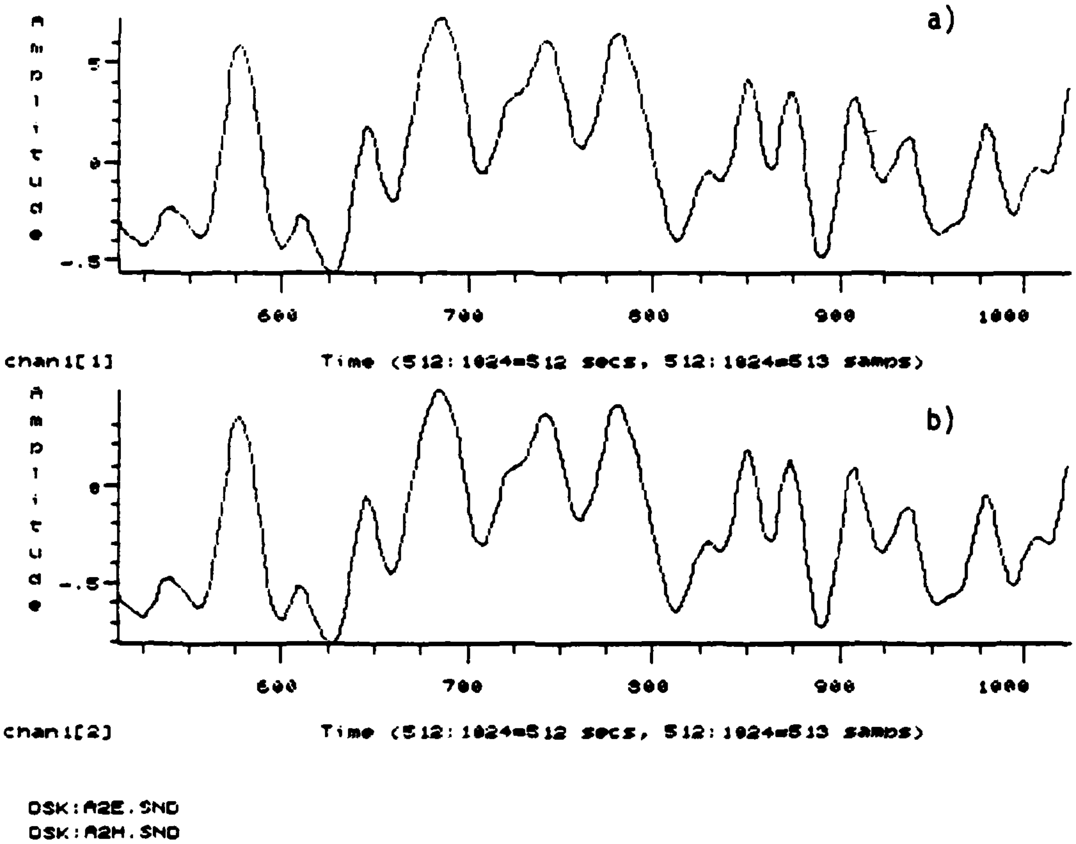


Figure 16

a) Angle modulation waveform

b) Demodulation of channel-distorted signal

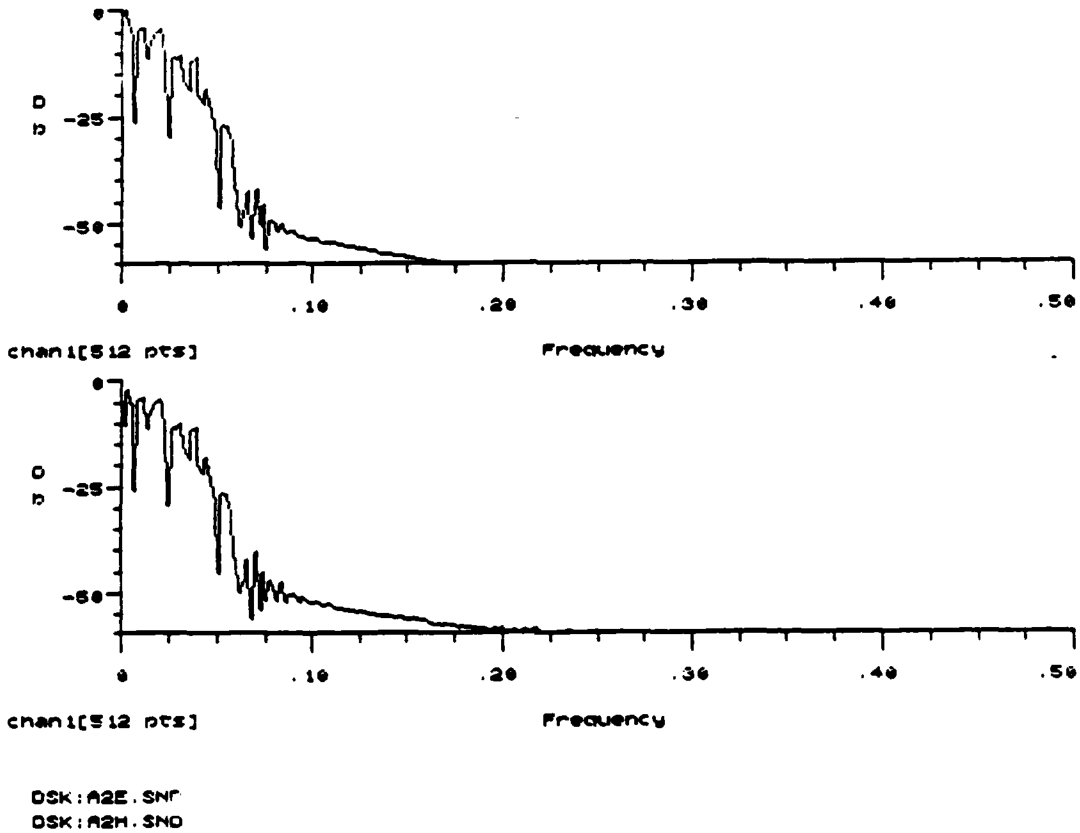
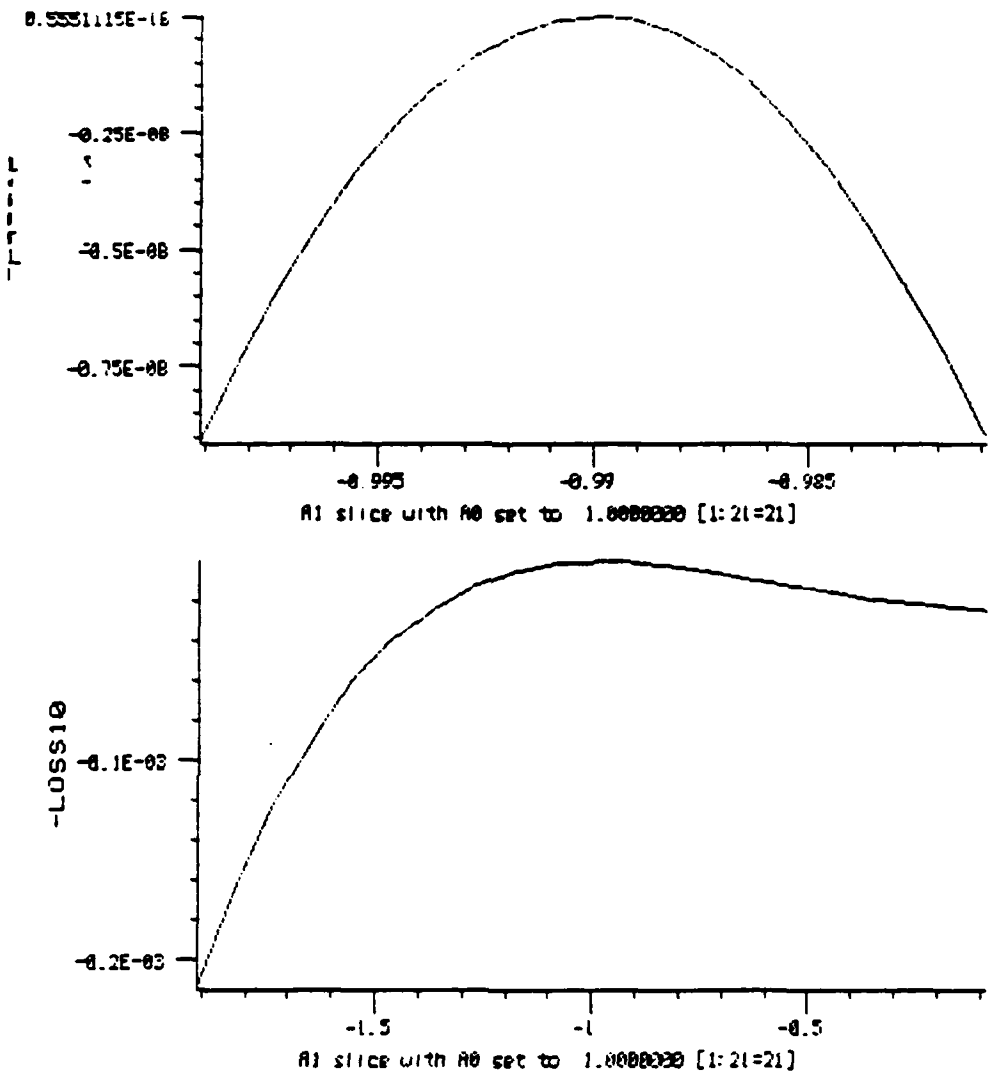

