

RL-TR-93-236
Final Technical Report
December 1993



AD-A275 759



THE BOOTSTRAPPED ALGORITHM: A FAST ALGORITHM FOR BLIND SIGNAL SEPARATION

New Jersey Institute of Technology

Yehezkel Bar-Ness, Abdulkadir Dinc, Hagit Messer-Yaron,
Zoran Siveski, Raafat Kamel, and Ruth Onn

DTIC
ELECTE
FEB 16 1994
S B D

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

RECEIVED

Rome Laboratory
Air Force Materiel Command
Griffiss Air Force Base, New York

ISJAL 94-05098



94 2 15 084

This report has been reviewed by the Rome Laboratory 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.

RL-TR-93-236 has been reviewed and is approved for publication.

APPROVED:



RICHARD N. SMITH
Project Engineer

FOR THE COMMANDER:



JOHN A. GRANIERO
Chief Scientist
Command, Control and Communications Directorate

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL (C3BA) Griffiss AFB NY 13441. 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 require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

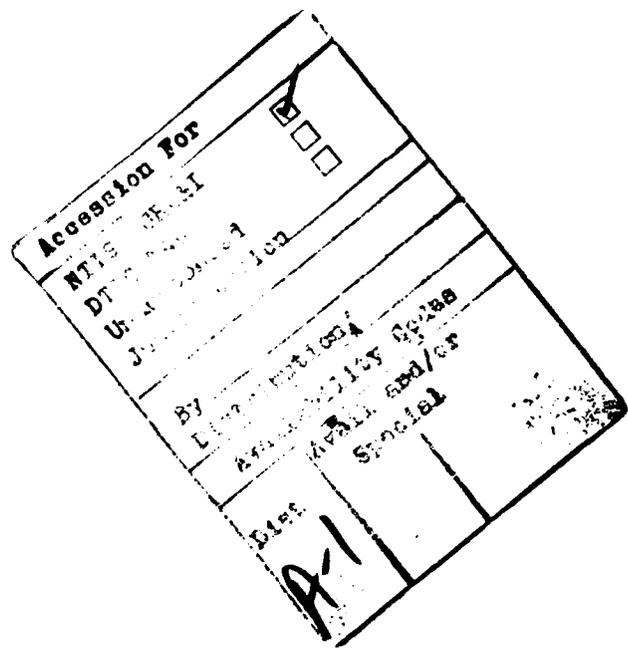
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE December 1993		3. REPORT TYPE AND DATES COVERED Final Apr 92 - Jun 93	
4. TITLE AND SUBTITLE THE BOOTSTRAPPED ALGORITHM: A FAST ALGORITHM FOR BLIND SIGNAL SEPARATION				5. FUNDING NUMBERS C - F30602-88-D-0025, Task C-2-2404 PE - 62702F PR - 4519 TA - 63 WU - P3	
6. AUTHOR(S) Yehekel Bar-Ness, Abdulkadir Dinc, Hagit Messer-Yaron, Zoran Siveski, Raafat Kamel, and Ruth Onn				8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) New Jersey Institute of Technology Dept of Electrical & Computer Engineering University Heights Newark NJ 07102				10. SPONSORING/MONITORING AGENCY REPORT NUMBER RL-TR-93-236	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Rome Laboratory (C3BA) 525 Brooks Road Griffiss AFB NY 13441-4505					
11. SUPPLEMENTARY NOTES Rome Laboratory Project Engineer: Richard N. Smith/C3BA/(315) 330-3091					
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The bootstrapped algorithm was introduced in the previous report. RL-TR-93-24, dated April 1993. Their different structures were proposed and shown to be suitable for blind signal separation and interference cancellation. It was demonstrated that the speed of convergence of these algorithms is much faster than other algorithms. Unlike other algorithms, the speed of convergence does not depend on the signals' power ratios and, hence, does not depend on the eigenvalue spread of the input correlation matrix. Further studies of these algorithms are reported in this section, which includes questions related to bandwidth, complexity, and dynamic studies for real-time control. Also reported are studies of spatial separation and direction-of-arrival estimation of wideband signals as well as application of the algorithm for broadband interference cancellation. Some preliminary results on applications to the blind decision feedback equalizer and separation of CDMA signals are also included.					
14. SUBJECT TERMS Algorithm, Signal Processing, Interference Cancellation				15. NUMBER OF PAGES 34	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED		18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED		19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	
				20. LIMITATION OF ABSTRACT UL	

