

DTIC FILE COPY

4

RADC-TR-89-306
In-House Report
November 1989



AD-A219 328

A PARAMETRIC DETECTION APPROACH USING MULTICHANNEL PROCESSES

James H. Michels

DTIC
ELECTE
MAR 19 1990
S E D

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

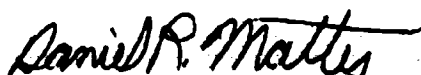
ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

90 08 16 026

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-89-306 has been reviewed and is approved for publication.

APPROVED:



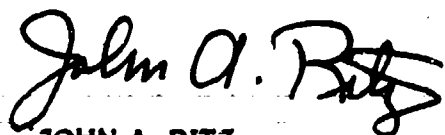
DANIEL R. MATTER, Capt, USAF
Asst Chief, Surveillance Technology Division
Directorate of Surveillance

APPROVED:



JAMES W. YOUNGBERG, Lt Col, USAF
Assistant Director
Directorate of Surveillance

FOR THE COMMANDER:



JOHN A. RITZ
Directorate of Plans and Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (OCTM) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188		
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) RADC-TR-89-306			5. MONITORING ORGANIZATION REPORT NUMBER(S) N/A			
6a. NAME OF PERFORMING ORGANIZATION Rome Air Development Center		6b. OFFICE SYMBOL (if applicable) OCTM	7a. NAME OF MONITORING ORGANIZATION Rome Air Development Center (OCTM)			
6c. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700			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) OCTM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N/A			
8c. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700			10. SOURCE OF FUNDING NUMBERS			
			PROGRAM ELEMENT NO. 62702F	PROJECT NO. 4506	TASK NO. 17	WORK UNIT ACCESSION NO. 67
11. TITLE (Include Security Classification) A PARAMETRIC DETECTION APPROACH USING MULTICHANNEL PROCESSES						
12. PERSONAL AUTHOR(S) James H. Michels						
13a. TYPE OF REPORT In-House		13b. TIME COVERED FROM Sep 88 to Aug 89		14. DATE OF REPORT (Year, Month, Day) November 1989		
15. PAGE COUNT 76						
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	Parametric Detection, Adaptive Filtering) Multichannel Detection, Generalized Likelihood Ratio, Parameter Estimation, Prediction Error Filtering (66)			
17	09					
19. ABSTRACT (Continue on reverse if necessary and identify by block number)						
This report considers the binary multichannel detection problem for an unknown random signal vector in additive nonwhite interference plus white Gaussian noise. A generalized likelihood ratio is derived based on the vector error residuals from multichannel prediction error filters designed as minimum mean squared error estimates under each hypothesis. The observation processes are considered to have an arbitrary correlation in time and across channels. The report outlines a research investigation currently in progress.						
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED			
22a. NAME OF RESPONSIBLE INDIVIDUAL JAMES H. MICHELS			22b. TELEPHONE (Include Area Code) (315) 330-4431		22c. OFFICE SYMBOL RADC(OCTM)	

DD Form 1473, JUN 86

Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

ACKNOWLEDGEMENTS

The author wishes to acknowledge the helpful guidance and encouragement of Drs. R. Srinivasan, P. Varshney and D. Weiner. Special thanks to Ms. Kimberly Swiss for her skillful preparation of this report.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Code	
Dist	Avail and/or Special
A-1	



Contents

I.	INTRODUCTION	
	A. BACKGROUND	1
	B. RESEARCH OBJECTIVE	5
II.	THE DETECTION PROBLEM	8
III.	AUTOREGRESSIVE PROCESS MODELS	
	A. DEFINITION OF THE AR PROCESS	10
	B. THE YULE-WALKER EQUATION	12
IV.	DATA TRANSFORMATION VIA LINEAR PREDICTION	14
V.	MULTI-CHANNEL LIKELIHOOD RATIO	
	A. DERIVATION	21
	B. SPECIAL CASE OF INCOHERENT INTEGRATION	27
VI.	LIKELIHOOD RATIO IMPLEMENTATION SCHEME	
	A. SYSTEM ARCHITECTURE	29
	B. FILTER COEFFICIENT DETERMINATION	29
	C. ERROR VARIANCE DETERMINATION	32
VII.	PROPOSED INVESTIGATION	
	A. STATIONARY GAUSSIAN PROCESS SYNTHESIS	37
	B. ANALYTIC PERFORMANCE ANALYSIS	41
	C. MONTE CARLO ANALYSIS (STATIONARY PROCESSES)	43
	D. POSSIBLE EXTENSIONS	43
	REFERENCES	45
	APPENDICES	
	A	53
	B	56
	C	60

I INTRODUCTION

A. BACKGROUND

The detection of a known deterministic signal in additive Gaussian noise is solved using a correlator or matched filter [45, 55, 62, 12 Part I]. For the case of a random signal in white noise, a two stage process is utilized to first estimate the signal and then correlate this estimate with the received waveform. This detection scheme is referred to as the estimator-correlator [47-50, 53, 12 Part III]. The advantage of this later approach is its practical implementation structure; ie, the best obtainable estimate of the signal is used in the correlation stage. As noted by Kailath [54], however, the method described in [50] fails when the signal and noise are correlated and when signals have nonzero means. These limitations were overcome for continuous-time processes by Kailath [53, 56] using the innovations representation. In addition, the approach applied to non-Gaussian desired signal processes. This detection problem is expressed as

$$H_1: x(t) = s(t) + w(t) \quad (1-1a)$$

$$H_0: x(t) = w(t) \quad (1-1b)$$

where H_i $i=0,1$ are the hypotheses for signal absent and present, respectively; $s(t)$ is the (not necessarily Gaussian) random signal and $w(t)$ is Gaussian white noise. Under H_1 , the optimal MMSE causal estimate of $s(t)$ is

$$\hat{s}(t) = E [s(t) | x(t'), t' < t] \quad (1-2)$$

Using the definition of the innovations [57]; ie.

$$v(t) = x(t) - \hat{s}(t), \quad (1-3)$$

the innovations theorem for continuous-time processes (ie, the covariance of $v(t)$ and $w(t)$ are identical), and the property that $v(t)$ is Gaussian if $w(t)$ is Gaussian [53], Kailath transformed the detection problem of eq (1-1) to the equivalent detection problem:

$$H_1: x(t) = \hat{s}(t) + v(t) \quad (1-4a)$$

$$H_0: x(t) = v(t) \quad (1-4b)$$

In this form of the detection problem, $\hat{s}(t)$ is viewed as a conditionally-known signal and the likelihood ratio becomes [53]

$$LR = \exp \left[\int \hat{s}(t) x(t) dt - 1/2 \int \hat{s}^2(t) dt \right]. \quad (1-5)$$

This likelihood ratio has the same form as the estimator-correlator for the Gaussian signal in the white, Gaussian noise case except that the first integral in eq(1-5) is an Ito integral. Eq(1-5) is significant because it shows the structure of the non-Gaussian detection problem. In addition, although eq(1-5) was derived in [53] assuming $s(t)$ and $w(t)$ as statistically independent, this restriction was later [56,58] relaxed to "future $w(\cdot)$ independent of past $s(\cdot)$ "; ie. the signal could be correlated with past noise. It is interesting to note that the innovations approach does not use a Karhunen-Loeve expansion which requires $s(t)$ and $w(t)$ to be statistically independent. This later generalization indicates that the innovations approach may offer a robustness in detection problems involving feedback, multipath, etc.

Unlike the continuous-time case, discrete innovations processes do not retain the property of Gaussianity unless both $w(t)$ and $s(t)$ are Gaussian. In addition, the covariance of the innovations is not equivalent to that of the white noise [53,57]. Thus, the likelihood ratio for the discrete case does not have an estimator correlator structure to provide a test statistic. For this case, a likelihood ratio using discrete-time innovations processes was developed [4, 6]. In [4], it is shown that for non-Gaussian processes, this likelihood ratio, called the Innovations-Based Detection Algorithm (IBDA), is a close approximation to the likelihood ratio for this detection problem. Furthermore, the structure of the likelihood ratio is obtained by assuming a specific parametric model of the observation processes for each hypothesis and designing a prediction error filtering stage for each. Adaptive algorithms are then used to estimate the unknown model parameters. Since the order of the model under H_1 is greater

than that under H_0 (and with possibly distinct coefficients) the error output from the filter designed for H_0 given that H_1 is true will be greater than the error output from the filter designed for H_1 . These error differences are used in the likelihood ratio to raise or lower its value relative to the threshold level. In [59], it is shown that this likelihood ratio is a generalization of several important special cases. These include (1) the detection of a deterministic signal in additive white Gaussian noise, (2) the detection of a non-white Gaussian signal process in additive white Gaussian noise, (3) the moving target detection (MTD) algorithm for the detection of a deterministic signal in non-white Gaussian noise [60, 61] and (4) the detection of a non-white Gaussian signal process in additive non-white Gaussian noise [62]. It is also shown [59] that the IBDA contains the algorithm developed [63] for the detection of a deterministic signal in additive non-white Gaussian noise of unknown correlation statistics. An experimental investigation using two implementations of the IBDA was conducted in [8] for an airport surveillance radar application and performance comparisons made with three MTD algorithms.

Several early analyses involving model-based parametric detection approaches are presented in [40, 44, 62, 64] with special notice given to [51]. Most of the interest, however, appears to have taken place within the past few years [4, 6, 8, 14, 21, 23, 41, 43, 45, 63]. The principal advantage of characterizing the observation processes for each hypothesis via a parametric model is that well known algorithms can be utilized to estimate the parameters. In [45], a likelihood ratio test was considered for two known autoregressive (AR) models. In [41], a more general formulation considers AR and autoregressive-moving average (ARMA) models to detect a Gaussian signal in white Gaussian noise, both with unknown statistics. Thus, the IBDA [4, 6] mentioned above considers the more general problem of detection on a non-Gaussian signal in non-white plus white noise. As noted in [16], however, different modeling approaches will generally yield differences in receiver performance. In addition, the problem of modeling observation processes in this detection scheme is more difficult than model-fitting a time series process. This is due to the fact that in the detection problem, one set of observation data is given and the problem is to determine which of the two filters is estimating the parameters properly. In [14, 16], Zhang investigated the detection performance improvement over the IBDA [6] in a radar application, when 'a priori' information was used to predetermine

the process parameters; specifically, a reference channel which provided data from range cells adjacent to the "test cell" was used to determine the filter coefficients under H_0 . A significant detection performance improvement over the IBDA was reported. It thus appears that performance improvements can be made in the model-based detection schemes through investigations of alternative algorithms and implementation schemes for the detection problem with unknown statistics.

In the development of the model-based detection approaches, the emphasis is placed on characterizing the observation data received under each hypothesis with approximate models. If the models fail to fit the physical processes, performance degradations will result. Model fitting of observation data via time series analysis has received considerable attention [19, 65, 66, 67, 68, 69, 71, 72]. The emphasis in the analysis proposed here will be the development of multichannel model-based detection approaches. Multichannel time series analyses have also been investigated [28, 35, 66 Part II, 67, 69, 71, 72] with emphasis in the areas of geophysics, biophysics and economics. Furthermore, multichannel algorithms for parameter estimation have received some attention [3, 9, 24, 25, 31, 70, 73]. The prime consideration for this analysis will be the investigation of the potential for improved detection performance of multichannel model-based detection. Thus, the applicability of multichannel time series models and the performance of multichannel parameter estimation algorithms to the detection problem proposed here is essential.

B. RESEARCH OBJECTIVE

The proposed investigation will consider the binary multichannel detection problem for an unknown random signal vector in additive non-white interference plus white Gaussian noise. The observation processes will be assumed to have an arbitrary correlation in time and space (ie, across the channels). Initially, Gaussian random processes will be considered, but a generalization to non-Gaussian processes will be developed.

The principal research objective is the investigation of multichannel model based detection methods utilizing estimation to determine their performance relative to the single channel case. In this approach to the detection problem, it is assumed that the underlying physical mechanisms which give rise to the observed random processes obtained under each hypothesis can be represented by a mathematical model which approximates its statistical characteristics. We therefore make a distinction between the model of the process (i.e., synthesis) and the estimation process (analysis). Specifically, we will give prime consideration to multichannel autoregressive processes [9, 24, 25, 28-31] as the process model description in the proposed investigation. The consideration of the multichannel approach is based on the contention that the coefficients of the AR processes are distinct for each of the two hypotheses (i.e., signal present or absent). The approach used in the model based detection method is the selection between the hypothesis based upon measures which are sensitive to the differences in the process coefficients for each hypothesis. Therefore, a likelihood ratio is sought which is sensitive to differences between the parameters of the processes. In this context, we view the multichannel processes arising from physical mechanisms such as those which may yield additional (although partially correlated) information about the processes. From the model based approach, this new information can be utilized to provide a better distinction between the process parameters under each hypothesis. The extraction of this new information is achieved in the processing of the observation data to remove the redundant (i.e. correlated) information. This is achieved via estimation methods which "whiten" the data in time and space (i.e., channels). These uncorrelated processes contain all the useful information about the processes in a compact form and are utilized to determine a sufficient statistic for the hypothesis determination; i.e., a likelihood ratio can be developed in terms of these transformed processes since

they are obtained via a causal and causally invertible transformation of the original observation data. As shown in sections IV and V, these processes enable one to efficiently perform a calculation of the likelihood ratio. Furthermore, the likelihood ratio considered here has an implementation framework amenable to adaptive processing methods.

Examples of application areas can be found in radar technology, biomedical applications, geophysical research, image processing, data compression, speech analysis and channel equalization. In the radar application, we may consider the processing of multi-sensor pre-detection data or dual-polarization data as the multichannel processes. A significant contribution of the detection algorithm considered here is the capability to utilize signal processing procedures to deal with partially correlated observation data. For the active radar case, we view the multichannel observation data as processes which arise through a simultaneous excitation of the surveillance volume with a multiplicity of waveforms. The approach here is to characterize these processes with a mathematical model (such as a multichannel AR description) and to implement a likelihood ratio whose magnitude is sensitive to the difference between the model parameters under each hypothesis. For passive detection applications, we view the processes as arising from internal physical mechanisms which give rise to the emission of radiation which may, in general, contain partial correlation when observed over specific bands. In biomedical applications, we might consider the processing of EEG waves where we may seek to detect a weak brain potential among other strong brain signals. For the purpose of validating the theoretical results developed here, at a later time, consideration will be given to the detection problem utilizing the data collected from three co-located radar systems simultaneously operating at three distinct frequencies. The resulting observation data will, however, be treated as partially correlated across the three channels. Performance evaluations will be determined in terms of receiver operating characteristics. Both analytic as well as Monte Carlo approaches will be utilized in the analysis.

A likelihood ratio for this detection problem with stationary, Gaussian signal and interference processes has been developed in section V. In section VI, implementation architectures for the likelihood ratio are briefly presented. Section VII outlines the future investigation. At a later time, we shall address the use of adaptive methods to consider the detection problem when the statistical

processes are unknown and time varying. The adaptive procedure is also approached by postulating an underlying parametric model for the observation processes. This approach enables the use of various adaptive algorithms to estimate the model coefficients and thus utilize these estimates to update the filter coefficients. The adaptive procedures are utilized to retain a robustness in non-stationary processes.