0
D
b
-25
-50
0
.10
.20
.30
.40
.50
chan1[512 pts]
Frequency
0
D
b
-25
-50
0
.10
.20
.30
.40
.50
chan1[512 pts]
Frequency
DSK:A2E.SNP
DSK:A2H.SNO

Figure 17



0.5551115E-16
-0.25E-08
-0.5E-08
-0.75E-08
-0.995
-0.99
-0.985
A1 slice with A0 set to 1.0000000 [1:21=21]
-LOSS10
-0.1E-03
-0.2E-03
-1.5
-1
-0.5
A1 slice with A0 set to 1.0000000 [1:21=21]

Figure 18

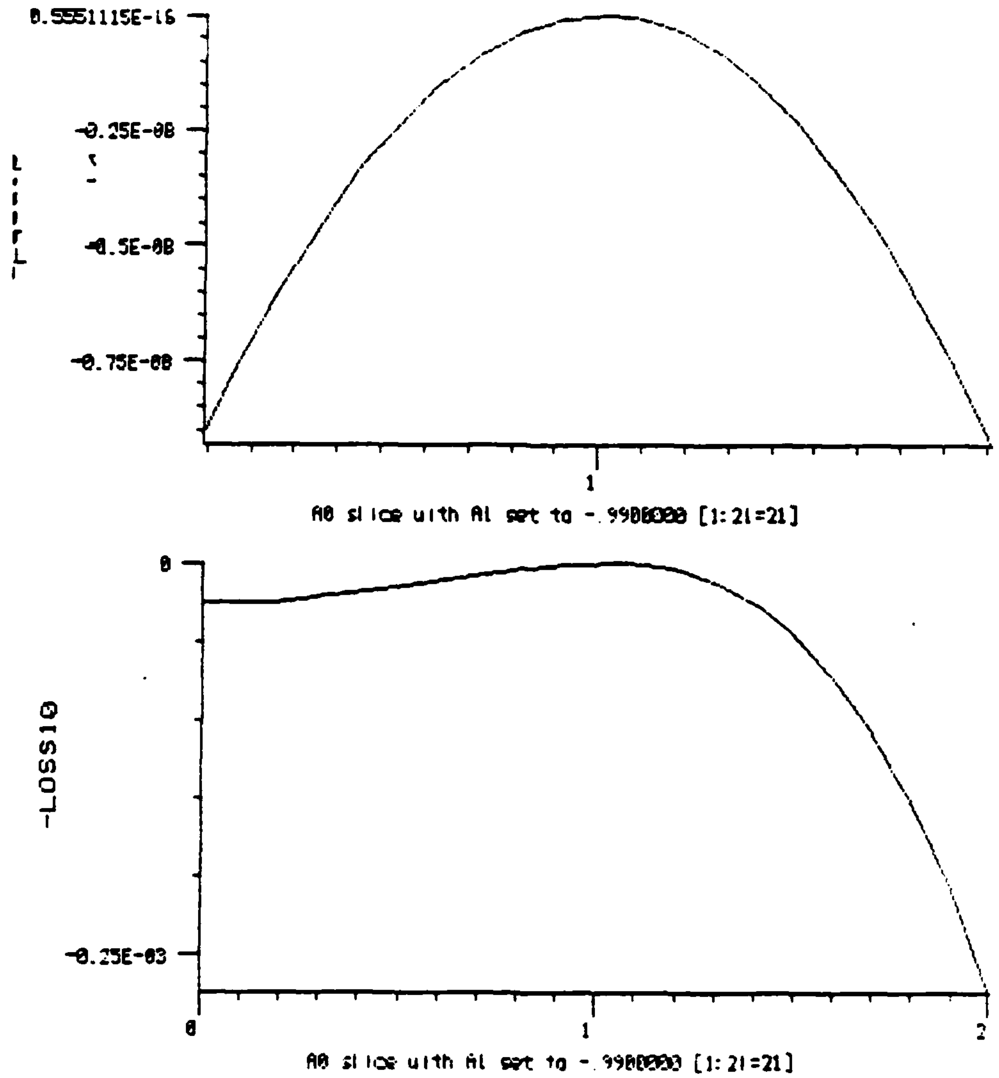

0.5551115E-16
-0.25E-08
-0.5E-08
-0.75E-08
1
A0 slice with A1 set to -.9900000 [1:21=21]
0
-LOSS10
-0.25E-03
0
1
2
A0 slice with A1 set to -.9900000 [1:21=21]