TABLE OF CONTENTS

Chapter	Page
1. Introduction	1
2. Technical Summary	4
2.1 Bootstrapped Algorithm - Separation of Two Superimposed Signals	4
2.1.1 Bandwidth Complexity Trade-Off	4
2.1.2 Stability Consideration	6
2.1.3 Dynamic Studies	9
2.1.3.1 Dynamic Study of B/B Power-Power Scheme Using Orthogonal Perturbation Sequences	9
2.1.3.2 Dynamic Study of B/F Power-Correlator Scheme Using Weight Dithering with PN Sequences	10
2.2 Bootstrapped Algorithms – Wideband Signal Separators	11
2.3 Bootstrapped Algorithms – Multi-Input, Multi-Output Separators	14
2.4 Bootstrapped Algorithm – Other Applications	14
3. Conclusions and Recommendations	15
4. References	18
Appendix A: Bootstrapped Adaptive Separation of Two Superimposed Signals – Bandwidth Complexity Trade-Off	29
Appendix B: Bootstrapped Adaptive Separation of Two Superimposed Signals – Stability Considerations	71
Appendix C: Bootstrapped Adaptive Separation of Two Superimposed Signals – Dynamic Study of Power-Power Scheme Using Orthogonal Perturbation Sequences	93

Chapter	Page
Appendix D: Bootstrapped Adaptive Separation of Two Superimposed Signals - Dynamic Study of Power-Correlator Scheme Using Weight Dithering with PN Sequences	115
Appendix E: Simultaneous Spatial Separation and Direction-Of-Arrival Estimation of Wideband Source Using Bootstrapped Algorithms	129
Appendix F: Broadband Interference Cancellation Using A Bootstrapped Approach ...	161
Appendix G: A Forward/Backward Bootstrapped Structure for Blind Separation of Signals in a Multi-Channel Dispersive Environment	173
Appendix H: Blind Decision Feedback Equalization Using The Decorrelation Criterion	187
Appendix I: Adaptive Two-Stage Detection Scheme in Synchronous Two-User CDMA Systems	213



I. INTRODUCTION

Reported below are the results of a study carried out at the Center for Communications and Signal Processing Research, at the New Jersey Institute of Technology, between April 1, 1992 and June 30, 1993. This research is a continuation of previous research performed during the same period in 91/92 [1]. The aim of the research is to further analyze and evaluate the performance of a fast algorithm known as the "bootstrapped algorithm." Historically, the idea of the algorithm as a way for cancelling interferences was first proposed by the principle investigator in 1981 [2], and later it was used for cancelling cross polarization in satellite communication [3] and in the microwave terrestrial radio line [4-6].

The bootstrapped interference canceler is principally composed of two separate cancelers, each using the output of the other canceler as its reference (desired signal) input. In fact, such a structure performs as a "Signal Separator" rather than an interference canceler. Since for its operation there is no need for a reference signal, it is sometimes justifiably called a "Blind Separator."

Three different structures were proposed in [1], namely, the Backward/Backward (BB), the Forward/Forward (FF) and the Forward/Backward (FB). They are depicted in Figures 1, 2, and 3. The different weights of these two-input/two-output separators can be controlled by minimizing the output power or minimizing the absolute value of the cross correlation between the two outputs.

During the first phase of this report [1], the steady state behavior of these separators was examined. It was shown that under acceptable conditions they converge to a state which represents the desirable signal separation. The effect of additive noise on the performance of the separator was also studied in [1].

It was also demonstrated that extensions of all three structures to multi-input/multi-output are possible. Such extensions are particularly important in multi-channel communication, such as Code Division Multiple Access (CDMA) applications or neural networks.

Some preliminary work on the applicability of the algorithms was also performed and some results were presented. Special emphasis was put on the usage of the three structures in handling dually-polarized signals where we examined the error probabilities of the bootstrapped cross-pol cancelers for M-ary QAM signals. Their performance was compared to other kinds of cancelers known in the literature.

It was concluded that the bootstrapped algorithms have many useful properties which make them excellent candidates for use as signal separators or interference cancelers when other algorithms have difficulties. In some cases, it clearly outperforms other algorithms. In particular, it was demonstrated that the algorithm has the property of converging to its steady state, where signal separation occurs, much faster than other algorithms. Unlike other algorithms, the speed of convergence does not depend on the signals' power ratios and, hence, does not depend on the eigenvalue spread of the input correlation matrix.

During the current phase of the research we continue to examine other properties and advantages of the algorithm. We propose and study some applications to communication systems.

First, it is stated that the different structure exhibits a different level of complexity, particularly when it is used in very high frequencies, as in microwave communication. They presents different delay path to the signals and, hence, have different system bandwidth. This raises the question of existing complexity - bandwidth trade-off.