II. THE DETECTION PROBLEM

In the multichannel simple binary detection problem, the discrete received baseband waveforms are

$$\begin{aligned} H_1: \quad \underline{x}(n) &= \underline{s}(n) + \underline{c}(n) + \underline{w}(n) & n &= 1, 2, \dots, N \\ H_0: \quad \underline{x}(n) &= \underline{c}(n) + \underline{w}(n) & n &= 1, 2, \dots, N \end{aligned} \quad (2-1)$$

where $\underline{x}(n)$ is a zero mean, stationary $J \times 1$ received observation vector consisting of J channels and $\underline{s}(n)$, $\underline{c}(n)$ and $\underline{w}(n)$ are zero mean, complex Gaussian random $J \times 1$ vector processes describing the signals, non-white noise and white noise, respectively. We will assume that the white noise process is uncorrelated with $\underline{s}(n)$ and $\underline{c}(n)$ and is furthermore uncorrelated with itself in time, but not across channels, so that

$$E[\underline{w}(n) \underline{w}^H(k)] = \begin{cases} [0] & n \neq k \\ R_{ww}(0) & n = k \end{cases} \quad (2-2)$$

where $R_{ww}(0)$ is the $J \times J$ correlation matrix of $\underline{w}(n)$. The vector processes $\underline{s}(n)$ and $\underline{c}(n)$, however, contain an arbitrary correlation in time and between channels. We will consider the condition where $\underline{s}(n)$, $\underline{c}(n)$ and $\underline{w}(n)$ are jointly wide-sense stationary processes. The correlation matrix for the observation data expressed in index ordered form [1] is

$$R_{\underline{x}\underline{x}} = E[\underline{x}_{1,N} \underline{x}_{1,N}^T] \quad (2-3)$$

where

$$\underline{x}_{1,N}^T = [\underline{x}^T(1) \quad \underline{x}^T(2) \dots \underline{x}^T(N)] \quad (2-4a)$$

$$\underline{x}^T(k) = [x_1(k) \quad x_2(k) \dots x_J(k)]. \quad (2-4b)$$

Under the condition of stationarity, $R_{\underline{x}\underline{x}}$ is a Hermitian, positive semi-definite matrix. Furthermore, this matrix can be written in block form as

Under the condition of stationarity, R_x is a Hermitian, positive semi-definite matrix. Furthermore, this matrix can be written in block form as

$$R_{xx}^B = \begin{bmatrix} R_{xx}(0) & R_{xx}(-1) \dots R_{xx}(-N+1) \\ R_{xx}(1) & R_{xx}(0) \dots R_{xx}(-N+2) \\ \vdots & \\ R_{xx}(N-1) & R_{xx}(N-2) \dots R_{xx}(0) \end{bmatrix} = \begin{bmatrix} R_{xx}^H(0) & R_{xx}^H(1) \dots R_{xx}^H(N-1) \\ R_{xx}^H(-1) & R_{xx}^H(0) \dots R_{xx}^H(N-2) \\ \vdots & \\ R_{xx}^H(-N+1) & R_{xx}^H(-N+2) \dots R_{xx}^H(0) \end{bmatrix} \quad (2-5)$$

where

$$R_{xx}(l) = E [\underline{x}(k) \underline{x}^H(k-l)] \quad \begin{array}{l} k = 1, 2, \dots, N \\ l = 0, \pm 1, \dots, \pm (N-1) \end{array} \quad (2-6)$$

and the last expression in eq (2-5) results because $R_{xx}(l) = R_{xx}^H(-l)$. It is noted, however, that each block matrix of R_{xx} is not Hermitian; i.e., $R_{xx}(l) \neq R_{xx}^H(l)$ for $l \neq 0$. We also note that R_{xx} is block Toeplitz. The superscript B denotes that R_{xx} is written in block form where each block as defined in eq (2-6) is a $J \times J$ correlation matrix over the J channels.

III. AUTOREGRESSIVE PROCESS MODELS

A. DEFINITION OF THE AR PROCESS

In this analysis, the multichannel observation processes obtained under hypotheses H_i with $i = 0, 1$ are assumed to be generated by multichannel autoregressive processes. For the single channel case, analyses have been conducted [26,27] which indicate the appropriateness of such models for radar applications. For the multichannel case, similar investigations remain as a potential area for future research.

The multichannel $J \times 1$ vector process $(n|H_i)$ with $i = 0, 1$ is expressed as

$$\underline{x}(n|H_i) = - \sum_{k=1}^{M_i} A_{M_i}^H(k|H_i) \underline{x}(n-k) + \underline{u}(n|H_i) \quad i = 0, 1 \quad (3-1)$$

where $A_{M_i}^H(k|H_i)$ is the k th $J \times J$ matrix coefficient for an AR process of model order M_i . We note that it is expressed in terms of the Hermitian operation for notational convenience, but is not treated here as a Hermitian matrix. The vector $\underline{u}(n)$ is a $J \times 1$ white noise driving vector which, in general, has an arbitrary correlation across the J channels so that

$$E [\underline{u}(n) \underline{u}^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ R_{uu}(0) & l = 0. \end{cases} \quad (3-2)$$

$R_{uu}(0)$ is a $J \times J$ covariance matrix of the vector process $\underline{u}(n)$ and may have off-diagonal components. Since $\underline{u}(n)$ is uncorrelated in time, but retains an arbitrary correlation across channels, then with wide-sense joint stationarity of the channel processes assumed, we can consider

$$\underline{u}(n) = C \underline{y}(n) \quad (3-3)$$

where the $J \times J$ matrix C is a constant matrix. This matrix gives rise to the channel correlation on $\underline{u}(n)$. The vector $\underline{y}(n)$ is a Gaussian white noise vector uncorrelated in time and across channels such that

$$E[\underline{y}(n) \underline{y}^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ D_v & l = 0 \end{cases} \quad (3-4)$$

The elements of the diagonal matrix D_v are the variance terms associated with the white noise driving term on each channel. And so, from eq (3-3) we can obtain the zero-lag correlation matrix (assuming wide-sense stationarity)

$$R_{uu}(0) = E[\underline{u}(n) \underline{u}^H(n)] \quad (3-5a)$$

$$= E[C \underline{y}(n) \underline{y}^H(n) C^H] \quad (3-5b)$$

$$= C D_v C^H. \quad (3-5c)$$

We could assume unit variance on all elements of D_v without loss of generality so that $D_v = I$ and eq(3-5c) implies the Cholesky decomposition. The significance of this discussion is that the correlation matrix $R_{uu}(0)$ is a constant matrix associated with the white noise driving term $\underline{u}(n)$. The correlation between the channel elements of $\underline{u}(n)$ gives rise to the off-diagonal terms in $R_{uu}(0)$. It will be shown that this correlation causes the error output vector $\underline{e}(n)$ resulting from a MMSE estimation process to retain some residual correlation across the J channels. Since $R_{uu}(0)$ expressed in eq (3-5) is Hermitian[†], positive semi-definite, we can perform an LDL^H decomposition* such that

$$R_{uu}(0) = L_u D_u L_u^H \quad (3-6)$$

where L_u is unit diagonal lower triangular. Solving for D_u , we obtain

$$D_u = L_u^{-1} R_{uu}(0) (L_u^{-1})^H \quad (3-7a)$$

$$= E[L_u^{-1} \underline{u}(n) \underline{u}^H(n) (L_u^{-1})^H] \quad (3-7b)$$

[†] It is noted that in general the correlation matrix $R_{uu}(l)$ is not Hermitian for $l \neq 0$.

*The motivation for using the LDL^H decomposition is noted at the end of section IV.

$$= E [z(n) z^H(n)] \quad (3-7c)$$

where

$$z(n) = L_u^{-1} u(n) \quad (3-8)$$

so that $z(n)$ is a $J \times 1$ vector containing uncorrelated elements. It represents an underlying process of the multichannel AR process which can be viewed as a "spatially-causal" white noise driving term. Since L_u^{-1} is also lower triangular unit diagonal, it is invertible so that from eq (3-8)

$$u(n) = L_u z(n). \quad (3-9)$$

Eq (3-9) indicates that $u(n)$, originally defined in eq (3-3), could identically be generated by the $z(n)$ process through the transformation matrix L_u ; i.e., eq (3-1) can be written in the equivalent form

$$x(n|H_i) = - \sum_{k=1}^{M_i} A_{M_i}^H(k|H_i) x(n-k) + L_u(H_i) z(n|H_i) \quad i = 0, 1 \quad (3-10)$$

where $L_u(H_i)$ denotes the specific matrix L_u under hypothesis H_i . In section IV a two stage multichannel prediction error filter is considered which uses estimates of the $A_{M_i}^H(k|H_i)$ coefficients to obtain an approximation of $u(n)$ in the first stage and an estimate of L_u^{-1} to obtain an approximation of the temporally and spatially uncorrelated process $z(n|H_i)$.

B. THE YULE-WALKER EQUATION

The relationship between the matrix coefficients $A_M^H(k)$, the covariance matrix $[\Sigma_f]_M$ of the forward AR driving noise vector and the known correlation matrix R_{xx} noted in eq (2-3) can be expressed [1] as

$$\Delta_M^H[\tilde{R}_{xx}] = \{[\Sigma_f]_M^H [0] \dots [0]\} \quad (3-11)$$

where

$$\Delta_M^H = [I \ A_M^H(1) \ A_M^H(2) \ \dots \ A_M^H(M)]. \quad (3-12)$$

The matrix $[\tilde{R}_{xx}]$ is the reversed order correlation matrix of $[R_{xx}]$; i.e., the correlation matrix obtained with the time order of the vector $\underline{x}_{1,N}$ from eq (2-4) reversed. The corresponding equations for the stationary, backward AR process is expressed as

$$B_M^H[\tilde{R}_{xx}] = \{[0] \dots [0] \ [\Sigma_b]_M^H\} \quad (3-13)$$

where

$$B_M^H = [B_M^H(M) \dots B_M^H(1) \ I] \quad (3-14)$$

and $[\Sigma_b]_M$ is the covariance matrix of the backward AR driving noise vector. Eqs(3-11) and (3-13) are the augmented forms of the multichannel Yule-Walker equations and are presented in more detail in [38].

IV DATA TRANSFORMATION VIA LINEAR PREDICTION

The output of a multichannel linear prediction error filter of order P is expressed as

$$\underline{\epsilon}(n) = \underline{x}(n) - \hat{\underline{x}}(n|n-1) \quad (4-1a)$$

$$= \underline{x}(n) + \sum_{k=1}^P \underline{A}_P^H(k) \underline{x}(n-k) \quad (4-1b)$$

$$= \sum_{k=0}^P \underline{A}_P^H(k) \underline{x}(n-k) \quad (4-1c)$$

where $\underline{A}_P^H(k)$ $k=1,2,\dots,P$ are the matrix coefficients of the linear predictor,

$\underline{A}_P^H(0) = I$ the $J \times J$ identity matrix, the subscript P distinguishes the matrix coefficients as belonging to a filter of order P , and H denotes the Hermitian operation (i.e., the complex conjugate transpose operation).

In Appendix A, it is shown that under the condition that the matrix coefficients in eq (4-1) satisfy the multichannel normal equations,

$$E[\underline{\epsilon}(n) \underline{\epsilon}^H(n-k)] = [0] \quad k > 0 \quad (4-2)$$

and the output vector process $\underline{\epsilon}(n)$ is a MMSE process. Eq (4-2) is the orthogonality principal and indicates that the sequential outputs of the MMSE filter are orthogonal in time. The multichannel normal equations which are to be satisfied to maintain this condition are expressed as

$$\underline{A}_P^H [\tilde{\underline{R}}_{\underline{xx}}] = \{[\Sigma_f]_P^H [0] \dots [0]\} \quad (4-3a)$$

where

$$\underline{A}_P^H = [I \ \underline{A}_P^H(1) \ \underline{A}_P^H(2) \ \dots \ \underline{A}_P^H(P)] \quad (4-3b)$$

$$[\tilde{\underline{R}}_{\underline{xx}}] = E[\tilde{\underline{x}}_{1,N} \tilde{\underline{x}}_{1,N}^H] = E[\underline{x}_{N,1} \underline{x}_{N,1}^H] \quad (4-3c)$$

and

$$[\Sigma_f]_P = E[\underline{\epsilon}(n) \underline{\epsilon}^H(n)] = R_{\underline{\epsilon}\underline{\epsilon}}(0). \quad (4-3d)$$

$[R_{xx}]$ is the reversed order known correlation matrix of eq (2-3); i.e., the correlation matrix expressed with the time order of the vector in eq (2-4) reversed. We also note that eqs (4-3a) and (3-11) are identical in form. This equality implies that the MMSE estimate of the AR observation process is obtained when the prediction error filter coefficients are identically equal to the AR process coefficients. Figure 4-1 shows the synthesis and analysis procedure.

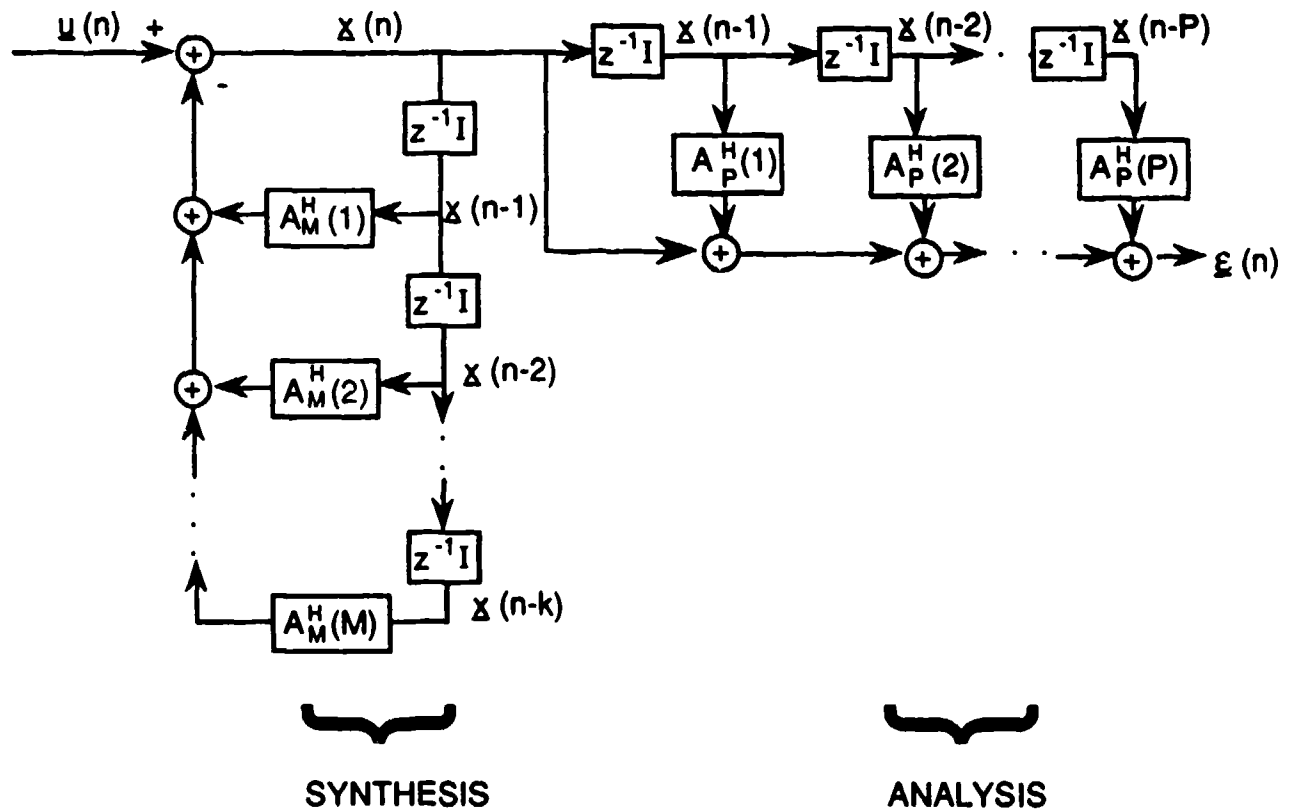


Figure 4-1

With $\underline{x}(n)$ and $\underline{u}(n)$ initially considered as non-random[†] processes, further insight can be obtained by considering the z-transform of

$$\underline{x}(n) = - \sum_{k=1}^M A_M^H(k) \underline{x}(n-k) + \underline{u}(n) \quad (4-4)$$

so that

$$\underline{X}(z) = - \sum_{k=1}^M A_M^H(k) \underline{X}(z) z^{-k} + \underline{U}(z). \quad (4-5)$$