Figure 19

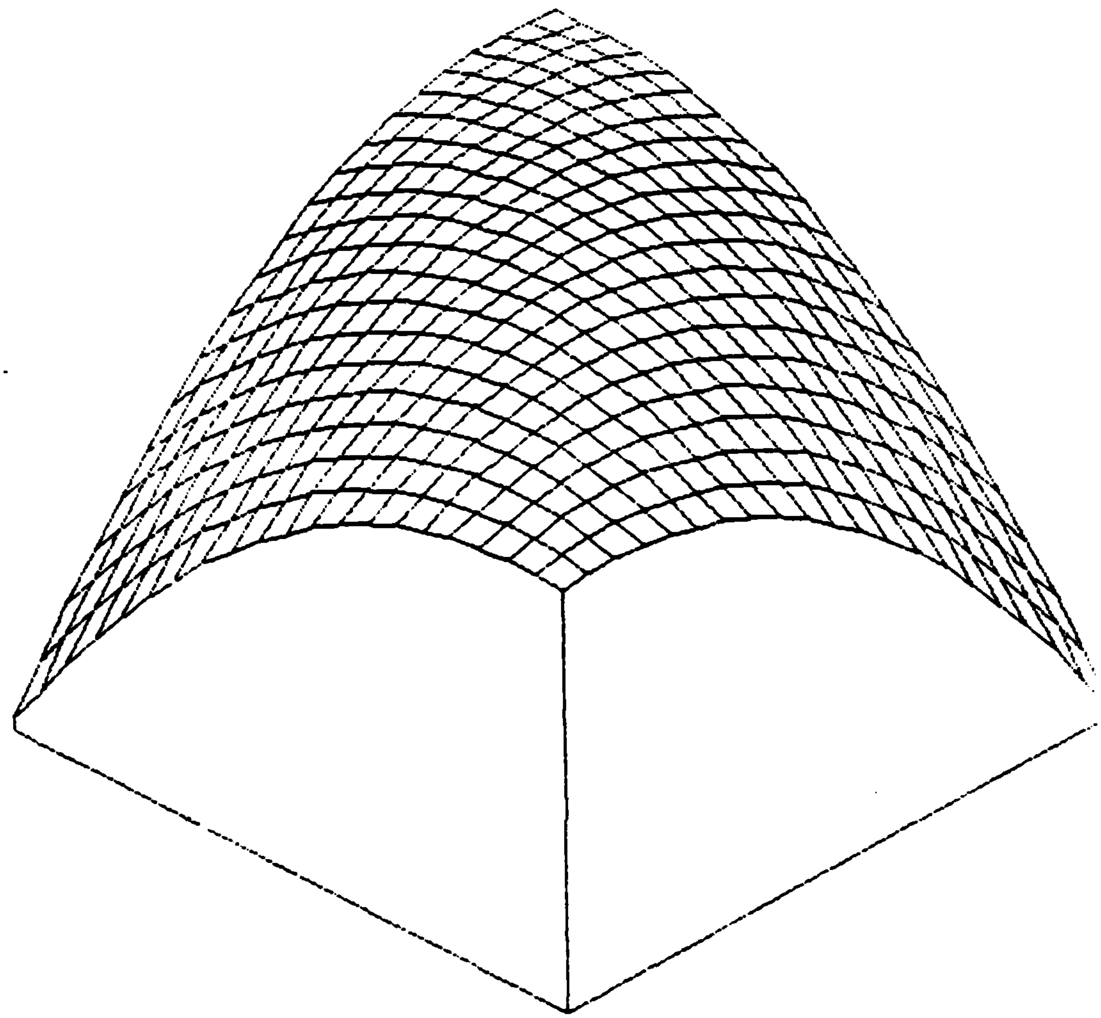


Figure 20

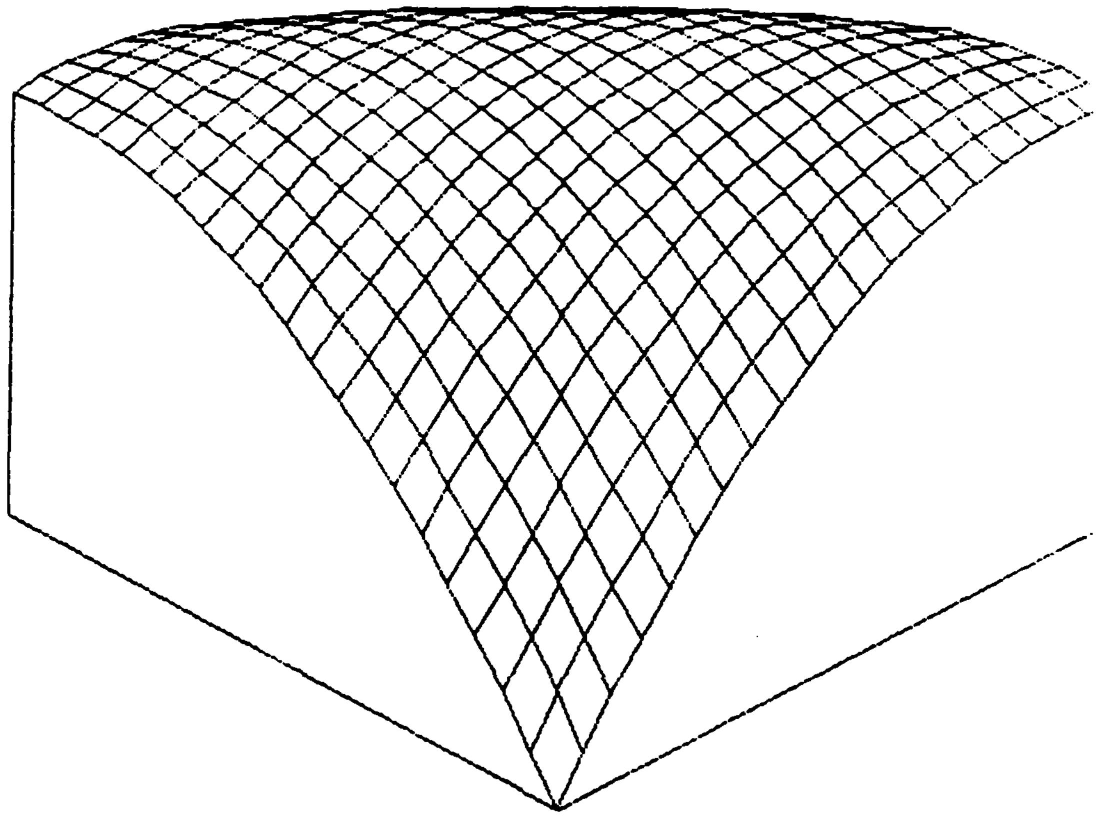


Figure 21

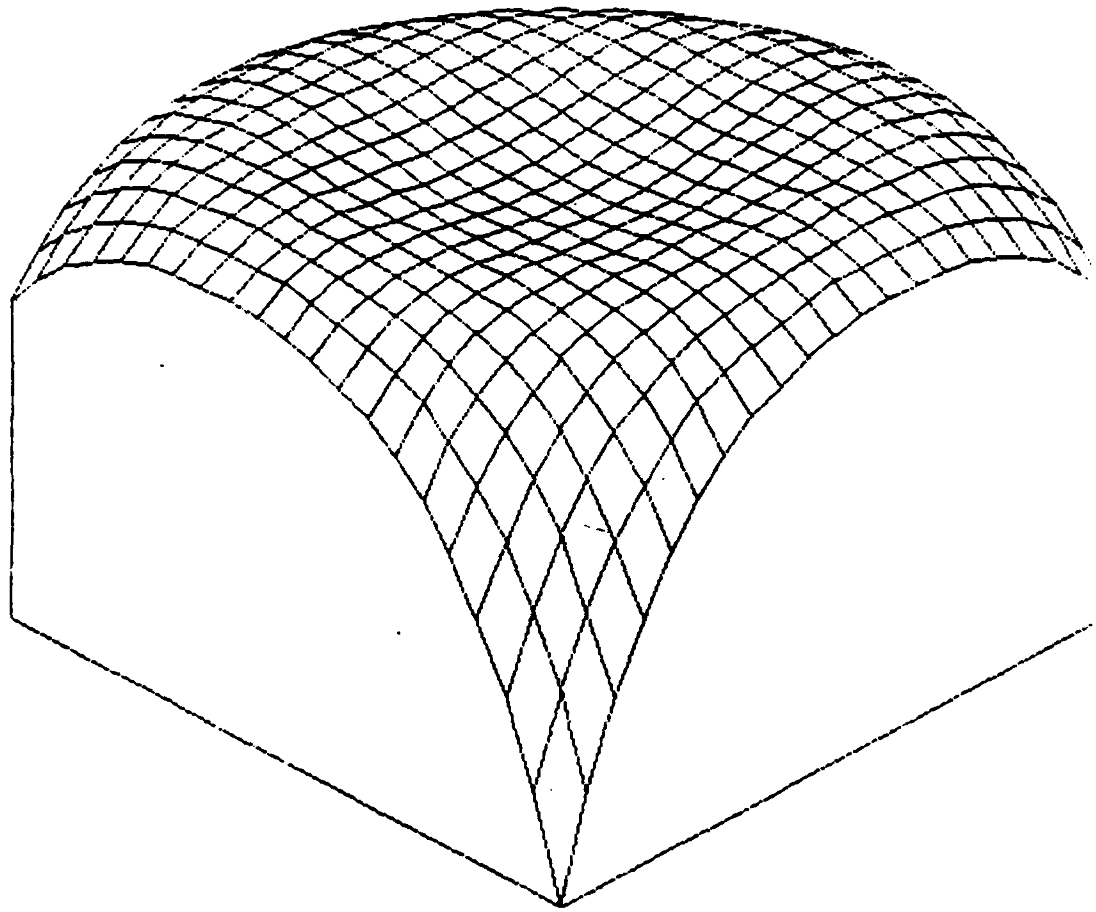


Figure 22

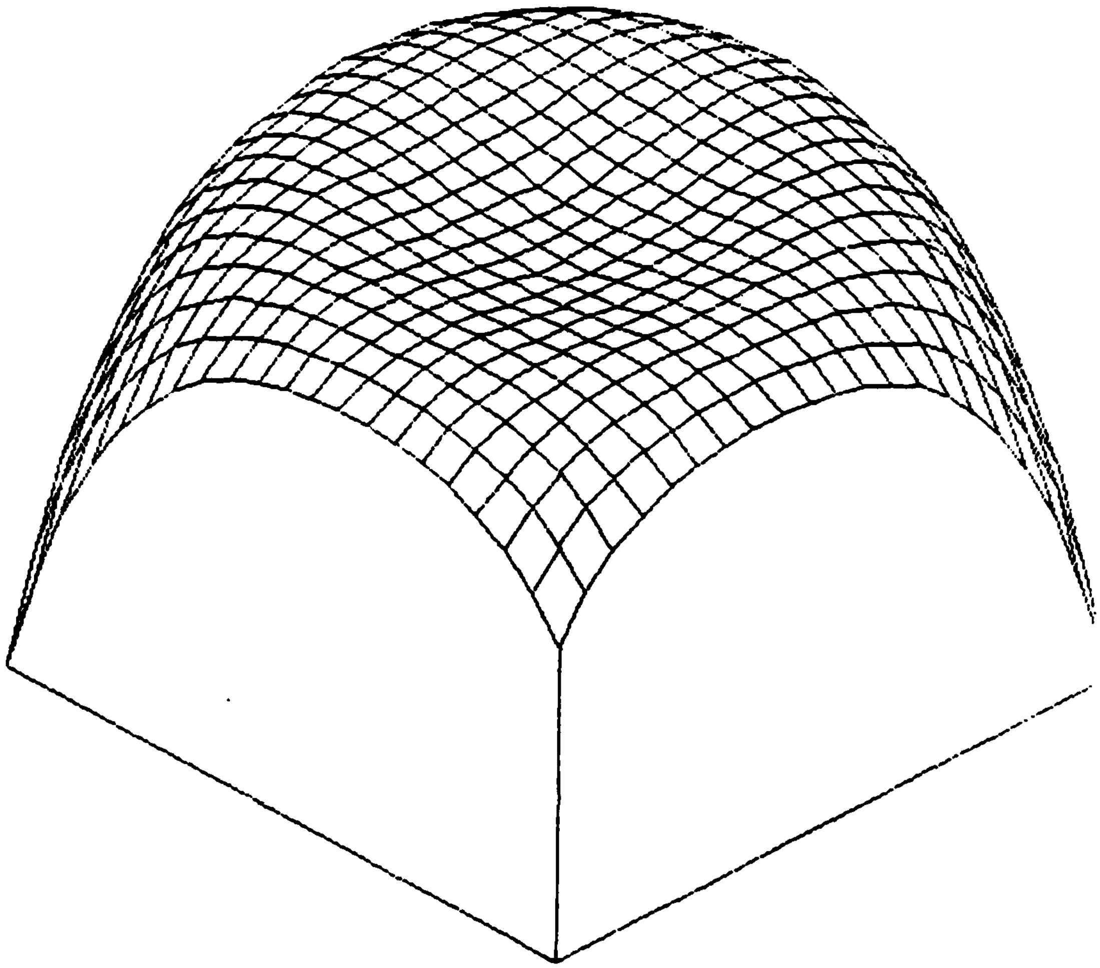