For every control system the question of stability of the steady state point is always asked. This is of particular interest with the bootstrapped algorithms, as they are highly nonlinear.

Clearly, when one proposes an adaptive algorithm, it is crucial to suggest an adequate real-time realization of the recursive control of the weights. Two methods of dynamic control are studied. One uses orthogonal perturbation sequences and is applied to the Backward/Backward structure and controlled via the power-power criterion. The other uses

weights dithering with a PN sequence and is applied to the Forward/Backward structure and controlled via the power-correlator criterion.

Delay controlled structures of the bootstrapped algorithms were also proposed in [1]. In this report we further study these structures and examine some of their applications. For the multi-input/multi-output application, we obtain new results in blind separation of signals when the environment is dispersive.

Possible applications of the algorithms are numerous. Two of such applications are in the fields of blind equalization and CDMA. In the second field, which will be studied further in the future, we will present some ideas which, although they are not implemented with the bootstrapped algorithm, would become a basis for comparison with these cancelers as they are applied in CDMA systems.

Section (2), below, is a technical summary of the study performed in this phase of the research and its results. Detailed reports on which this summary is based are given in the appendices of this document. Each appendix covers a specific part of the research and is written in a way that can be read independently of the other parts. Section (3) contains the conclusions and recommendations for further studies.

II. TECHNICAL SUMMARY

2.1 Bootstrapped Algorithm – Separation of Two Superimposed Signals

Figures 1, 2, and 3, the configuration of the three bootstrapped algorithms for separating two superimposed signals,

$$x_1(t) = a_{11}I_1(t) + a_{12}I_2(t) + n_1(t) \quad (1)$$

$$x_2(t) = a_{21}I_1(t) + a_{22}I_2(t) + n_2(t), \quad (2)$$

where $x_1(t)$ and $x_2(t)$ are the received superimposed signals at the first and second channels, respectively; $I_i(t)$ are the two data signals such that for $i=1,2$ I_i is an M-ary signal from set $\{\pm 1, \pm 3, \dots, \pm(\sqrt{M} - 1)\}$; and a_{ij} $i=1,2$ are channel coupling.

The performance criterion used in controlling the algorithm weights are, respectively, the power-power, correlation-correlation and power-correlation. A detailed discussion of these criteria and their corresponding steady state values is given in [1 of Appendix A] and the effect of noise on these values is given in [1 of Appendix B]. Note the insertion of the discriminator, which is crucial in obtaining unique steady state points.

2.1.1 Bandwidth Complexity Trade-Off

Comparing the steady state performance for the three systems [1 of Appendix A] showed that the symmetric B/B and F/F correlation systems have identical output signals and signal-to-interference ratios, while the asymmetric B/F system has slightly different output signals and signal-to-interference ratios at its output ports. Nevertheless, they are all equivalent in that the desired signal separation is ideal provided certain conditions or assumptions are valid; specifically, when the effects of added noise, quantization error, and non-zero circuit delay are assumed to be negligible.

An examination of Figures 1-3 illustrates that the three configurations require different

levels of hardware complexity. The F/F, correlation-correlation scheme is expected to be the most complex since it requires correlators with zero offsets, while the B/B, power-power scheme is the least complex. Also, the three canceler systems have different signal paths through the individual circuits and, thus, when they are applied for separation of dually polarized signals directly at microwave frequencies [5], or high I.F. frequencies (as in a dual polarized microwave terrestrial link [4]), different system delays are expected. Therefore, in Appendix A of this report we present a thorough study of the effects of these delay differences on the performance of each of the different cross-pol separator configurations. In particular, we analyze and discuss system bandwidth limitation due to these delays. For each of the three different schemes, we study the effect of delay on the optimal weights, on the outputs' powers and ratios. These effects are presented in terms of the autocorrelation of the two signals as well as the mixing ratios at the superimposed inputs. Using the relation between autocorrelation and signal bandwidth, we estimate the different separator bandwidths and, hence, we obtain the level of signal bandwidth limitation that each separator shows. These bandwidth limitations are presented in the form of a lower bound on the depth of cancellation as a function of both signals' bandwidths and system delays. The results obtained in this research are summarized in Tables 1-4 of Appendix A.