Bringing the summation to the LHS, we can write

$$\left[\sum_{k=0}^M A_M^H(k) z^{-k} \right] \underline{X}(z) = \underline{U}(z) \quad (4-6)$$

where $A_M^H(0) = I$. We now define a filter representation for the model process as

$$H_M(z) = \left[\sum_{k=0}^M A_M^H(k) z^{-k} \right]^{-1} \quad (4-7)$$

so that

$$\underline{X}(z) = H_M(z) \underline{U}(z). \quad (4-8)$$

Eq (4-8) indicates that $\underline{X}(z)$ is the output of the filter $H_M(z)$ with input $\underline{U}(z)$. Similarly, the z-transform of eq (4-1b) where $\underline{x}(n)$ and $\underline{g}(n)$ are considered non-random is expressed as

$$\underline{E}(z) = \left[\sum_{k=0}^P A_P^H(k) z^{-k} \right] \underline{X}(z) \quad (4-9a)$$

[†] We initially consider non-random processes since the z-transform of a random process is not defined.

$$= H_F(z) \underline{X}(z) \quad (4-9b)$$

where

$$H_F(z) = \left[\sum_{k=0}^P A_P^H(k) z^{-k} \right] \quad (4-10)$$

and

$$A_P^H(0) = I. \quad (4-11)$$

When $P \geq M$, consider the case

$$A_P^H = \begin{cases} A_M^H(k) & k \leq M \\ [0] & k > M. \end{cases} \quad (4-12)$$

We then have

$$H_F(z) = H_M^{-1}(z). \quad (4-13)$$

At this point, we can now consider the input and outputs of these filters to be random. Using eqs (4-8) and (4-13) in eq (4-9b)

$$\underline{E}(z) = H_M^{-1}(z) H_M(z) \underline{U}(z) = \underline{U}(z) \quad (4-14)$$

In the time domain, eq (4-14) is equivalent to

$$\underline{E}(n) = \underline{U}(n). \quad (4-15)$$

And so, under the condition that the prediction error coefficients are identical to the coefficients of the AR model process and under the assumption that the AR process is the exact model of the observed process, the prediction error filter output $\underline{e}(n)$ is a white noise vector equivalent to the AR model white noise driving vector. However, it must be emphasized that the use of an AR process with a white noise driving function is usually an approximate representation; i.e., it is not used to describe the underlying physical mechanisms which give rise to the random processes. Rather, it is a representation which has a system transfer function given by eq (4-7). We must therefore make a distinction between the model of the processes (synthesis) and the estimation

process (analysis)[1]. In general, the output $\underline{g}(n)$ of the linear predictor is not a white noise vector output due to the approximate representation of physical processes by an AR model. It is also due to the fact that we often do not have 'a priori' knowledge regarding the values of the coefficients of this approximate model. As a result, we must estimate these coefficients from the observation data as we obtain it. With a limited amount of data, the filter coefficients are only estimates of the AR process coefficients.

For stationary processes, these coefficients could be determined through estimates of the multichannel correlation matrix lag values and the Levinson-Wiggins-Robinson algorithm [3]. Although, other methods proposed by Strand-Nuttall and Morf-Vieira have been developed [24,25,30,31] with improved performance with limited data. For non-stationary processes, adaptive schemes must be considered. We will address this topic in a subsequent report.

At this point, we note that eq (4-15) resulted from the analysis procedure[via a linear prediction error filter with coefficients given by eq(4-12)] of the process synthesized by eq (4-4). If $\underline{u}(n)$ is assumed to be uncorrelated across channels, the resulting $\underline{g}(n)$ is also uncorrelated in time and space (i.e., channels). In general, as noted in the previous section, $\underline{u}(n)$ may possess arbitrary correlation between the J channel elements. Therefore, the vector $\underline{g}(n)$ will retain a residual correlation over the channels due to the spatial correlation of $\underline{u}(n)$.

Since the matrix $R_{\underline{g}\underline{g}}(0) = E [\underline{g}(n) \underline{g}^H(n)]$ is Hermitian[†], and positive semi-definite, we can perform an LDL^H decomposition* such that

$$R_{\underline{g}\underline{g}}(0) = L_{\gamma} D_{\gamma} L_{\gamma}^H. \quad (4-16)$$

Solving for D_{γ}

$$D_{\gamma} = L_{\gamma}^{-1} R_{\underline{g}\underline{g}}(0) (L_{\gamma}^{-1})^H \quad (4-17a)$$

$$= L_{\gamma}^{-1} E [\underline{g}(n) \underline{g}^H(n)] (L_{\gamma}^{-1})^H \quad (4-17b)$$

[†] It is noted that in general the correlation matrix $R_{\underline{g}\underline{g}}(l)$ is not Hermitian for $l \neq 0$.

* Other decompositions could have been used such as Cholesky or unitary[51], however, the motivation for the LDL^H decomposition is noted at the end of this section.

$$=E [L_{\gamma}^{-1} \underline{\epsilon}(n) \underline{\epsilon}^H(n) (L_{\gamma}^{-1})^H] \quad (4-17c)$$

$$=E [\gamma(n) \gamma^H(n)] \quad (4-17d)$$

where

$$\gamma(n) = L_{\gamma}^{-1} \underline{\epsilon}(n) \quad (4-18)$$

so that the vector $\gamma(n)$ contains uncorrelated elements. Also,

$$E [\underline{\epsilon}(n) \underline{\epsilon}^H(n')] = [0] \quad n \neq n'. \quad (4-19)$$

Then, using eq (4-18) to solve for $\underline{\epsilon}(n)$ and substituting this result in (4-19), we obtain

$$E [L_{\gamma} \gamma(n) \gamma^H(n') L_{\gamma}^H] = [0] \quad n \neq n' \quad (4-20)$$

so that

$$L_{\gamma} E [\gamma(n) \gamma^H(n')] L_{\gamma}^H = [0] \quad n \neq n'. \quad (4-21)$$

Finally,

$$E [\gamma(n) \gamma^H(n')] = [0] \quad n \neq n'. \quad (4-22)$$

Eq (4-22) implies that $\gamma(n)$ retains its temporal decorrelation while eq(4-17d) denotes its spatial decorrelation. When

$$A_p^H \approx \begin{cases} A_M^H(k) & k \leq M \\ [0] & k > M \end{cases} \quad (4-23)$$

the output of the first filter stage converges toward $\underline{u}(n)$, so that eq (4-18) becomes (noting that $R_{\epsilon\epsilon}(0) \approx R_{uu}(0)$ and the uniqueness of the LDL^H decomposition)

$$\gamma(n) = L_g^{-1} \underline{\epsilon}(n) \approx L_u^{-1} \underline{u}(n) \quad (4-24a)$$

$$\approx L_u^{-1} L_u \underline{z}(n) \quad (4-24b)$$

$$\approx \underline{z}(n) \quad (4-24c)$$

where eq (3-9) was used to obtain eq (4-24b). Thus, as the filter coefficients converge to the coefficients of the AR process, $\gamma(n)$ approximates the spatially and temporally whitened process $\underline{z}(n)$. In addition, $\gamma(n)$ has been obtained through a causal and causally invertible transformation of the original observation process

$\underline{x}(n)$; ie. the input vector $\underline{x}(n)$ could be recovered through an inverse filter operation on $\underline{y}(n)$. This "information preserving" feature of $\underline{y}(n)$ justifies its use in a likelihood ratio test.

A procedure that could be used to obtain the estimate $\hat{\underline{L}}_u^{-1}$ consists of estimating the correlation lag values of $R_{\underline{\epsilon}\underline{\epsilon}}(0)$ using time samples of $\underline{\epsilon}(n)$. An LU decomposition of this matrix would provide L_γ . The inverse matrix L_γ^{-1} would then be the required estimate $\hat{\underline{L}}_u^{-1}$. An alternate approach based on a Gram-Schmidt procedure will be investigated using correlation lag estimates of $R_{\underline{\epsilon}\underline{\epsilon}}(0)$. The motivation for considering the LDL^H decomposition is based on the anticipation of utilizing a single stage recursive procedure to obtain the filter coefficients required to estimate $\underline{y}(n)$ {see section VI and [36]}.

V. MULTICHANNEL LIKELIHOOD RATIO

A. DERIVATION

In this section, we develop a multichannel likelihood ratio for the detection problem of eq.(2-1);ie.

$$H_1: \underline{x}(n) = \underline{s}(n) + \underline{c}(n) + \underline{w}(n) \quad n = 1, 2, \dots, N \quad (5-1a)$$

$$H_0: \underline{x}(n) = \underline{c}(n) + \underline{w}(n) \quad n = 1, 2, \dots, N \quad (5-1b)$$

where each of the complex vectors are Gaussian, $J \times 1$ vectors and $\underline{x}(n)$ is a baseband observation vector. The methodology derived here stems from the considerations presented in [51] for real processes. Under hypothesis H_1 , the multivariate joint Gaussian density can be written as the product of conditional densities so that

$$p_{\underline{x}}(\underline{x}_{1,n}|H_i) = p[\underline{x}(1)|H_i] \prod_{n=2}^N p[\underline{x}(n)|\underline{x}_{1,n-1}, H_i] \quad i=0,1 \quad (5-2)$$

where

$$\underline{x}_{1,n}^T = [\underline{x}^T(1) \underline{x}^T(2) \dots \underline{x}^T(n)] \quad (5-3)$$

$$\underline{x}^T(k) = [x_1(k) \ x_2(k) \dots x_J(k)] \quad (5-4)$$

and

$$\underline{x}_{1,1} = \underline{x}(1) \quad (5-5)$$

and all the conditional densities are Gaussian. The mean of the multivariate conditional density $p[\underline{x}(n)|\underline{x}_{1,n-1}, H_i]$ is $\hat{\underline{x}}(n|n-1, H_i)$; ie. the linear MMSE predictor of $\underline{x}(n)$ using past data $\underline{x}_{1,n-1}$ and assuming H_i is true. The $J \times J$ covariance matrix of this density function is the conditional covariance matrix $K_{\underline{x}}(n|n-1, H_i)$ such that

$$K_{\underline{x}}(n|n-1, H_i) = E\{[\underline{x}(n) - \hat{\underline{x}}(n|n-1, H_i)][\underline{x}(n) - \hat{\underline{x}}(n|n-1, H_i)]^H\} \quad (5-6a)$$

$$= E[\underline{e}(n|H_i)(\underline{e}^H(n|H_i))] \quad n = 1, 2, \dots, N$$

$$i = 0, 1 \quad (5-6b)$$

where

$$\underline{\epsilon}(n|H_i) = \underline{x}(n) - \underline{A}(n|n-1, H_i) \quad i = 0, 1 \quad (5-7)$$

is the zero-mean MMSE vector. Assuming wide-sense joint stationarity on the bandpass processes, the conditional density functions can be expressed in terms of the quadratic form* such that

$$\begin{aligned} p[\underline{x}(n)|\underline{x}_{1,n-1}, H_i] &= \\ &= \frac{1}{(\pi)^J |K_x(n|n-1, H_i)|} \exp\{[\underline{x}(n) - \underline{A}(n|n-1, H_i)][K_x(n|n-1, H_i)]^{-1}[\underline{x}(n) - \underline{A}(n|n-1, H_i)]^H\} \\ & \quad i = 0, 1. \end{aligned} \quad (5-8)$$

Using eq(5-7) in (5-8), we obtain

$$\begin{aligned} p[\underline{x}(n)|\underline{x}_{1,n-1}, H_i] &= \\ &= \frac{1}{(\pi)^J |K_x(n|n-1, H_i)|} \exp\{-\underline{\epsilon}^H(n|H_i)[K_x(n|n-1, H_i)]^{-1}\underline{\epsilon}(n|H_i)\} \quad i = 0, 1 \end{aligned} \quad (5-9)$$

where

$$\begin{aligned} p[\underline{x}(1)|H_i] &= p[\underline{\epsilon}(1)|H_i] = \\ &= \frac{1}{(\pi)^J |K_x(1|0, H_i)|} \exp\{-\underline{\epsilon}^H(1|H_i)[K_x(1|0, H_i)]^{-1}\underline{\epsilon}(1|H_i)\} \quad i = 0, 1 \end{aligned} \quad (5-10)$$

and

$$\underline{A}(1|0, H_i) = \underline{0} \quad i = 0, 1. \quad (5-11)$$

We can now express the likelihood ratio for the multivariate joint Gaussian density functions as

* The condition of wide-sense stationarity provides specific relationships between the auto- and cross-correlation functions of the in-phase and quadrature components. These relationships enable the multivariate Gaussian density function to be expressed directly in terms of the complex random vectors as given by eq(5-8).

$$\Lambda_{H_1, H_0} = \frac{p_X(\underline{x}_{1,N}|H_1)}{p_X(\underline{x}_{1,N}|H_0)} \quad (5-12a)$$

$$= \frac{(p[\underline{x}(1)|H_1] \prod_{n=2}^N p[\underline{x}(n)|\underline{x}_{1,n-1}, H_1])}{(p[\underline{x}(1)|H_0] \prod_{n=2}^N p[\underline{x}(n)|\underline{x}_{1,n-1}, H_0])} \quad (5-12b)$$

where the last equality results from eq(5-2). Substituting eqs(5-9) and (5-10) into (5-12b) and taking the natural logarithm, we have

$$\ln \Lambda_{H_1, H_0} = \ln \left[\frac{\prod_{n=1}^N |K_X(n|n-1, H_0)| \exp\{-\underline{\epsilon}^H(n|H_1)[K_X(n|n-1, H_1)]^{-1} \underline{\epsilon}(n|H_1)\}}{\prod_{n=1}^N |K_X(n|n-1, H_1)| \exp\{-\underline{\epsilon}^H(n|H_0)[K_X(n|n-1, H_0)]^{-1} \underline{\epsilon}(n|H_0)\}} \right] \quad (5-13a)$$

$$= \sum_{n=1}^N \left[\ln \frac{|K_X(n|n-1, H_0)|}{|K_X(n|n-1, H_1)|} + \underline{\epsilon}^H(n|H_0)[K_X(n|n-1, H_0)]^{-1} \underline{\epsilon}(n|H_0) - \underline{\epsilon}^H(n|H_1)[K_X(n|n-1, H_1)]^{-1} \underline{\epsilon}(n|H_1) \right] \quad (5-13b)$$

Eq (5-13b) can be simplified further by a diagonalization of the conditional covariance matrices. Since these matrices are Hermitian, we can perform their LDL^H decomposition such that

$$K_X(n|n-1, H_i) = L_\gamma(n|H_i) D_\gamma(n|H_i) L_\gamma^H(n|H_i) \quad i = 0, 1 \quad (5-14)$$

where the matrix $D_\gamma(n|H_i)$ $i=0,1$ is a diagonal matrix with elements $d_{jj}^\gamma(n|H_i)$. In the discussion below, we note that $K_X(n|n-1, H_i)$ is a deterministic quantity[59]. We therefore consider this matrix as well as those on the RHS of eq(5-14) as non-random. Solving for $D_\gamma(n|H_i)$, we have from eqs(5-14) and (5-6b)

$$D_{\gamma}(n|H_i) = L_{\gamma}^{-1}(n|H_i) K_x(n|n-1, H_i) [L_{\gamma}^H(n|H_i)]^{-1} \quad i = 0, 1 \quad (5-15a)$$

$$= L_{\gamma}^{-1}(n|H_i) E[\xi(n|H_i) \xi^H(n|H_i)] [L_{\gamma}^{-1}(n|H_i)]^H \quad (5-15b)$$

$$= E \left\{ L_{\gamma}^{-1}(n|H_i) \xi(n|H_i) \xi^H(n|H_i) [L_{\gamma}^{-1}(n|H_i)]^H \right\} \quad (5-15c)$$

$$= E[\chi(n|H_i) \chi^H(n|H_i)] \quad i = 0, 1 \quad (5-15d)$$

where

$$\chi(n|H_i) = L_{\gamma}^{-1}(n|H_i) \xi(n|H_i) \quad i = 0, 1. \quad (5-16)$$

Eq(5-15d) implies that $\chi(n|H_i)$ contains uncorrelated elements across the channels.