Figure 23

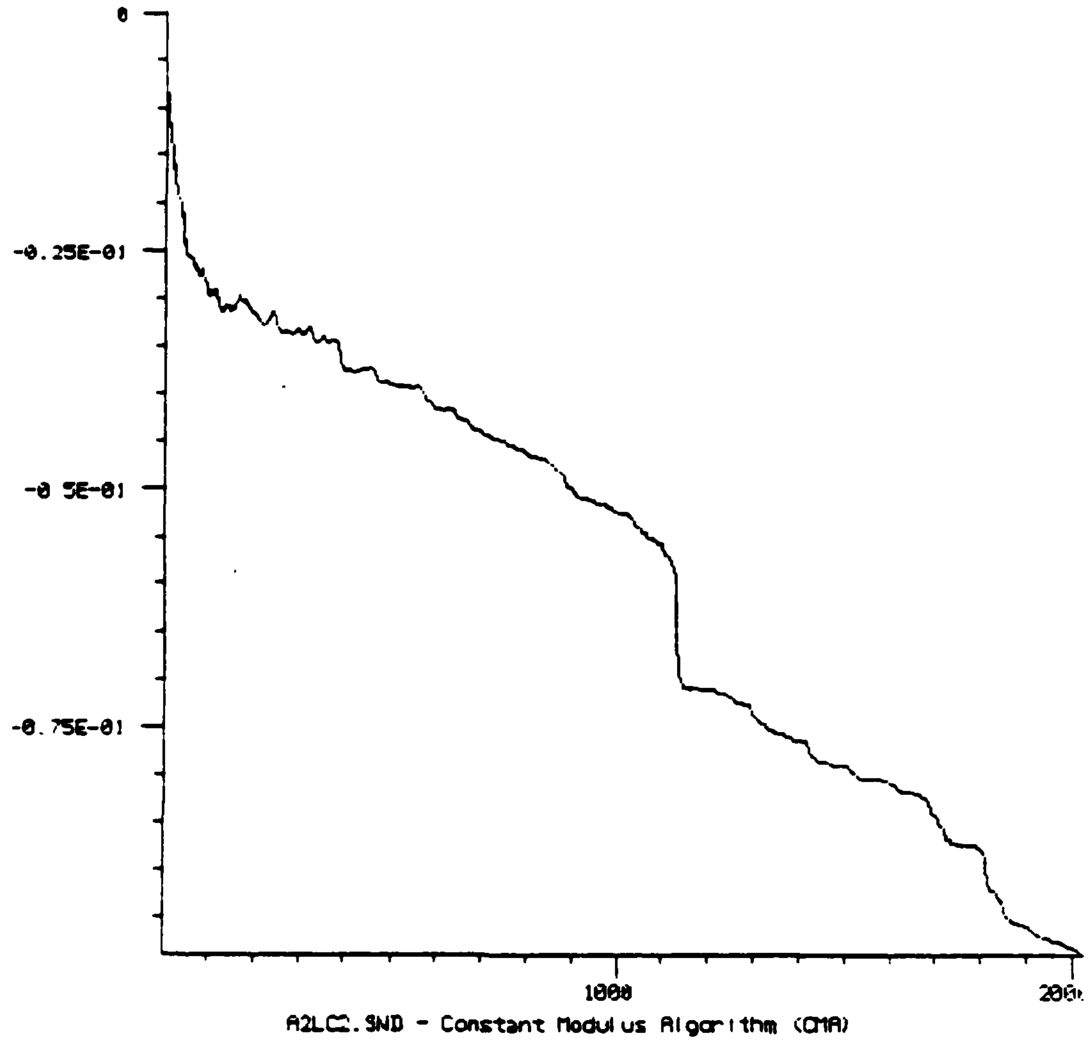
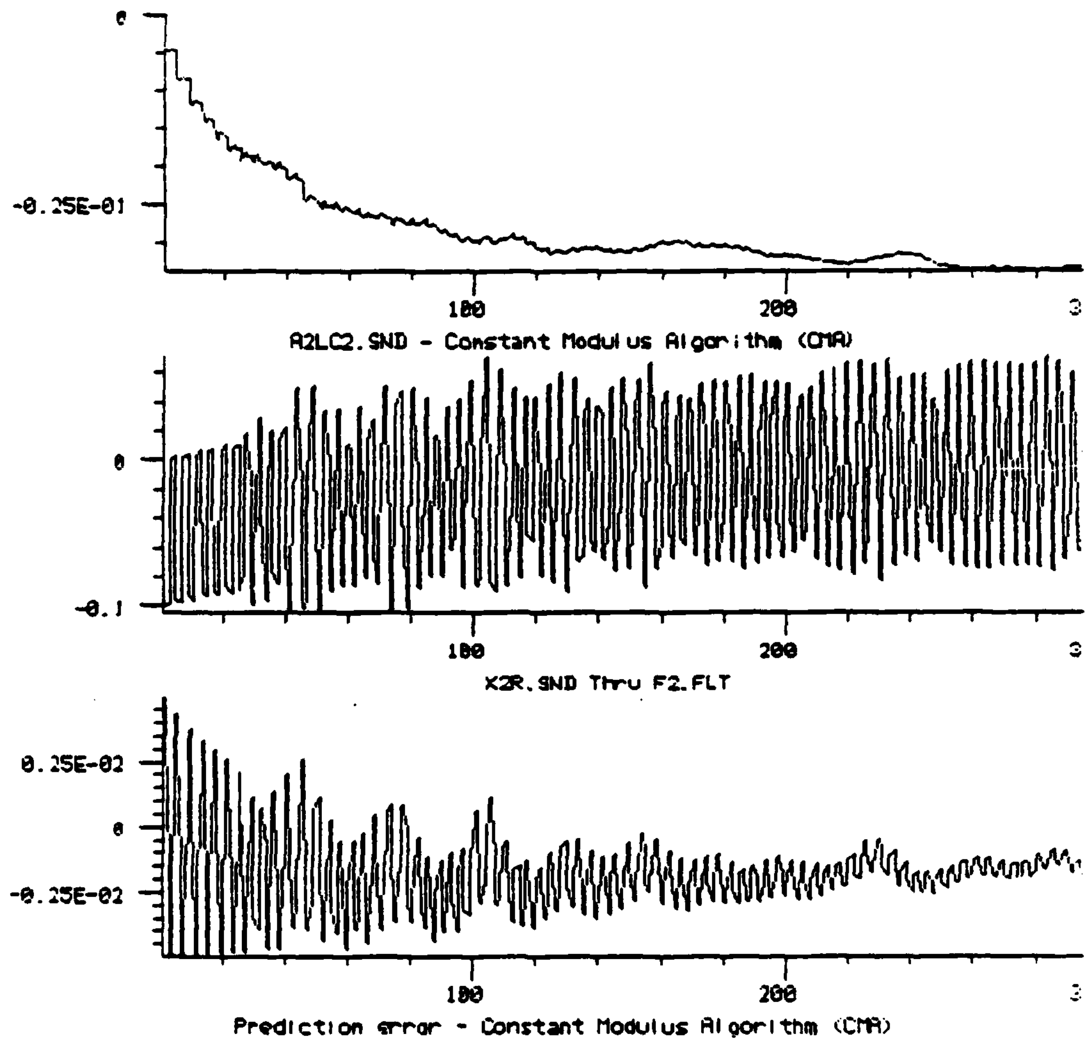

0
-0.25E-01
-0 5E-01
-0.75E-01
1000
200
A2LC2.SND - Constant Modulus Algorithm (CMA)

Figure 24



0
-0.25E-01
100
200
3
A2LC2.SND - Constant Modulus Algorithm (CMA)
0
-0.1
100
200
3
X2R.SND Thru F2.FLT
0.25E-02
0
-0.25E-02
100
200
3
Prediction error - Constant Modulus Algorithm (CMA)