From the orthogonality condition, we have

$$E[\xi(n|H_i) \xi^H(n'|H_i)] = [0] \quad n \neq n' \quad i = 0, 1. \quad (5-17)$$

Solving for $\xi(n|H_i)$ in eq(5-16), eq(5-17) becomes

$$E \{ L_{\gamma}(n|H_i) \chi(n|H_i) \chi^H(n'|H_i) [L_{\gamma}^{-1}(n'|H_i)]^H \} = [0] \quad n \neq n' \quad i = 0, 1 \quad (5-18a)$$

so that

$$L_{\gamma}(n|H_i) E[\chi(n|H_i) \chi^H(n'|H_i)] [L_{\gamma}^{-1}(n'|H_i)]^H = [0] \quad n \neq n' \quad i = 0, 1. \quad (5-18b)$$

Since the matrices $L_{\gamma}(n|H_i)$ are, in general non-zero, it follows that

$$E[\chi(n|H_i) \chi^H(n'|H_i)] = [0] \quad n \neq n' \quad i = 0, 1. \quad (5-19)$$

In summary,

$$E[\chi(n|H_i) \chi^H(n'|H_i)] = \begin{cases} D_{\gamma}(n|H_i) & n=n' \\ [0] & n \neq n' \end{cases} \quad i = 0, 1. \quad (5-20)$$

From eq(5-14), we now have

$$K_x^{-1}(n|n-1, H_i) = [L_{\gamma}^H(n|H_i)]^{-1} D_{\gamma}^{-1}(n|H_i) L_{\gamma}^{-1}(n|H_i) \quad i=0, 1 \quad (5-21)$$

and

$$|K_X(n|n-1, H_i)| = |L_\gamma(n|H_i)| |D_\gamma(n|H_i)| |L_\gamma^H(n|H_i)| \quad i=0,1 \quad (5-22a)$$

$$= |D_\gamma(n|H_i)| \quad (5-22b)$$

$$= \prod_{j=1}^J d_{jj}^\gamma(n|H_i) = \prod_{j=1}^J \sigma_{\gamma_j}^2(n|H_i) \quad i=0,1 \quad (5-22c)$$

where

$$|L_\gamma(n|H_i)| = 1 \quad i = 0,1 \quad (5-23)$$

and

$$d_{jj}^\gamma(n|H_i) = E[|\gamma_j(n|H_i)|^2] = \sigma_{\gamma_j}^2(n|H_i) \quad i = 0,1. \quad (5-24)$$

The quantity $\sigma_{\gamma_j}^2(n|H_i)$ denotes the variance of $\gamma_j(n|H_i)$. Using eq(5-21), the quadratic terms in eq.(5-13b) become

$$\begin{aligned} \xi^H(n|H_i) [K_X(n|n-1, H_i)]^{-1} \xi(n|H_i) &= \\ &= \xi^H(n|H_i) [L_\gamma^H(n|H_i)]^{-1} D_\gamma^{-1}(n|H_i) L_\gamma^{-1}(n|H_i) \xi(n|H_i) \quad i=0,1 \end{aligned} \quad (5-25a)$$

$$= [L_\gamma^{-1}(n|H_i) \xi(n|H_i)]^H D_\gamma^{-1}(n|H_i) [L_\gamma^{-1}(n|H_i) \xi(n|H_i)] \quad i=0,1 \quad (5-25b)$$

$$= \chi^H(n|H_i) D_\gamma^{-1}(n|H_i) \chi(n|H_i) \quad i = 0,1 \quad (5-25c)$$

$$= \prod_{j=1}^J \frac{|\gamma_j(n|H_i)|^2}{\sigma_{\gamma_j}^2(n|H_i)} \quad i = 0,1 \quad (5-25d)$$

where the last two equations result from eqs(5-16) and (5-24), respectively. Using eqs(5-22c) and (5-25d) in eq(5-13b), we obtain

$$\ln \Lambda_{H_1, H_0} = \sum_{j=1}^J \sum_{n=1}^N \left\{ \ln \frac{\sigma_j^2(n|H_0)}{\sigma_j^2(n|H_1)} + \frac{|\gamma_j(n|H_0)|^2}{\sigma_j^2(n|H_0)} - \frac{|\gamma_j(n|H_1)|^2}{\sigma_j^2(n|H_1)} \right\}. \quad (5-21)$$

Eq(5-21) is the multichannel likelihood ratio for wide-sense, joint stationary, Gaussian processes. It represents a generalization of the single channel likelihood ratio reported in [6,7,23,39] for Gaussian processes. In section IV, we discussed a two stage filtering method using multichannel linear prediction filtering to obtain $\gamma(n|H_i)$. For a known correlation function under each of the two hypotheses, two sets of filter coefficients and error variances can be obtained exactly, through the multichannel Yule-Walker equations. In this case, $\gamma(n|H_i)$ will be a MMSE output of the filter for which the hypothesis H_i is true. The other filter output, however, will increase in terms of its average output magnitude. Eq(5-21) indicates that when H_0 is true, the last term will contribute a larger value in a negative sense, causing the likelihood ratio to decrease. Alternatively, when H_1 is true, the second term increases in a positive sense so that the likelihood ratio increases. For the unknown correlation case, the filter coefficients and error variances must be estimated. In this case, eq(5-21) becomes a generalized likelihood ratio and is therefore suboptimal. From one set of observation data, we must estimate the parameters (ie. the filter coefficients and error variances) for each filter assuming the appropriate hypothesis is true. In the practical implementation for this case, we must assume that the MMSE filters under each hypothesis have distinct orders so that the likelihood ratio will, in general, have a value other than zero. This assumption is justified, for example, when characterizing the observation data as autoregressive processes where the order of the process under hypothesis H_1 (signal present) is larger than that under H_0 (signal absent). For single channel processes, these considerations have been treated in [6,8].

B. THE SPECIAL CASE OF INCOHERENT INTEGRATION

In [59], several limiting forms of the single channel likelihood ratio are discussed. They include (1) the detection of a deterministic signal in additive white Gaussian noise, (2) the detection of a non-white Gaussian signal process in additive white Gaussian noise, (3) the moving target detection (MTD) algorithm for the detection of a deterministic signal in non-white Gaussian noise [60, 61] and (4) the detection of a non-white Gaussian signal process in additive non-white Gaussian noise [62]. It is also shown [59] that this likelihood ratio contains the algorithm developed [63] for the detection of a deterministic signal in additive non-white Gaussian noise of unknown correlation statistics.

One limiting form of this likelihood ratio which has not been noted is the case of uncorrelated signal processes in additive white noise. In this case, the signal is a white noise process. Since past values of such a process are uncorrelated with present and future values, the coefficients for the MMSE estimation are zeroes in the single channel case. Thus, the error signals in this case are identical to the observation processes. For $J=1$ and known variance terms, we have

$$\sigma^2(H_0) = \sigma_w^2 \quad (5-22a)$$

and

$$\sigma^2(H_1) = \sigma_s^2 + \sigma_w^2 \quad (5-22b)$$

where σ_s^2 and σ_w^2 are the known variances associated with the signal and white noise, respectively. Absorbing the constant term into the threshold, eq(5-21) becomes

$$\ln \Lambda_{H_1, H_0} = \sum_{n=1}^N \left\{ \frac{|x(n)|^2}{\sigma_w^2} - \frac{|x(n)|^2}{\sigma_s^2 + \sigma_w^2} \right\} \quad (5-23a)$$

$$= \frac{\sigma_s^2}{\sigma_w^2(\sigma_s^2 + \sigma_w^2)} \sum_{n=1}^N |x(n)|^2. \quad (5-23b)$$

Eq(5-23b) is the likelihood ratio for complex, uncorrelated signal processes in additive white noise. Its form is noted in [12,Part I,Chap. 2] for the case of real processes. Eq(5-23b) is also informative from another point of view. We recognize that it has the form of an incoherent integrator. In section VI, we discuss the implementation of eq(5-21) with a dual filter strategy, where each filter is designed for the hypothesis H_i $i=0,1$. Eq(5-23b) indicates that if the coefficients of these filters are set to zero, the likelihood implementation becomes a simple incoherent integrator. More importantly, this consideration reveals the useful role of the filtering process; ie., when the filters are used as prediction error filters, they provide the means to obtain coherent integration. This integration gain is achieved when the signal process contains pulse-to-pulse amplitude and phase correlation. The detection performance for various signal correlation levels will be described in forthcoming reports.

VI. LIKELIHOOD RATIO IMPLEMENTATION SCHEME

A. SYSTEM ARCHITECTURE

A block diagram of a system that approximates the likelihood ratio of eq (5-22) is shown in Fig 6-1. It is recognized as the multichannel extension of the implementations reported in [8,23]. The specific choice of the prediction error filter structure will depend on the assumed underlying model of the observation process [14]. A forward prediction error filter using a tapped delay line architecture is shown in Figs. 6-2 and 6-3. The lattice structure which utilizes both forward and backward coefficients is shown in Figs. 6-4 and 6-5.

We note that the architecture in Fig 6-1 utilizes two prediction error filters (PEF) implemented in parallel with each designed to be an optimal estimator under the given hypothesis. As noted in section V.A, a linear filter of a higher order may be used on the filter selected under H_1 as compared to that for H_0 if the underlying process is assumed to be an AR process of a higher order with the signal present [8,23]. Under this condition, when hypothesis H_0 is true (i.e., no signal present), both filters provide a MMSE output when the optimum filter coefficients are determined. When H_1 is true, the filter designed for H_1 adjusts to the new process. However, the filter for H_0 produces a much larger error output since this lower order filter now operates on a higher order AR process and therefore cannot adapt to the underlying process coefficients. As a result, the second summation in eq (5-22) increases. Since this term provides a positive contribution to the likelihood ratio, it is the mechanism which raises the likelihood above a predetermined threshold under H_1 .

B. FILTER COEFFICIENT DETERMINATION

The determination of the filter coefficients can be realized via several approaches. For stationary processes, the most straight forward would be a solution of the $A_p^H(k)$ coefficients via the Levinson-Wiggins-Robinson (LWR) method; however, the Strand-Nuttall and the Vieira-Morf methods may also be considered [24,25,29,30,31]. For the LWR method, estimates of the correlation lag values as well as the prediction error variances are obtained by recursive update. Since this method utilizes correlation lag value estimates, it is anticipated

to yield less accurate estimates of the coefficients than the Strand-Nuttall or Vieira-Morf approaches for the same reasons as noted by Burg [35] in the single channel case; ie, the unbiased estimate of the correlation matrix may not be positive semi-definite and therefore physically unrealizable. On the other hand, the biased estimate may yield inaccurate estimates, especially for limited data. The Strand-Nuttall method (the multichannel generalization of the Burg algorithm), however, estimates the filter coefficients directly from the data thus bypassing the requirement to estimate the correlation matrix. Actually, the reflection coefficients for the single channel or multi-channel cases are determined by this algorithm and applied to the lattice filter structures shown in Figures (6-4) and (6-5), respectively. From these values, autoregressive coefficient matrices $A^H(k)$ [as well as the backward coefficients] can be determined.

The additional matrix coefficient L_γ^{-1} discussed in section IV must also be determined in order to use the simplified form of eq(5-21). This coefficient completes the spatial whitening of the $\underline{x}(n)$ process via eq(4-18). As noted in section IV, the elements of this matrix could be obtained by performing estimates of the correlation matrix $R_{\underline{x}\underline{x}}(0) = E[\underline{x}(n)\underline{x}^H(n)]$ using the output $\underline{x}(n)$ from the first stage of processing. An LU decomposition of this matrix would yield the lower triangular unit diagonal L_γ matrix which is an estimate of L_μ . In [36], a recursive procedure is being considered which computes the filter coefficients for the structure shown in Figure 6-3. This single stage filter is equivalent to that shown in Figure 6-2; however, these coefficients are obtained via a single recursive procedure.

For the LWR method, the unbiased cross-correlation function at lag l between channel i and j at the N th data sample is calculated as

$$\hat{r}_{ij}^N(l) = \frac{1}{N-l} \sum_{k=l+1}^N x_i(k)x_j^*(k-l) \quad i,j = 1,2,\dots,J$$

$$l = 0,1,2,\dots,P \quad (6-1)$$

where P is the filter order and N is the total length of the time series with $N \gg P$. Since we are considering jointly stationary processes when using this approach, we also have

$$\hat{r}_{ij}^N(l) = [\hat{r}_{ij}^N(-l)]^* \quad (6-2)$$

Eqs (6-1) and (6-2) enable us to fill the entire correlation matrix of eq (2-3) with updated estimates. The expression in eq (6-1) can be made computationally efficient by re-expressing it in terms of the recursive form

$$\hat{r}_{ij}^N(l) = \frac{1}{N-l} x_i(N) x_j^*(N-l) + \frac{1}{N-l} \sum_{k=l+1}^{N-l} x_i(k) x_j^*(k-l) \quad (6-3a)$$

$$= \frac{x_i(N) x_j^*(N-l)}{N-l} + \left[\frac{N-l-1}{N-l} \right] \frac{1}{N-l-1} \sum_{k=l+1}^{N-l} x_i(k) x_j^*(k-l) \quad (6-3b)$$

$$= \frac{x_i(N) x_j^*(N-l)}{N-l} + \left[\frac{N-l-1}{N-l} \right] \hat{r}_{ij}^{N-1}(l) \quad \begin{array}{l} i,j = 1,2,\dots,J \\ l = 0,1,\dots,P. \end{array} \quad (6-3c)$$

Likewise, the biased* estimate of $r_{ij}(l)$ is expressed as

$$\hat{r}_{ij}^N(l) = \frac{1}{N} \sum_{k=l+1}^N x_i(k) x_j^*(k-l) \quad (6-3d)$$

$$= \frac{x_i(N) x_j^*(N-l)}{N} + \left[\frac{N-l}{N} \right] \hat{r}_{ij}^{N-1}(l) \quad \begin{array}{l} i,j = 1,2,\dots,J \\ l = 0,1,\dots,P. \end{array} \quad (6-3e)$$

Eqs (6-3c) and (6-3e) enable the current estimate of $r_{ij}(l)$ to be made in terms of past estimates $\hat{r}_{ij}^N(l)$ without a complete recalculation using all prior data.

* The biased estimate of $r_{ij}(l)$ is most often used since it ensures positive semi-definiteness of the correlation matrix[25,35].

C. ERROR VARIANCE DETERMINATION

The unbiased estimate of the prediction error variance for the j th channel under H_i for $i = 0, 1$, at the N th data sample could be calculated using the sample variance; ie.

$$\hat{\sigma}_j^2 (N|H_i) = \frac{1}{N-1} \sum_{k=1}^N \left| \gamma_j (k|H_i) - \bar{\gamma}_j (N|H_i) \right|^2 \quad i = 0, 1 \quad (6-4a)$$

where

$$\bar{\gamma}_j (N|H_i) = \frac{1}{N} \sum_{k=1}^N \gamma_j (k|H_i) \quad i = 0, 1. \quad (6-4b)$$

A recursive form for $\hat{\sigma}_j^2 (N|H_i)$ can be expressed as (see Appendix C)

$$\begin{aligned} \hat{\sigma}_j^2 (N|H_i) &= \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2 (N-1|H_i) + \frac{1}{N} |\gamma_j(N|H_i)|^2 \\ &\quad - \frac{1}{N} \left[\gamma_j(N|H_i) \bar{\gamma}_j^* (N-1|H_i) + \gamma_j^* (N|H_i) \bar{\gamma}_j (N-1|H_i) \right] \\ &\quad + \frac{1}{N} |\bar{\gamma}_j (N-1|H_i)|^2. \end{aligned} \quad (6-4c)$$

A very informative discussion regarding the effect of error variance estimates on the single channel likelihood ratio is presented in [14].

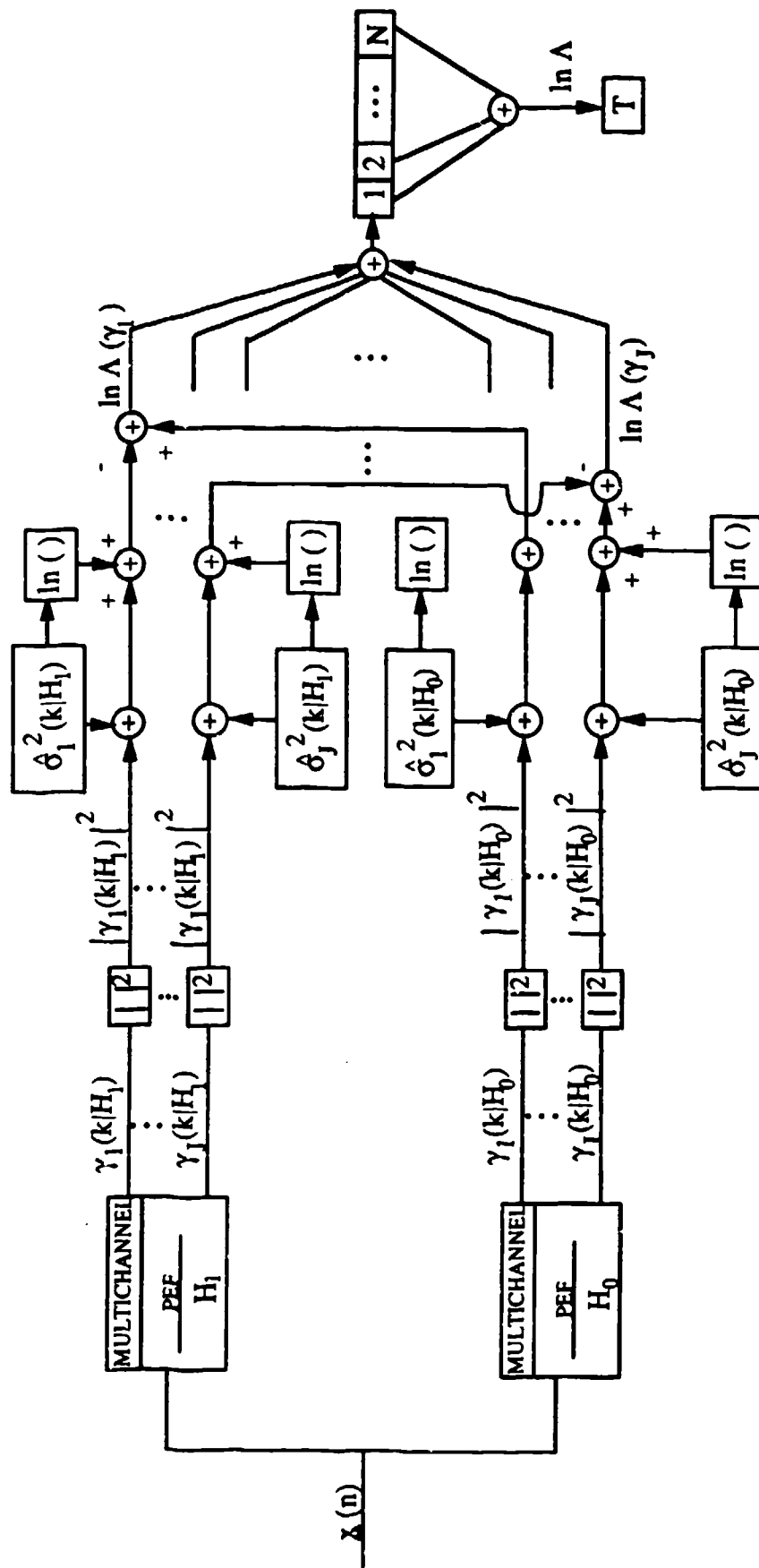


Figure 6-1 MULTI-CHANNEL LIKELIHOOD RATIO ARCHITECTURE

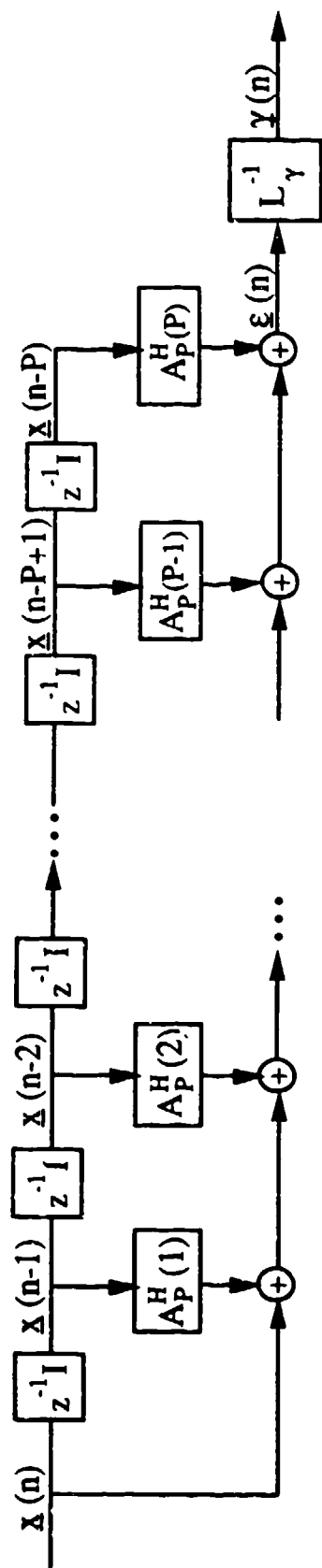


Figure 6-2 MULTI-CHANNEL PREDICTION ERROR FILTER (PEF)

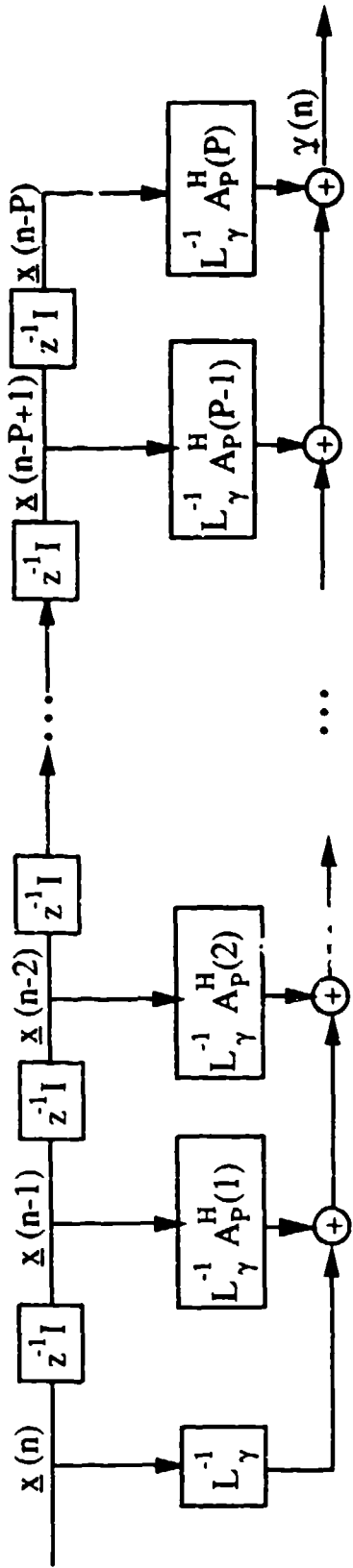


Figure 6-3 EQUIVALENT MULTI-CHANNEL PEF STRUCTURE

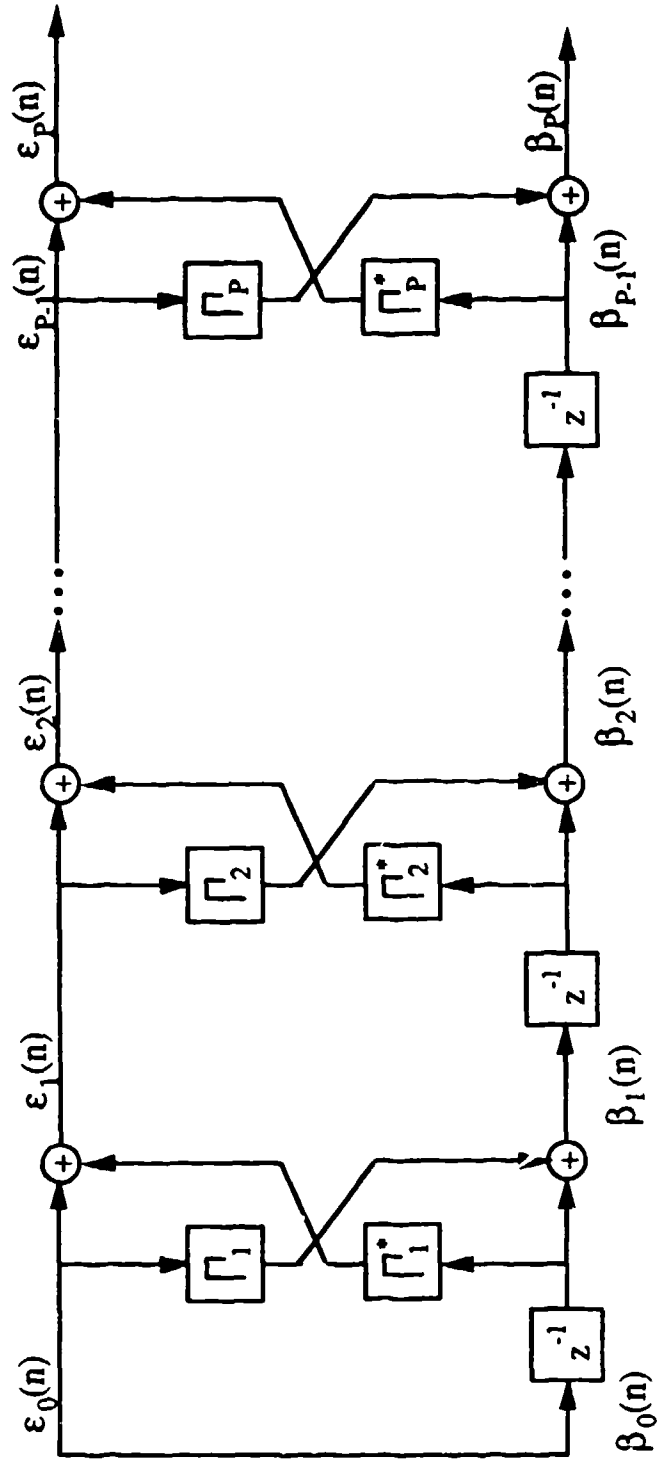


Figure 6-4 SINGLE CHANNEL LATTICE FILTER

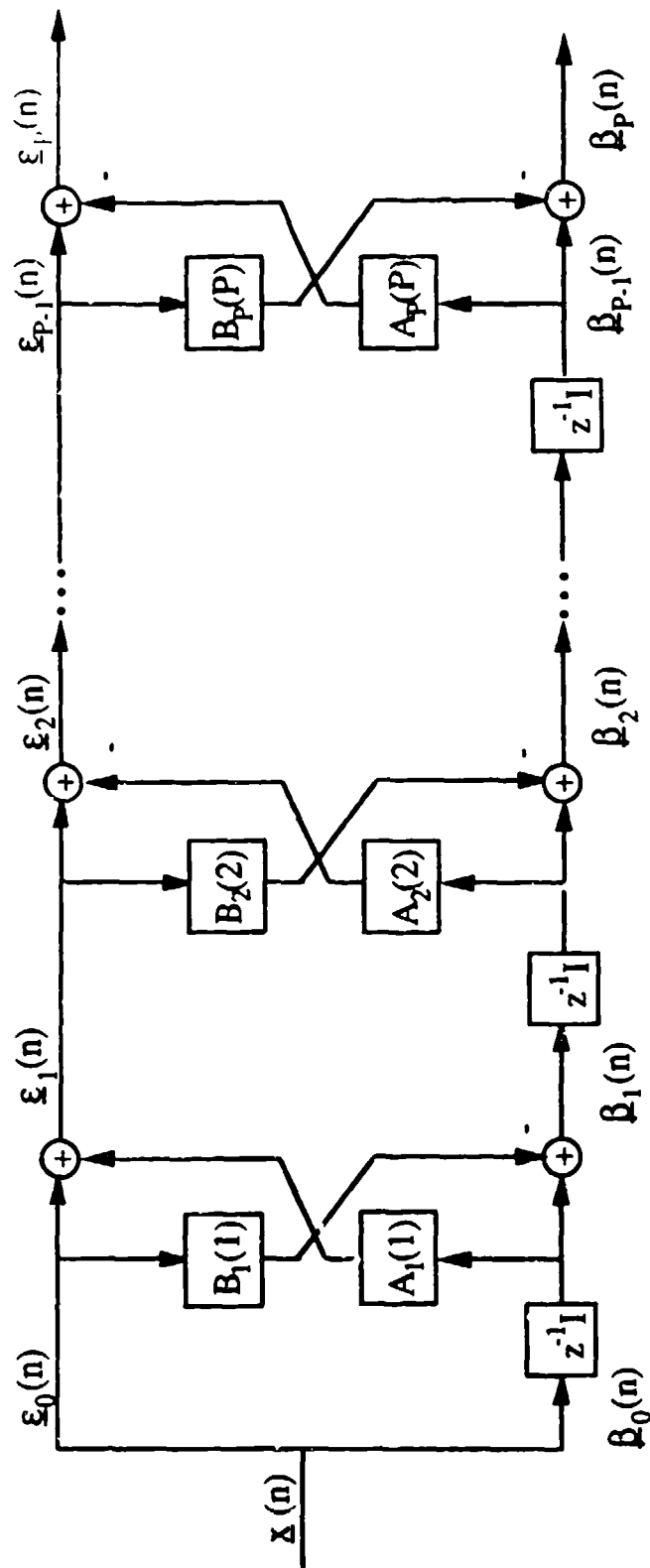


Figure 6-5 MULTI-CHANNEL LATTICE FILTER

VII PROPOSED INVESTIGATION

A. STATIONARY GAUSSIAN PROCESS SYNTHESIS

In this part of the report, a future investigation is proposed. We will assume the presence of stationary, Gaussian processes for both signal and noise. The analysis will be further divided into two areas: (1) signal in additive white noise and (2) signal in additive non-white (clutter) plus white noise. The principal objective will be the determination of receiver operating characteristic (ROC) plots in terms of probability of detection (P_d), probability of false-alarm (P_{fa}), the signal-to-noise (S/N) and signal-to-clutter (S/C) ratios associated with multichannel processes. The analysis will also consider the number of channels and the number of pulses as parameters. Performance comparisons will be made between the single channel case and the multichannel case as well as to an analytic evaluation described later in this section.

The synthesis of the random observation processes under H_i for $i=0, 1$ is now addressed. These processes will be utilized in the performance of a Monte-Carlo analyses for the determination of the ROC plots. Two methods will be considered.

METHOD 1

In the first approach we can characterize the observation processes as multichannel AR processes under each hypothesis. We then have

$$x(n|H_i) = - \sum_{k=1}^{M_i} A_{M_i}(k|H_i) x(n-k) + u(n|H_i) \quad i = 0,1 \quad (7-1)$$

where M_i , $A_{M_i}(k|H_i)$ and $u(n|H_i)$ denote the model order, the matrix coefficients and the white noise driving term under each hypothesis, respectively. In the single channel radar problem, analyses have been conducted to model radar clutter with AR processes of relatively low order [20,21]. It has also been noted [8,19,22] that the sum of two AR processes yields an autoregressive moving average (ARMA) process which in turn can be modelled as a higher-order AR process. Therefore, assuming that the signal and clutter noise are each

characterized by an AR process, we would expect that $M_1 > M_0$. For single channel processes, values of $M_0 = 1$ or 2 and $M_1 = 4$ or 5 have been reported [8] for radar applications. The extension of this work to the multichannel case remains an open area of research. Eq (7-1) could be utilized to generate the processes under each hypothesis using predetermined values for the coefficients. This approach would be useful in the diagnostics of the filtering scheme; i.e., one could validate that the filter coefficients converge to the known preassigned model coefficients as well as assess the convergence rate and final error variance. This approach, however, does not allow control over the variations of the signal-to-noise (S/N) and/or signal-to-clutter (S/C) ratios for parametric performance evaluations.