Figure 25

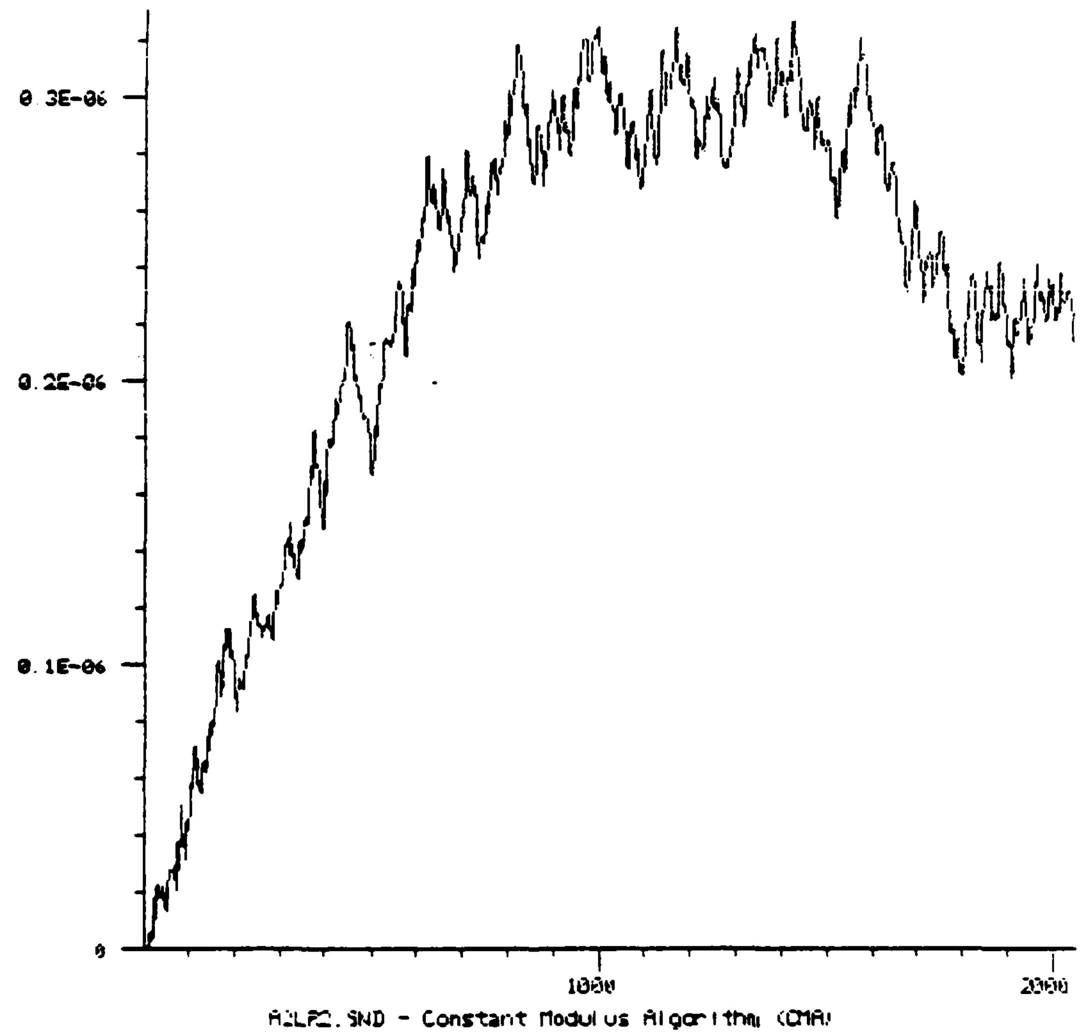

0.3E-06
0.2E-06
0.1E-06
0
1000
2000
A2LP2.SND - Constant Modulus Algorithm (CMA)

Figure 26

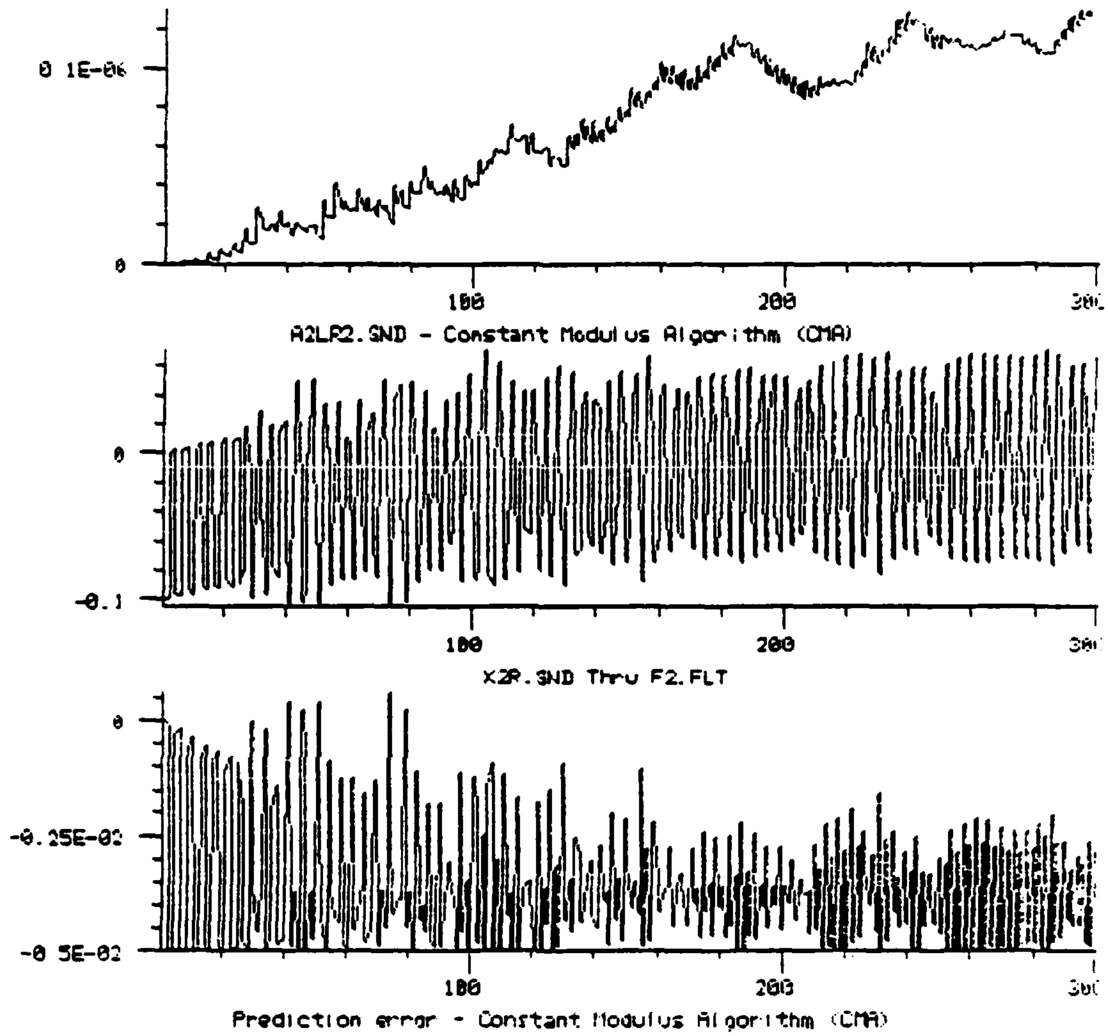

0 1E-06
0
100
200
300
A2LR2.SND - Constant Modulus Algorithm (CMA)
0
-0.1
100
200
300
X2R.SND Thru F2.FLT
0
-0.25E-02
-0 5E-02
100
200
300
Prediction error - Constant Modulus Algorithm (CMA)