METHOD 2

In the second approach, we consider the multichannel extension of the method suggested in [23]. A complete description of this approach is presented in [74]. This is shown in Figure 7-1 where we generate the $J \times 1$ vectors $\underline{s}(n)$ and $\underline{c}(n)$ as distinct multichannel AR processes.

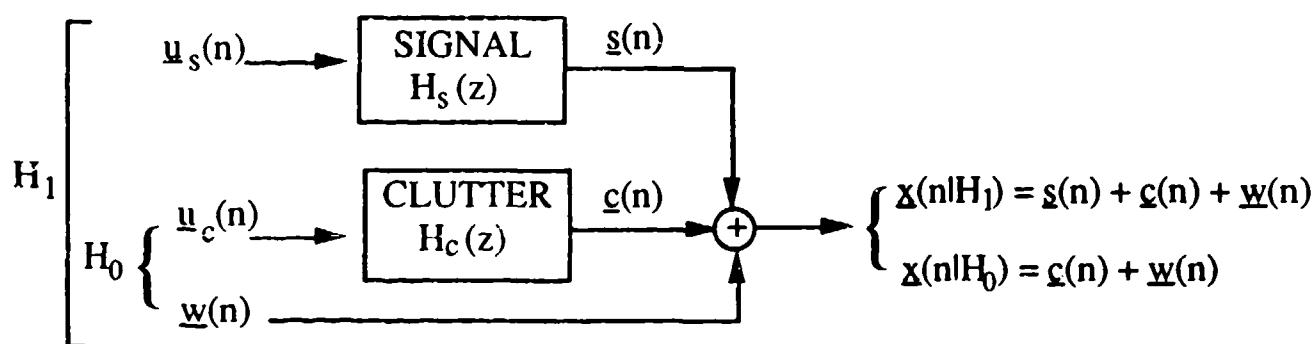


Figure 7-1

The vectors $\underline{u}_s(n)$, $\underline{u}_c(n)$ and $\underline{w}(n)$ are zero-mean, Gaussian white noise vectors uncorrelated in time but have an arbitrary correlation across channels as described in sections II and III. In this case,

$$E [\underline{u}_s(n) \underline{u}_s^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ {}^s R_{uu}(0) & l = 0 \end{cases} \quad (7-2)$$

$$E [\underline{u}_c(n) \underline{u}_c^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ {}^c R_{uu}(0) & l = 0 \end{cases} \quad (7-3)$$

$$E [\underline{w}(n) \underline{w}^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ R_{ww}(0) & l = 0 \end{cases} \quad (7-4)$$

where ${}^s R_{uu}(0)$ and ${}^c R_{uu}(0)$ are the $J \times J$ correlation matrices of the white noise driving vectors $\underline{u}_s(n)$ and $\underline{u}_c(n)$, respectively. The $J \times J$ matrix $R_{ww}(0)$ is the corresponding matrix for the additive white noise vector $\underline{w}(n)$. In general, they have off-diagonal components. The functions $H_s(z)$ and $H_c(z)$ are the filter representations for the synthesis of the signal and clutter processes, respectively. For the generation of signal AR processes we would use

$$\underline{s}(n) = - \sum_{k=1}^{M_s} A_s^H(k) \underline{s}(n-k) + \underline{u}_s(n) \quad (7-5)$$

where M_s is the order of the signal model and $A_s^H(k)$ is the matrix coefficient of the process.

We must now establish a procedure to determine the matrix coefficients $A_s^H(k)$ which when used in eq (7-5) will yield realizable values for $\underline{s}(n)$. The cross-correlation functions for the signal and non-white noise processes are to be selected using the functional forms [74]:

$$R_{ij}^g(l) = \frac{(\rho_{gij}) \sigma_{gii} \sigma_{gjj} f_g(\lambda_{gij}, |l - l_{gij}|)}{f_g(\lambda_{gij}, |l - l_{gij}|) |_{l=0}} \exp\{j[\theta_{gij}(l) - \theta_{gij}(0)]\} \quad (7-6a)$$

$g = s, c$

where $g=s,c$ refers to signal and non-white noise, respectively; ρ_{gij} is the cross correlation coefficient; σ_{gii} and σ_{gjj} are the standard deviations for the channel i and j processes; λ_{gij} is the temporal cross-correlation coefficient and is a measure of the correlation between pulses (relative to l_{gij}) across the channels i and j for $i \neq j$ [74]; for $i=j$, λ_{gii} is a temporal autocorrelation coefficient for the pulse-to-pulse correlation on channel i [23]; $\theta_{gij}(l)$ is the phase of the cross-correlation function and allows for the modeling of processes with uneven spectral shape including Doppler shifts. The functions $f(\lambda_{gij}, l - l_{gij})$ are selected to appropriately shape the autocorrelation ($i=j$) and cross-correlation ($i \neq j$) functions. The parameter l_{gij} is the lag value for which the function $f(\cdot)$ has a peak value of unity and accounts for the fact that the cross-correlation function does not peak at $l=0$ as the autocorrelation function does. Eq(7-6a) could be modified using the relation

$$|\rho_{gij}| = \rho_{gij} \exp[-j\theta_{gij}(0)]. \quad (7-6b)$$

For the autocorrelation function ($i=j$), we note that $|\rho_{gii}|=1$, $l_{gii}=0$ and $\theta_{gii}(0)=0$ [see 74] so that eq(7-6) reduces to the expression

$$R_{ii}^g(l) = \sigma_{gii}^2 f_g(\lambda_{gii}, l) \exp\{j[\theta_{gii}(l)]\} \quad g = s.c. \quad (7-7)$$

In reference [74], the selection of the $f_g(\cdot)$ functions as well as restrictions on the correlation function values are discussed to ensure conditions such as the positive semi-definiteness of the correlation matrix.

The following procedure is then used to generate the time sequenced values [23]:

- (1) the desired shapes of the autocorrelation and cross-correlation values are obtained using the functional forms above.
- (2) the order of the AR process (for synthesis) is selected based upon requirements to fit the desired spectrum.
- (3) the values of $R_{ij}^s(l)$ and $R_{ij}^c(l)$ are used to form the matrix elements of

R_{ss} and R_{cc} .

- (4) the multichannel Yule-Walker equations are solved to determine the matrix coefficients $A_g^H(k)$ for $k = 1, 2, \dots, M_g$; i.e.,

$$\Delta_g^H [R_{gg}] = \{[\Sigma_f]_g^H [0] \dots [0]\} \quad g = s, c \quad (7-8)$$

where

$$\Delta_g^H = [I \ A_g^H(1) \ A_g^H(2) \dots A_g^H(M_g)] \quad g = s, c \quad (7-9)$$

$$[\Sigma_f]_g = E[\underline{u}_g \underline{u}_g^H] \quad g = s, c \quad (7-10)$$

$$\underline{u}_g^T = [\underline{u}^T(1) \ \underline{u}^T(2) \dots \underline{u}^T(N)] \quad g = s, c \quad (7-11)$$

$$\underline{u}^T(m) = [u_1(m) \ u_2(m) \ \dots u_J(m)] \quad m = 1, 2, \dots, N. \quad (7-12)$$

- (5) the values of $A_s^H(k)$ are now substituted into eq (7-5) and used in the generation of $\underline{s}(n)$.

The vector $\underline{x}(n|H_j) \ i=0, 1$ is thus obtained as described in Fig 7-1.

B. ANALYTIC PERFORMANCE ANALYSIS

With the matrices R_{ss} and R_{cc} specified using elements determined from eq (7-6), it may be possible to proceed with an analytic determination of the probability of detection (P_d) and probability of false alarm (P_{fa}). For the single channel case, analytic expressions for these quantities are discussed in [23]. In our notation

$$P_d = 1 - \frac{1}{\pi} \int_0^\infty \text{Re} \left\{ \frac{\left(-\frac{1}{j\xi}\right) [\exp(-j\xi T) - 1]}{\det [I - 2j\xi (R_d + R_s)]^* R^*} d\xi \right\} \quad (7-13)$$

and

$$P_{fa} = 1 - \frac{1}{\pi} \int_0^\infty \text{Re} \left\{ \frac{\left(-\frac{1}{j\xi}\right) [\exp(-j\xi T) - 1]}{\det [I - 2j\xi R_d^* R]} d\xi \right\} \quad (7-14)$$

where T is the threshold and the single channel correlation matrices are expressed as

$$R_d = R_c + R_w \quad (7-15)$$

$$R = [R_d^{-1} - (R_d + R_s)^{-1}] \quad (7-16)$$

$$R_s = E[\underline{s} \underline{s}^H] \quad (7-17)$$

$$R_c = E[\underline{c} \underline{c}^H] \quad (7-18)$$

$$\underline{S}^T = [S(1) \ S(2) \dots S(N)] \quad (7-19)$$

$$\underline{C}^T = [C(1) \ C(2) \dots C(N)] \quad (7-20)$$

Equations (7-13) and (7-14) will be considered to determine if they can be generalized to the multichannel case using the matrices $R_{\underline{s}\underline{s}}$, $R_{\underline{c}\underline{c}}$ and $R_{\underline{w}\underline{w}}$. Receiver operating characteristic curves can then be generated in terms of P_d versus P_{fa} for the parameters noted in eq(7-6).

C. MONTE CARLO ANALYSIS (STATIONARY PROCESSES)

The simulation procedure discussed in VII-A will be used to synthesize multichannel observation processes under hypotheses H_0 and H_1 . A sequence of N data samples will be generated for the parameters specified in eq(7-6). The threshold level T for a given trial will be established by generating $\underline{x}(n|H_0)$ using n_s samples where n_s satisfies the condition[17] $n_s \geq 10/P_{fa}$. The simulation will then be rerun to generate N observation processes $\underline{x}(n|H_1)$ as inputs to the likelihood ratio. For stationary processes, the filter weights will be determined as discussed in section VI B. The processes $\underline{g}(n)$ or $\underline{\gamma}(n)$ will then be computed in terms of the likelihood ratio [see eq (5-21) as well as Fig. 6-1]. The number of threshold crossings shall be used to compute P_d .

Since the number of samples required to establish a given P_{fa} varies inversely with P_{fa} , low P_{fa} levels will require long computer run times. Therefore, the Monte Carlo approach may have to be limited to values such as $P_{fa} \sim 10^{-3}$ with the resulting analyses used to confirm the analytic solutions of section VII - B.

D. POSSIBLE EXTENSIONS

The generalization to non-Gaussian processes can follow the approach utilized in [4,6]. For the single channel case, Metford formed a process consisting of a variance normalized partial summation of innovations processes. In [6], he proved, (through a lengthy algebraic manipulation), that this process satisfied a central limit theorem approaching a Gaussian distribution with a rate of convergence of $N^{-1/2}$ (where N is the number of pulses). It was shown that the resulting likelihood ratio was identical in form to the single channel innovations based detection algorithm (IBDA) established for Gaussian processes, but with the additional requirement of processing over N pulses; i.e., the likelihood ratio converged to the IBDA with a rate of convergence $1/\sqrt{N}$ for large N . The exact size of N was not investigated specifically in [6], but in [4] it is indicated that these processes are expected to be Gaussian even for "a small number of samples."

In the previous sections, stationary observation processes were assumed. The condition of stationarity will be maintained for the major portion of the

performance evaluations. However, the extension to adaptive methods to update the filter coefficients can be considered. These methods are important when considering non-stationary processes with changing statistics. Consideration will be given to methods such as the least-mean-square (LMS) algorithm, recursive-least-squares (RLS) [70] and adaptive lattice filter methods.

The potential for detection performance improvement when utilizing 'a priori' information to predetermine the filter coefficients under hypothesis H_0 , could be considered. In this case, reference data collected on the non-white noise processes could be utilized to preset the filter weights. In the radar problem, a procedure similar to the CFAR approach could be utilized in which this reference data is collected from range cells adjacent to the test cell. Preliminary consideration was given to this procedure in the single channel case [14, 16] with considerable detection performance improvement noted.

Another area of investigation would involve model-fitting of vector observation processes with multichannel time series for specific applications. For the radar case, single channel clutter processes have been considered in terms of autoregressive processes [20, 21]. The extension of these analyses to multichannel processes holds potential for further research investigations which may provide performance improvement. In addition, the multichannel model-based approach appears to offer the potential to utilize data from distinct, yet correlated, processes in the detection problem; ie. if a vector time series exists which models the individual processes, the likelihood ratio presented here should be applicable. The above extensions will be investigated as time permits.

REFERENCES

- [1] Therrien, C.W., "The Analysis of Multichannel Two-Dimensional Random Signals", Naval Postgraduate School technical report NPS62-87-002, 31 Oct. 1986.
- [2] Griffiths, L. "Adaptive Structures for Multiple-Input Noise Cancelling Applications", in Proc. Int. Conf. Acoust, Speech, Signal Processing; Washington, DC, Apr. 1979, pp. 925-928.
- [3] Wiggins, R., Robinson, E., "Recursive solution to the multichannel filtering problem," JGR, Vol 70, 1965.
- [4] Metford, P.A.S., Haykin, S., Taylor, D., "An innovations approach to discrete-time detection theory," IEEE Trans. on Information Theory, Vol IT-28, no 2, March 1982.
- [5] Halmos, P.R., Measure Theory, New York: Van Nostrand, 1950.
- [6] Metford, P.A.S., "An innovations approach to discrete-time detection theory with application to radar", Ph.D. thesis, McMaster University, Sept. 1984.
- [7] Basseville, M., and Benveniste, A., "Sequential detection of abrupt changes in spectral characteristics of digital signals," IEEE Trans. Information Theory, Vol IT-29, pp. 709-724, Sept. 1983.
- [8] Haykin, S., Metford, P.A.S., "Experimental analysis of an innovations-based detection algorithm for surveillance radar," IEE Proc., Vol 132, Pt F, no 1, Feb. 1985.
- [9] Marple, S. L., Digital Spectral Analysis with Applications, Prentice Hall, N.J., 1987.

- [10] Therrien, C.W, "On the relation between triangular matrix decomposition and linear prediction", Proc. IEEE, Vol 71, no. 12, DEC. 1983, pg. 1459-1460
- [11] Papoulis, A., Probability, Random Variables and Stochastic Processes McGraw Hill, 1965.
- [12] Van Trees, H., Detection, Estimation and Modulation Theory, Part I, III Wiley and Sons, 1968.
- [13] Kadota, T., Shepp, L. "Conditions for Absolute Continuity Between a Certain Pair of Probability Measures" Z. Wahr, Verw. Feb. Vol. 16, pg 250-260 (1970).
- [14] Haykin, S., Zhang, Q.T. "A model-based detection approach", IEEE Radar Conference, 1986.
- [15] Kailath, T. "A view of three decades of linear filter theory", IEEE Trans on Inform Theory, Vol IT-20, no. 2, March 1974.
- [16] Zhang, Q.T. "Adaptive Detection for Radar" Ph.D. thesis, McMaster University, 1985.
- [17] Schleher, D.C., 1980.
- [18] Klein, A. "Polarization Processing Techniques Study", RADC-TR-79-285, Final Tech. Rpt., Nov. 1979.
- [19] Box, G. and Jenkins, G., Time Series Analysis -Forecasting and Control, Holden Day, 1976.
- [20] Gibson, C.S., Haykin, S. and Kesler, S.B., "Maximum entropy (adaptive) filtering applied to radar clutter," Proc. IEEE Intern. Conf. Acoust. Speech and Signal Processing, Washington, D.C., pg. 166-169.

- [21] Bowyer, D.E., Rajasekaran, P.K., and Gebhar, W.W., "Adaptive clutter filtering using auto-regressive spectral estimation," IEEE Trans. Aerospace and Electronic Systems, Vol. AES-15, pg. 538-546.
- [22] Haykin, S., Adaptive filter theory Prentice-Hall, 1986.
- [23] Farina, A., Russo, A., "Radar detection of correlated targets in clutter", IEEE Trans. on Aerospace and Electronics Systems, vol AES-22, no 5, Sept. 1986.
- [24] Morf, M., Vierra, A. Lee, D., Kailath, T., "Recursive Multichannel maximum entropy spectral estimation", IEEE Trans. on Geoscience Electronic, vol GE-16, no 2, April 1978.
- [25] Strand, O. N., "Multichannel complex maximum entropy (auto-regressive) spectral analysis", IEEE Trans. Antom. Control, vol, AC-22, pg 634-640, Aug 1977.
- [26] Haykin, S., Currie, B.W., Kesler, S.B., "Maximum entropy spectral analysis of radar clutter", Proc. IEEE, vol. 70, 1982, pg 953-962.
- [27] Hanson, D. S., "AR spectral modeling of simulated incoherent Doppler sonar data", MPL Technical Memorandum 367, Marine Physical Laboratory, Scripps of Oceanography, San Diego, Calif.
- [28] Inouye, Y., "Modeling of Multichannel Time Series and Extrapolation of Matrix Valued Auto-correlation Sequences", IEEE Trans on Acoustics, Speech, and Signal Processing", Vol. ASSP-31, no. 1, Feb. 1983.
- [29] Morf, M., Vieira, A., Kailath, T., "Covariance characterization by partial auto-correlation matrices", Ann. Statistics, no 6, pg 643-648, 1978.
- [30] Nuttall, A.H., FORTRAN Program for multivariate linear predictive spectral analysis, employing forward and backward averaging", Naval Underwater Systems Center TR-5419, New London, Conn., May 1976.

- [31] Nuttall, A.H., "Multivariate linear predictive spectral analysis employing weighted forward and backward averaging: A generalization of Burg's algorithm, Naval Underwater Systems Center TR-5501, New London, Conn. Oct. 1976.
- [32] Lee, D.T.L., "Canonical Ladder Form Realizations and Fast Estimation Algorithms," Ph.D. thesis, Stanford University, 1980.
- [33] Woodward, P.M., Probability and Information Theory, with Applications to Radar, McGraw Hill Book Co., 1953.
- [34] Oppenheim, A., Schafer, R., Digital Signal Processing, Prentice-Hall Inc. 1975.
- [35] Burg, J.P., "Maximum Entropy Spectral Analysis," Ph.D. dissertation, Stanford University, Stanford, California, 1975.
- [36] Michels, J. " Multichannel linear prediction and its association with triangular matrix decomposition," to be published as an RADC technical report.
- [37] Middleton, D., "Statistical Communication Theory,..." McGraw Hill, 1960.
- [38] Michels, J. "A summary of single and multichannel prediction error filters and associated algorithms," to be published as an RADC technical report.
- [39] Therrien, C.W. "A Sequential Approach to Target Discrimination" IEEE Trans. on Aerospace and Electronic Systems, vol AES-14, no.3, pp. 443-440, May 1978.
- [40] Birdsall, T.G., Gobien, J.O., "Sufficient Statistics and Reproducing Densities in Simultaneous Sequential Detection and Estimation," IEEE Trans. on Information Theory, vol IT-19, no.6, Nov. 1973.

- [41] Porat, B., Friedlander, B., "Parametric Techniques for Adaptive Detection of Gaussian Signals," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-32, no.4, Aug. 1984.
- [42] Rubinstein, R.Y., Simulation and the Monte Carlo Method, John Wiley and Sons, 1981.
- [43] Kay, S.M. "Robust detection by autoregressive spectrum analysis," IEEE Trans. on Acoustic, Speech and Signal Processing, vol. ASSP-20, pp. 256-269, April 1982.
- [44] Burdic, W.S. "Detection of narrowband signals using time domain adaptive filters," IEEE Trans. Aerospace and Electronic Systems, vol AES-14, pp. 578-591, July 1978.
- [45] Dickinson, B.W., "Properties and Applications of Gaussian Autoregressive Processes in Detection Theory," IEEE Trans on Information Theory, vol. IT-27, pp. 343-3467, May 1981.
- [46] North, D.O., "An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems," Proc. IEEE, vol. 51, pp. 1016-1027, ,1963; Report PTR-6C, RCA Laboratories, June 1943.
- [47] Price, R., "Statistical Theory Applied to Communication through Multipath Disturbances," Massachusetts Institute of Technology Research Labroatory of Electronics, Tech. Pept. 266 Spetember 3, 1953.
- [48] Price, R., "The Detection of Signals Perturbed by Scatter and Noise," IRE Trans. PGIT-4, 163-170 (Sept. 1954).
- [49] Price, R., "Notes on Ideal Receivers for Scatter Multipath," Group Rept. 34-39, Lincoln Laboratory, Massachusetts Institute of Technology, May 12, 1955.

- [50] Price, R., "Optimum Detection of Random Signals in Noise, with Application to Scatter-Multipath Communication. I" IRE Trans. PGIT-6 125-137 (Dec. 1956).
- [51] Schweppe, F., "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Trans. IT-11, No. 1, 61-70 (Jan 1965).
- [52] Kalman, R.E. and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory," ASME J. Basic Eng., 83, 95-108 (March 1961).
- [53] Kailath, T., "A General Likelihood-Ratio Formula for Random Signals in Noise," IEEE Trans. Information Theory IT-5, No. 3, 350-361 (May 1969).
- [54] Kailath, T., "The innovations approach to detection and estimation theory," Proc. IEEE, vol. 58, no. 5, May 1970.
- [55] Zadeh, L.A. and Ragazzini, J.R., "Optimum filters for the detection of signals in noise," Proc. IRE, vol. 40, pp 1223-1231, Oct 1952.
- [56] Kailath, T., "A further note on a general likelihood formula for random signals in Gaussian noise," IEEE Trans. on Information Theory, vol. IT-16, pp. 393-396, July, 1970.
- [57] Kailath, T., "An innovations approach to least-squares estimation, Part I" IEEE Trans. on Auto. Control, vol. AC-13, no. 6, Dec. 1968.
- [58] Kailath, T., "Likelihood ratios for Gaussian Processes," IEEE Trans. on Information Theory, vol. IT-16, no. 3, May 1970.
- [59] Haykin, S., Metford, P.A.S. "Some limiting forms of an Innovations-Based Discrete-Time Detection Algorithm," IEEE Trans. on Aerospace and Electronic Systems, AES-23, no.3, May 1987.

- [60] Rummler, W.D, "Clutter suppression by complex weighting of coherent pulse trains," IEEE Trans. on Aerospace and Electronic Systems, vol. AES-2, pp. 689-699, Nov, 1966.
- [61] Spafford, L.J., "Optimal radar signal processing in clutter," IEEE Trans. on Information Theory, vol IT-14, pp 734-743, Sept, 1968.
- [62] VanVleck, J., Middleton, D., "A Theoretical comparison of the visual, aural, and meter reception of pulsed signals in the presence of noise," Jour. Appl. Physics, vol. 17, pp. 940-971; Nov., 1946.
- [63] Kay, S.M., "Asymptotically optimal detection in unknown colored noise via autoregressive modeling," IEEE Trans. on Acoustics, Speech and Signal Processing, vol ASSP-31, pp. 927-940, Aug, 1983.
- [64] Van Trees, H.L., "Applications of State-Variable Techniques in detection Theory, Proc. of the IEEE, vol 58, no.5, May 1970.
- [65] Anderson, T.W., The Statistical Analysis of Time Series, John Wiley & Sons, Inc., 1971.
- [66] Priestly, M.B., Spectral Analysis and Time Series, vols 1,2, Academic Press, 1981.
- [67] Anderson, T.W., An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, 1984.
- [68] Brillinger, D.R., Time Series: Data Analysis and Theory, Holt, Rinhard and Winston, Inc., New York 1975.
- [69] Aoki, M., State Space Modeling of Time Series. Springer-Verlag, 1987.
- [70] Cioffi, J.M., Kailath, T., "Fast, recursive-least squares transversal filters for adaptive filtering," IEEE Trans on Acoustics, Speech, and Signal Processing, vol. ASSP-32, no.2, April 1984.

- [71] Wiener, N., Extrapolation, Interpolation and Smoothing of Stationary Time Series, Technology Press Books, 1957.
- [72] Robinson, E., Multichannel Time series Analysis, Goose Pond Press, 1983.
- [73] Aasnaes, H.B., Kailath, T. "An Innovations Approach to Least Squares Estimation-Part VII: Some Applications of Vector Autogressive-Moving Average Models," IEE Trans. on Automatic Control, vol AC-18, no.6, Dec., 1973.
- [74] Michels, J. "Synthesis of multichannel autoregressive random processes", to be published at an RADC technical report.

APPENDIX A

In this section, we determine the conditions under which $\underline{\epsilon}(n)$ as expressed in eq (3-1) is a temporal white noise MMSE output of a linear prediction error filter. We first follow an argument similar to the abbreviated discussion in [9]. The linear prediction error of $\underline{x}(n)$ as defined in eq (4-1) is expressed as

$$\underline{\epsilon}(n) = \underline{x}(n) - \hat{\underline{x}}(n|n-1) \quad (\text{A-1})$$

where $\hat{\underline{x}}(n|n-1)$ represents the estimated vector of $\underline{x}(n)$ using past data values with the initial condition $\underline{x}(1|0) = \underline{0}$. Using a linear predictor with P past values, we define

$$\hat{\underline{x}}(n|n-1) = - \sum_{k=1}^P \underline{A}_P^H(k) \underline{x}(n-k) \quad (\text{A-2})$$

where $\underline{A}_P^H(k)$, $k = 1, 2, \dots, P$ are $J \times J$ matrices representing the coefficients of the linear predictor. Substituting eq (A-2) in (A-1)

$$\underline{\epsilon}(n) = \underline{x}(n) + \sum_{k=1}^P \underline{A}_P^H(k) \underline{x}(n-k) \quad (\text{A-3a})$$

$$= \sum_{k=0}^P \underline{A}_P^H(k) \underline{x}(n-k) \quad (\text{A-3b})$$

where $\underline{A}_P^H(0) = \underline{I}$. Let the concatenated column vector of $P+1$ vectors (each of dimension J) be defined as

$$\underline{\underline{x}}_{n,n-P}^T = [\underline{x}^T(n) \ \underline{x}^T(n-1) \dots \underline{x}^T(n-P)]. \quad (\text{A-4})$$

Post multiplying eq (A-3) by $\underline{\underline{x}}_{n,n-P}^H$ and taking the expected value

$$\begin{aligned} E[\underline{\epsilon}(n) \underline{\underline{x}}_{n,n-P}^H] &= E[\underline{\Delta}_P^H \underline{\underline{x}}_{n,n-P} \underline{\underline{x}}_{n,n-P}^H] \\ &= \underline{\Delta}_P^H E[\underline{\underline{x}}_{n,n-P} \underline{\underline{x}}_{n,n-P}^H] \end{aligned}$$

$$= \Delta_P^H [\tilde{R}_{xx}] \quad (A-5)$$

where Δ_P^H and $[\tilde{R}_{xx}]$ are defined in eq (4-3b) and (4-3c). The LHS of eq (A-5) can also be written using the Hermitian of $\underline{x}_{n,n-P}^T$ so that

$$\begin{aligned} E [\underline{\epsilon}(n) \underline{x}_{n,n-P}^H] &= E \{ \underline{\epsilon}(n) [\underline{x}^H(n) \underline{x}^H(n-1) \dots \underline{x}^H(n-P)] \} \\ &= \{ E [\underline{\epsilon}(n) \underline{x}^H(n)] E [\underline{\epsilon}(n) \underline{x}^H(n-1)] \dots E [\underline{\epsilon}(n) \underline{x}^H(n-P)] \}. \end{aligned} \quad (A-6)$$

We now determine the coefficients of Δ_P^H subject to the condition that $\underline{\epsilon}(n)$ is a MMSE vector. Under this condition,

$$E [\underline{\epsilon}(n) \underline{x}^H(n-k)] = [0] \quad k > 0. \quad (A-7)$$

Eq (A-7) is the orthogonality condition which states that the error vector is orthogonal to past observation values. Using this condition, eq (A-6) becomes

$$E [\underline{\epsilon}(n) \underline{x}_{n,n-P}^H] = \{ E [\underline{\epsilon}(n) \underline{x}^H(n)] [0] [0] \dots [0] \}. \quad (A-8)$$

From eq (A-3a)

$$\underline{x}(n) = \underline{\epsilon}(n) - \sum_{k=1}^P A_P^H(k) \underline{x}(n-k) \quad (A-9)$$

so that

$$\underline{x}^H(n) = \underline{\epsilon}^H(n) - \sum_{k=1}^P \underline{x}^H(n-k) A_P(k). \quad (A-10)$$

Using eq (A-10) in the RHS of eq (A-8)

$$E[\underline{\epsilon}(n) \underline{x}_{n,n-P}^H] = \left\{ \left[E[\underline{\epsilon}(n) \underline{\epsilon}^H(n)] - \sum_{k=1}^P E[\underline{\epsilon}(n) \underline{x}^H(n-k)] A_P(k) \right] [0] \dots [0] \right\}$$

$$= \{[\Sigma_f]_{\mathbf{e}} \ [0] \ [0] \dots [0]\} \quad (\text{A-11})$$

where we have again used eq (A-7) to yield eq (A-11) and $[\Sigma_f]_{\mathbf{e}}$ is the forward prediction error covariance matrix. Combining eq (A-5) and (A-11), we have

$$\Delta_P^H [\hat{\mathbf{R}}_{\mathbf{xx}}] = \{[\Sigma_f]_{\mathbf{e}} \ [0] \ [0] \dots [0]\}. \quad (\text{A-12})$$

Eq (A-12) is the multichannel AR Yule Walker normal equation in augmented form. It provides a set of JP linear equations to solve for the values of the matrix coefficients which minimize the mean square error vector. Although eq (A-12) has often been presented in the literature, the reversed order form of the correlation matrix has not often been sighted [1]. We will utilize this feature in Appendix B.

We now show that the vector process $\mathbf{x}(n)$ is uncorrelated in time. At an arbitrary time $(n-l)$ where $l > 0$, eq (A-10) becomes

$$\mathbf{x}^H(n-l) = \mathbf{x}^H(n-l) - \sum_{k=1}^P \mathbf{x}^H(n-k-l) A_P(k). \quad l > 0. \quad (\text{A-13})$$

Using eq (A-13) in eq (A-7)

$$E[\mathbf{x}(n) \mathbf{x}^H(n-l)] = E[\mathbf{x}(n) \mathbf{x}^H(n-l)] - \sum_{k=1}^P E[\mathbf{x}(n) \mathbf{x}^H(n-k-l)] A_P(k) = [0]. \quad (\text{A-14})$$

From eq (A-7), we have

$$E[\mathbf{x}(n) \mathbf{x}^H(n-k-l)] = [0] \quad k > 0. \quad (\text{A-15})$$

And so, eq (A-14) yields

$$E[\mathbf{x}(n) \mathbf{x}^H(n-l)] = [0] \quad l > 0. \quad (\text{A-16})$$

Thus, the sequence of outputs from the MMSE prediction error filter are orthogonal. Since $\mathbf{x}(n)$ is a zero mean Gaussian process, its sequence of values $\{\mathbf{x}(n)\}$ are mutually independent so that $\{\mathbf{x}(n)\}$ is a white, Gaussian noise sequence.

APPENDIX B

In this Appendix, we show that the matrix coefficients of multichannel linear prediction for a multichannel random process and the prediction error covariance matrices are related to the covariance matrix through a block triangular decomposition. The procedure is a straight forward generalization of [10]. In this discussion, we are considering the first stage of processing as noted in sections IV through VI. This stage results in the $\underline{x}(n)$ vector output. Therefore, we only consider the the $A(k)$ matrix coefficients here. A treatment which includes the second stage of processing using the matrix L_γ to obtain $\underline{y}(n)$ is developed in [36].