Figure 27

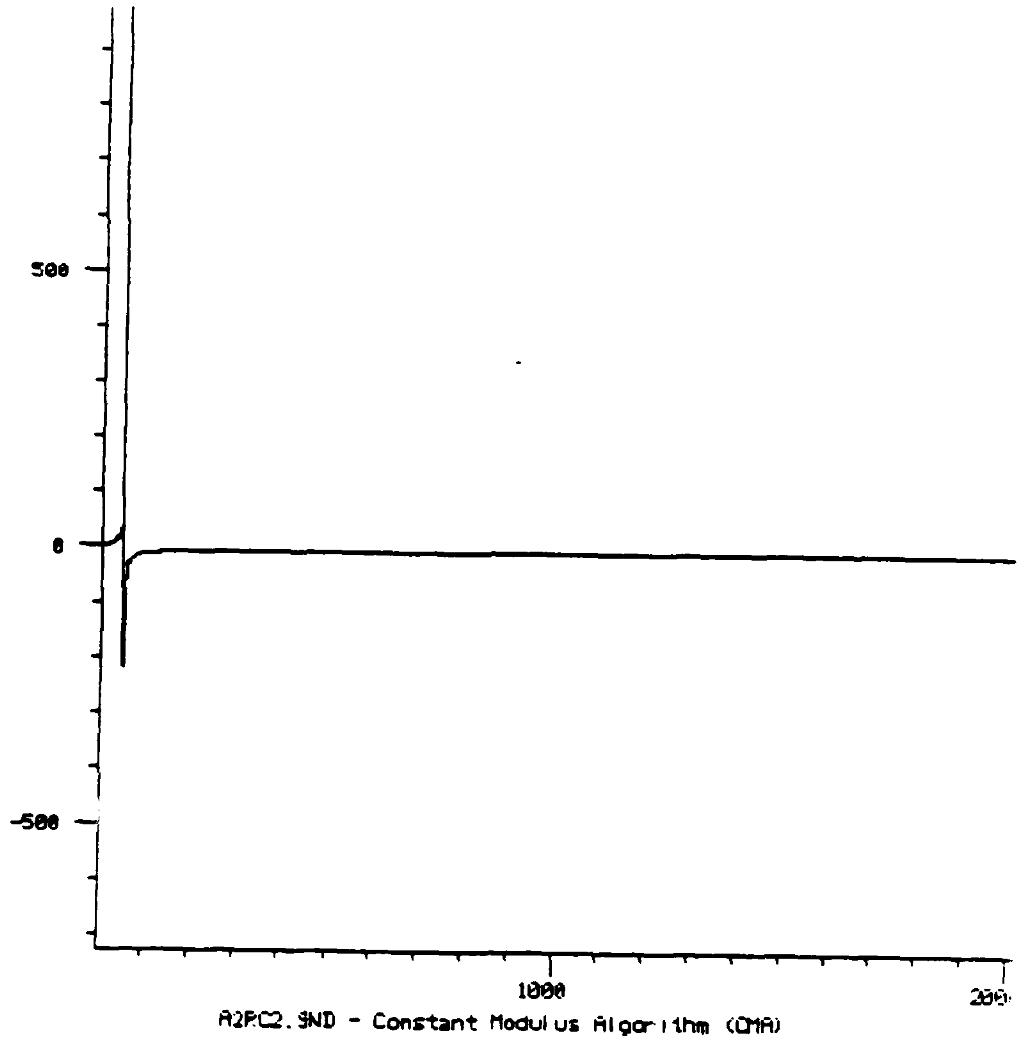
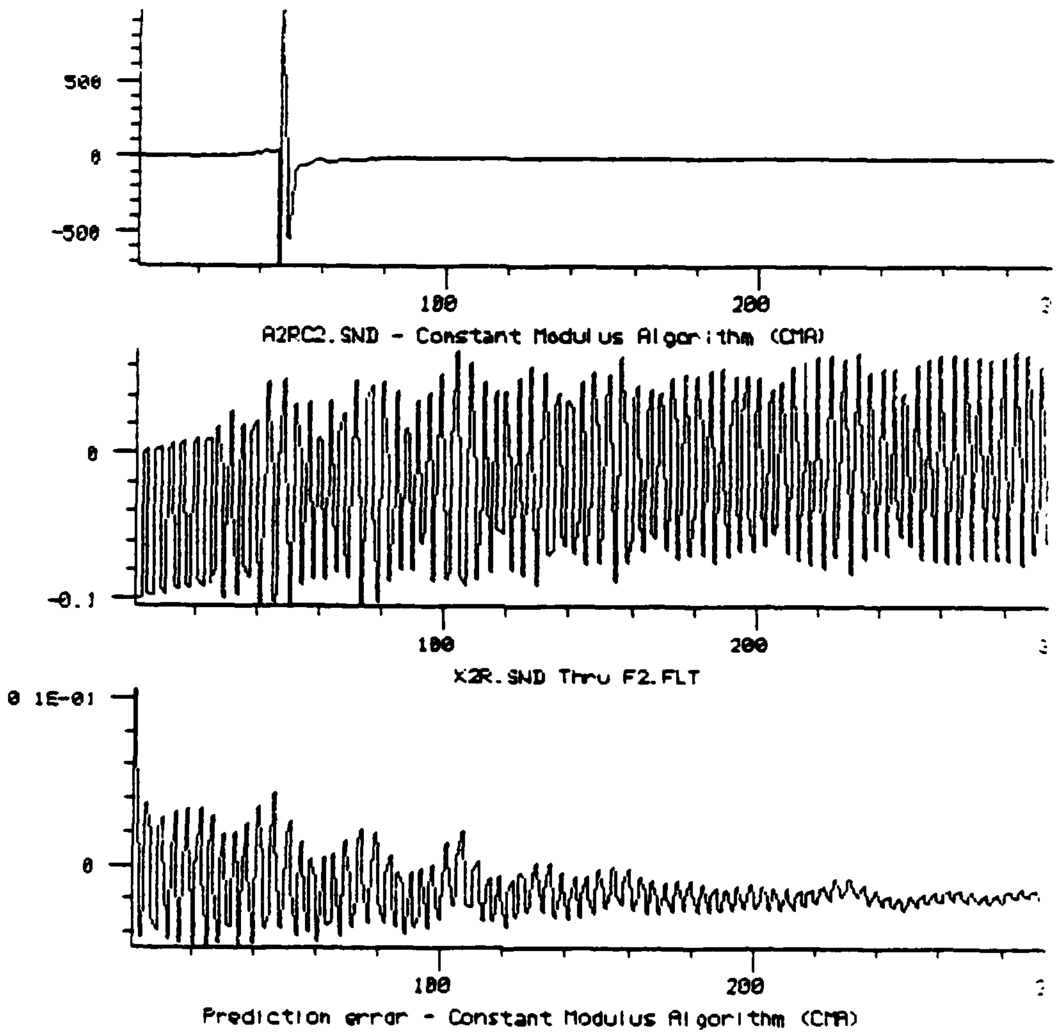


Figure 28



500
0
-500
100
200
A2RC2.SND - Constant Modulus Algorithm (CMA)
0
-0.1
100
200
X2R.SND Thru F2.FLT
0.1E-01
0
100
200
Prediction error - Constant Modulus Algorithm (CMA)