Recognizing that $R_{\underline{x}\underline{x}}^B$ defined in eq (2-5) is Hermitian with non-singular upper left principal minors, we can obtain

$$R_{\underline{x}\underline{x}}^B = L_{\underline{\epsilon}} D_{\underline{\epsilon}} L_{\underline{\epsilon}}^H \quad (B-1)$$

where $L_{\underline{\epsilon}}$ is lower block triangular with the identity matrix I forming the block diagonal matrices and $D_{\underline{\epsilon}}$ is a real block diagonal matrix. Solving eq (B-1) for $D_{\underline{\epsilon}}$

$$D_{\underline{\epsilon}} = L_{\underline{\epsilon}}^{-1} R_{\underline{x}\underline{x}}^B (L_{\underline{\epsilon}}^{-1})^H. \quad (B-2)$$

If we consider

$$\underline{\epsilon}_{1,N} = L_{\underline{\epsilon}}^{-1} \underline{x}_{1,N} \quad (B-3)$$

we can easily show that

$$D_{\underline{\epsilon}} = E [\underline{\epsilon}_{1,N} \underline{\epsilon}_{1,N}^H] \quad (B-4)$$

where

$$\underline{\epsilon}_{1,N}^T = [\underline{\epsilon}(1) \underline{\epsilon}(2) \dots \underline{\epsilon}(N)]^T \quad (\text{B-5})$$

and $\underline{\epsilon}(k)$ is a $J \times 1$ channel vector. Since $L_{\underline{\epsilon}}$ is lower block triangular with unit diagonal elements, $L_{\underline{\epsilon}}^{-1}$ has the same form so that eq (B-3) is a causal and causally invertible transformation of the data.

We now consider the normal equations for a multichannel predictor of order p such that

$$\underline{\Delta}_P^H [\tilde{R}_{XX}] = [\Sigma_f]_P^H I^T \quad (\text{B-6})$$

where $[\tilde{R}_{XX}]_P$ is the reversed order multichannel covariance matrix,

$$\underline{\Delta}_P^H = [I \ A_P^H(1) \ A_P^H(2) \dots A_P^H(P)] \quad (\text{B-7})$$

and

$$I^T = \{I \ [0] \ [0] \dots [0]\} \quad (\text{B-8})$$

where I is a $J \times J$ identity matrix. The vector of matrices $\underline{\Delta}_P^H$ is the vector of multichannel p th order linear prediction coefficients and $[\Sigma_f]_P$ is the Hermitian, p th order, multichannel prediction error covariance matrix. Post multiplying eq (B-6) by $\underline{\Delta}_P$, and recognizing that $[\Sigma_f]_P$ is Hermitian, we obtain

$$[I \ A_P^H(1) \ A_P^H(2) \dots A_P^H(P)] [\tilde{R}_{XX}] \begin{bmatrix} I \\ A_P(1) \\ A_P(2) \\ \vdots \\ A_P(P) \end{bmatrix} = [\Sigma_f]_P. \quad (\text{B-9})$$

Using the relation

$$G_B^{-1} G_B = I \quad (\text{B-10})$$

where G_B is the block reflection matrix (i.e., the square $J(p+1) \times J(p+1)$ matrix with $J \times J$ identity matrices along the block cross diagonal), one can easily show that

$$[A_P^H(P) \dots A_P^H(2) A_P^H(1) I][R_{\Delta\Delta}] \begin{bmatrix} A_P(P) \\ \vdots \\ A_P(2) \\ A_P(1) \\ I \end{bmatrix} = [\Sigma_f]_P. \quad (B-11)$$

At this point, it is noted that eq (B-11) results because of the reversed order of the correlation matrix [1].

We now write eq (B-11) for $p = 0, 1, \dots, P = N-1$ in the combined form as

$$\begin{bmatrix} I & & & & \\ A_1^H(1) I & & & & \\ A_2^H(2) A_2^H(1) I & & & & \\ A_3^H(3) A_3^H(2) A_3^H(1) I & \ddots & & & \\ \vdots & & & & \\ A_P^H(P) A_P^H(P-1) A_P^H(P-2) \dots A_P^H(1) I & & & & \end{bmatrix} \begin{bmatrix} R(0) R(-1) R(-2) \dots R(-P) \\ R(1) R(0) R(1) \dots R(-P+1) \\ R(2) R(1) R(0) \dots R(-P+2) \\ \vdots \quad \ddots \\ R(P) R(P-1) R(P-2) \dots R(0) \end{bmatrix} \\ \cdot \begin{bmatrix} I A_1(1) A_2(2) A_3(3) \dots A_P(P) \\ I A_2(1) A_3(2) \dots A_P(P-1) \\ I A_3(1) \dots A_P(P-2) \\ \vdots \\ [0] \quad I \quad \ddots \quad I \end{bmatrix} = \begin{bmatrix} [\Sigma_f]_0 & & & & \\ & [\Sigma_f]_1 & & & \\ & & [\Sigma_f]_2 & & \\ & & & \ddots & \\ [0] & & & & [\Sigma_f]_P \end{bmatrix} \quad (B-12)$$

Eq (B-12) is of the same form as eq (B-2). Since the causal decomposition in eq (B-2) is unique[†], then L_E^{-1} can be identified with the lower triangular matrix in eq (B-12). Thus the block rows of L_E^{-1} are the multichannel coefficients for linear predictive filters of orders zero through $P = N-1$ and the block diagonal matrix elements of D_E are the prediction error variances associated with the multichannel filter orders.

[†] The uniqueness of this decomposition is based upon the specified block form of R_{XX}^B defined in eq (2-5). However, other block forms of R_{XX} could have been made which still retain the Hermitian property.

APPENDIX C

In this Appendix, a recursive expression for the sample variance estimate of the complex quantity $\gamma_j(n|H_i)$ is derived. The sample mean of the j th channel sequence $\gamma_j(n|H_i)$ is given at the N th time sample as

$$\bar{\gamma}_j(N|H_i) = \frac{1}{N} \sum_{l=1}^N \gamma_j(l|H_i) \quad i=0,1. \quad (C-1)$$

By definition, the sample variance of the complex quantity $\gamma_j(n|H_i)$ is expressed as

$$\delta_j^2(N|H_i) = \frac{1}{N-1} \sum_{k=1}^N |\gamma_j(k|H_i) - \bar{\gamma}_j(N|H_i)|^2 \quad i=0,1 \quad (C-2a)$$

$$= \frac{1}{N-1} \sum_{k=1}^N [\gamma_j(k|H_i) - \bar{\gamma}_j(N|H_i)] [\gamma_j^*(k|H_i) - \bar{\gamma}_j^*(N|H_i)] \quad i=0,1 \quad (C-2b)$$

$$= \frac{1}{N-1} \sum_{k=1}^N \{ |\gamma_j(k|H_i)|^2 - \gamma_j(k|H_i) \bar{\gamma}_j^*(N|H_i) - \bar{\gamma}_j(N|H_i) \gamma_j^*(k|H_i) + |\bar{\gamma}_j(N|H_i)|^2 \} \quad i=0,1. \quad (C-2c)$$

Using eq (C-1) in (C-2c)

$$\begin{aligned}
\hat{\sigma}_j^2(N|H_i) &= \frac{1}{N-1} \sum_{k=1}^N |\gamma_j(k|H_i)|^2 \frac{1}{(N-1)} \sum_{k=1}^N \gamma_j(k|H_i) \frac{1}{N} \sum_{l=1}^N \gamma_j^*(l|H_i) \\
&\quad - \frac{1}{N} \sum_{l=1}^N \gamma_j(l|H_i) \frac{1}{(N-1)} \sum_{k=1}^N \gamma_j^*(k|H_i) \\
&\quad + \frac{1}{N-1} \sum_{k=1}^N \frac{1}{N^2} \sum_{l=1}^N \gamma_j(l|H_i) \sum_{m=1}^N \gamma_j^*(m|H_i) \quad i=0,1 \quad (C-3)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{(N-1)} \sum_{k=1}^N |\gamma_j(k|H_i)|^2 - \frac{1}{N(N-1)} \sum_{k=1}^N \sum_{l=1}^N \gamma_j(k|H_i) \gamma_j^*(l|H_i) \\
&\quad - \frac{1}{N(N-1)} \sum_{l=1}^N \sum_{k=1}^N \gamma_j(l|H_i) \gamma_j^*(k|H_i) + \frac{1}{N^2(N-1)} \sum_{k=1}^N \sum_{l=1}^N \sum_{m=1}^N \gamma_j(l|H_i) \gamma_j^*(m|H_i) \\
&\quad i=0,1. \quad (C-4)
\end{aligned}$$

Interchanging k and l in the second summation as well as k and m in the fourth does not change the expression. And so,

$$\begin{aligned}
\hat{\sigma}_j^2(N|H_i) &= \frac{1}{(N-1)} \sum_{k=1}^N |\gamma_j(k|H_i)|^2 - \frac{2}{N(N-1)} \sum_{l=1}^N \sum_{k=1}^N \gamma_j(l|H_i) \gamma_j^*(k|H_i) \\
&\quad + \sum_{m=1}^N \frac{1}{N^2(N-1)} \sum_{l=1}^N \sum_{k=1}^N \gamma_j(l|H_i) \gamma_j^*(k|H_i) \quad i=0,1. \quad (C-5)
\end{aligned}$$

But the last summation in eq (C-5) is just

$$\frac{1}{N(N-1)} \sum_{l=1}^N \sum_{k=1}^N \gamma_j(l|H_i) \gamma_j^*(k|H_i)$$

so that

$$\hat{\sigma}_j^2(N|H_i) = \frac{1}{(N-1)} \sum_{k=1}^N |\gamma_j(k|H_i)|^2 - \frac{1}{N(N-1)} \sum_{k=1}^N \sum_{l=1}^N \gamma_j(l|H_i) \gamma_j^*(k|H_i) \quad i=0,1. \quad (C-6)$$

To obtain a recursive form for $\sigma_j^2(N|H_i)$ we separate out the terms involving $\gamma_j(N|H_i)$ so that

$$\begin{aligned} \sigma_j^2(N|H_i) &= \frac{1}{(N-1)} |\gamma_j(N|H_i)|^2 + \frac{1}{N-1} \sum_{k=1}^{N-1} |\gamma_j(k|H_i)|^2 - \frac{1}{N(N-1)} \gamma_j(N|H_i) \gamma_j^*(N|H_i) \\ &\quad - \frac{1}{N(N-1)} \gamma_j(N|H_i) \sum_{k=1}^{N-1} \gamma_j^*(k|H_i) \\ &\quad - \frac{1}{N(N-1)} \gamma_j^*(N|H_i) \sum_{l=1}^{N-1} \gamma_j(l|H_i) - \frac{1}{N(N-1)} \sum_{l=1}^{N-1} \sum_{k=1}^{N-1} \gamma_j(l|H_i) \gamma_j^*(k|H_i) \end{aligned} \quad i=0,1 \quad (C-7a)$$

$$\begin{aligned}
&= \frac{1}{(N-1)} |\gamma_j(N|H_i)|^2 + \frac{1}{N-1} \sum_{k=1}^{N-1} |\gamma_j(k|H_i)|^2 - \frac{1}{N(N-1)} |\gamma_j(N|H_i)|^2 \\
&- \frac{1}{N(N-1)} \gamma_j(N|H_i) \sum_{k=1}^{N-1} \gamma_j^*(k|H_i) - \frac{1}{N(N-1)} \gamma_j^*(N|H_i) \sum_{l=1}^{N-1} \gamma_j(l|H_i) \\
&- \frac{1}{N(N-1)} \sum_{l=1}^{N-1} \sum_{k=1}^{N-1} \gamma_j(l|H_i) \gamma_j^*(k|H_i) \quad i=0,1
\end{aligned} \tag{C-7b}$$

so that

$$\begin{aligned}
\hat{\sigma}_j^2(N|H_i) &= \frac{1}{N} |\gamma_j(N|H_i)|^2 + \frac{1}{N-1} \sum_{k=1}^{N-1} |\gamma_j(k|H_i)|^2 - \frac{1}{N} \gamma_j(N|H_i) \bar{\gamma}_j^*(N-1|H_i) \\
&- \frac{1}{N} \gamma_j^*(N|H_i) \bar{\gamma}_j(N-1|H_i) - \frac{1}{N(N-1)} \sum_{l=1}^{N-1} \sum_{k=1}^{N-1} \gamma_j(l|H_i) \gamma_j^*(k|H_i) \quad i=0,1.
\end{aligned} \tag{C-7c}$$

From eq (C-6)

$$\begin{aligned}
\hat{\sigma}_j^2(N-1|H_i) &= \frac{1}{(N-2)} \sum_{k=1}^{N-1} |\gamma_j(k|H_i)|^2 - \frac{1}{(N-1)(N-2)} \sum_{k=1}^{N-1} \sum_{l=1}^{N-1} \gamma_j(l|H_i) \gamma_j^*(k|H_i) \\
&\quad i=0,1.
\end{aligned} \tag{C-8}$$

Multiplying both sides by $N-2/N-1$,

$$\frac{N-2}{N-1} \hat{\sigma}_j^2(N-1|H_i) = \frac{1}{(N-1)} \sum_{k=1}^{N-1} |\gamma_j(k|H_i)|^2 - \frac{1}{(N-1)} \sum_{l=1}^{N-1} \gamma_j(l|H_i) \frac{1}{(N-1)} \sum_{k=1}^{N-1} \gamma_j^*(k|H_i) \quad i=0,1$$

(C-9)

so that

$$\frac{1}{N-1} \sum_{k=1}^{N-1} |\gamma_j(k|H_i)|^2 = \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2(N-1|H_i) + \bar{\gamma}_j(N-1|H_i) \bar{\gamma}_j^*(N-1|H_i) \quad i=0,1$$

(C-10a)

$$= \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2(N-1|H_i) + |\bar{\gamma}_j(N-1|H_i)|^2 \quad i=0,1.$$

(C-10b)

Using eq (C-10b) in eq (C-7c)

$$\begin{aligned} \hat{\sigma}_j^2(N|H_i) &= \frac{1}{N} |\gamma_j(N|H_i)|^2 + \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2(N-1|H_i) + |\bar{\gamma}_j(N-1|H_i)|^2 \\ &\quad - \frac{\gamma_j(N|H_i) \bar{\gamma}_j^*(N-1|H_i)}{N} - \frac{\gamma_j^*(N|H_i) \bar{\gamma}_j(N-1|H_i)}{N} \\ &\quad - \frac{(N-1)}{(N)} \left[\frac{1}{(N-1)} \sum_{l=1}^{N-1} \gamma_j(l|H_i) \right] \left[\frac{1}{(N-1)} \sum_{k=1}^{N-1} \gamma_j^*(k|H_i) \right] \quad i=0,1 \end{aligned}$$

(C-11)

$$= \frac{1}{N} |\gamma_j(N|H_i)|^2 + \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2(N-1|H_i) + |\bar{\gamma}_j(N-1|H_i)|^2 \\ - \frac{1}{N} \left[\gamma_j(N|H_i) \bar{\gamma}_j^*(N-1|H_i) + \gamma_j^*(N|H_i) \bar{\gamma}_j(N-1|H_i) \right] - \frac{(N-1)}{N} |\bar{\gamma}_j(N-1|H_i)|^2$$

so that

$$\hat{\sigma}_j^2(N|H_i) = \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2(N-1|H_i) + \frac{1}{N} |\gamma_j(N|H_i)|^2 \\ - \frac{1}{N} [\gamma_j(N|H_i) \bar{\gamma}_j^*(N-1|H_i) + \gamma_j^*(N|H_i) \bar{\gamma}_j(N-1|H_i)] \\ + \frac{1}{N} |\bar{\gamma}_j(N-1|H_i)|^2 \quad (C-12a)$$

$$= \frac{(N-2)}{(N-1)} \hat{\sigma}_j^2(N-1|H_i) + \frac{1}{N} |\gamma_j(N|H_i) - \bar{\gamma}_j(N-1|H_i)|^2 \quad (C-12b)$$

with initial conditions

$$\hat{\sigma}_j^2(1|H_i) = 0 \quad i = 0, 1. \quad (C-13)$$