Figure 29

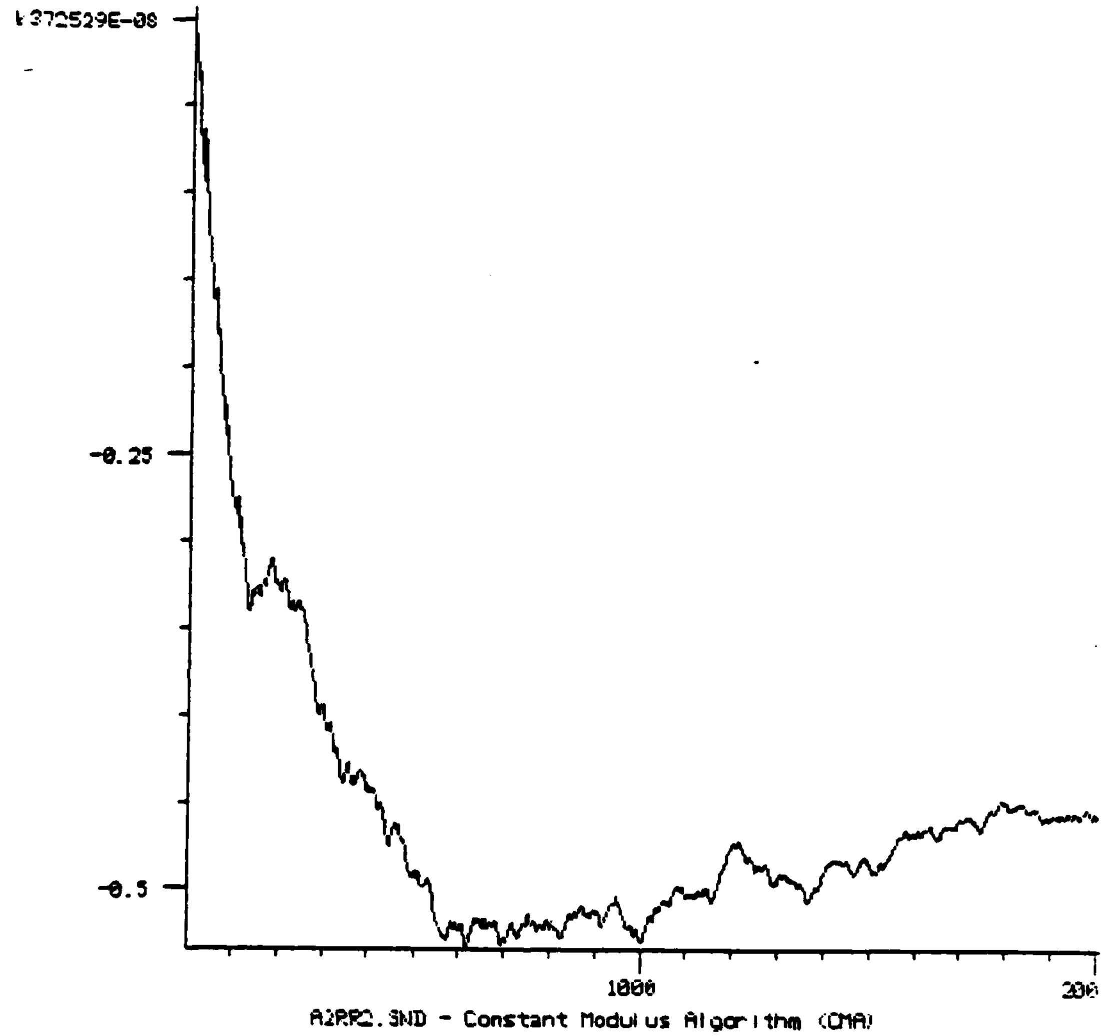
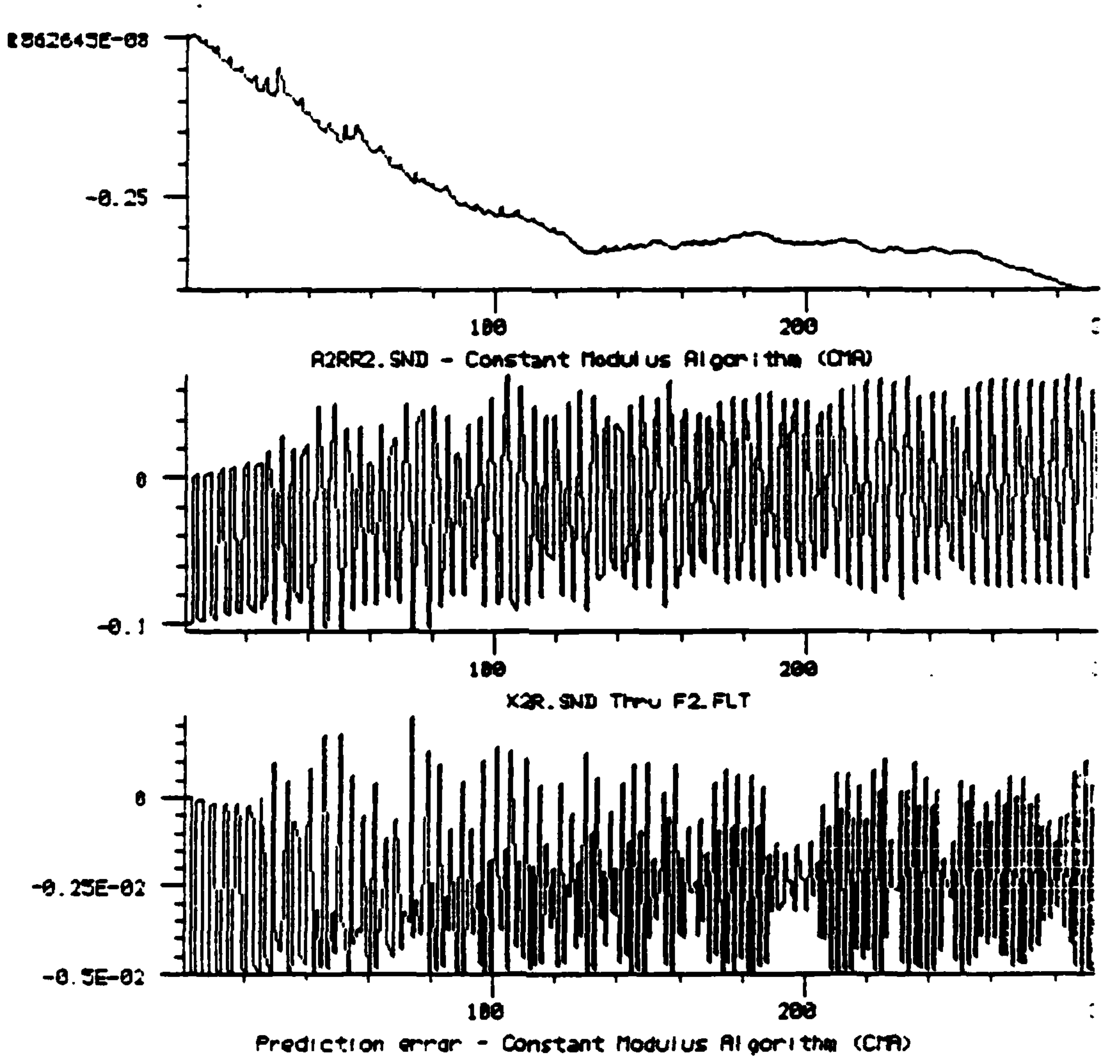


Figure 30



.8562645E-08
-0.25
100
200
A2RR2.SND - Constant Modulus Algorithm (CMA)
0
-0.1
100
200
X2R.SND Thru F2.FLT
0
-0.25E-02
-0.5E-02
100
200
Prediction error - Constant Modulus Algorithm (CMA)

Figure 31

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.

END

FILMED

2-86

DTIC