Based on these results, a practical example from a typical 14-18 GHz microwave circuitry shows that in the design of a B/F power-correlation separator, one can obtain as much as 20dB input-output improvement in the signal-to-interference ratio (in corresponding to a 40dB signal-to-interference ratio at the output when the input cross-pol ratio $|a_{12}| \simeq |a_{21}| \simeq 0.1$ or 20dB), when the signals' bandwidths are of the order of 500 MHz. On the other hand, with the same assumption regarding the accuracy of the delay timing, we can achieve almost 34dB improvement (54dB output signal-to-interference ratio), if the signal's bandwidths are only 100 MHz.

The B/B power-power scheme can handle only a small bandwidth. In fact, the power-

power scheme might be applied with reasonable success in lower frequency low bandwidth applications. For example, if the frequency band is in the 4-6 GHz range then the delays are of the order of 1 ns, so that 20dB input-output improvement in signal-to-interference can be accomplished with approximately 8 MHz. If the cancellation is performed in the 70 MHz IF frequency then the $\pi/4$ wavelength delay needed for the in-phase-quadrature weighting will restrict the delays to greater than 3.5 ns. If we assume delays of 5 ns, the bandwidth will be limited to approximately 2 MHz. Notice, however, that the power-power scheme requires much simple hardware than the other two schemes, so that the hardware complexity bandwidth trade-off is exhibited.

2.1.2 Stability Consideration

In every control system, adaptive or non-adaptive, one is always concerned in finding whether the steady state optimal weight solution is stable or unstable. For a linear dynamic system this is done by examining the eigenvalue of the corresponding weight control equation. It is of interest to examine and classify the steady state point of the three structures proposed in [1 of Appendix A]. However, since the control equations for the bootstrapped algorithms are highly nonlinear they require special attention. This is done in Appendix B of this report.

For each of the structures, we first derive the steady state equilibrium points for the no noise case. These are respectively;

1. The Backward/Backward, Power-power Scheme

$$w_{12opt1} = -\frac{a_{12}}{a_{22}} \triangleq -r_1 \quad w_{21opt1} = -\frac{a_{21}}{a_{11}} \triangleq -r_2 \quad (3)$$

and

$$w_{12opt2} = -\frac{a_{11}}{a_{21}} \triangleq -\frac{1}{r_2} \quad w_{21opt2} = -\frac{a_{22}}{a_{12}} \triangleq -\frac{1}{r_1} \quad (4)$$

2. The Forward/Forward, Correlator-Correlator Scheme

$$w_{12opt1} = -\frac{a_{12}}{a_{22}} \triangleq -r_1 \quad w_{21opt1} = -\frac{a_{21}}{a_{11}} \triangleq -r_2 \quad (5)$$

$$\text{and} \quad w_{12\text{opt}2} = -\frac{a_{11}}{a_{21}} \triangleq -\frac{1}{r_2} \quad w_{21\text{opt}2} = -\frac{a_{22}}{a_{12}} \triangleq -\frac{1}{r_1} \quad (6)$$

3. The Forward/Backward, Power-Correlator Scheme (assuming $\frac{a_{21}}{a_{11}} \neq \frac{a_{22}}{a_{12}}$);

$$w_{12\text{opt}1} = -\frac{a_{12}}{a_{22}(1 - \frac{a_{21}a_{12}}{a_{11}a_{22}})} \quad w_{21\text{opt}1} = -\frac{a_{21}}{a_{11}} = -r_2 \quad (7)$$

and;

$$w_{12\text{opt}2} = -\frac{a_{11}}{a_{21}(1 - \frac{a_{22}a_{11}}{a_{21}a_{12}})} \quad w_{21\text{opt}2} = -\frac{a_{22}}{a_{12}} = \frac{1}{r_1}. \quad (8)$$

The stability parameters are found by linearizing the control equation around the equilibrium point of equations (3) to (8). This linearization process leads to a matrix differential equation of the form

$$\dot{\underline{\mathbf{X}}} = \mathbf{A}\Delta\underline{\mathbf{w}} \quad (9)$$

where $\Delta\underline{\mathbf{w}} = [\Delta w_{12}, \Delta w_{21}]^T$ and

$$\dot{\underline{\mathbf{X}}} = \left[\frac{\partial X_1(w_{12} + \Delta w_{12}, w_{21} + \Delta w_{21})}{\partial w_{12}}, \frac{\partial X_2(w_{12} + \Delta w_{12}, w_{21} + \Delta w_{21})}{\partial w_{21}} \right]^T \quad (10)$$

where

$$\mathbf{A} = \begin{bmatrix} \frac{\partial^2 X_1}{\partial w_{12}^2} \Big|_{\mathbf{w}_{\text{opt}}} & \frac{\partial^2 X_1}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} \\ \frac{\partial^2 X_2}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} & \frac{\partial^2 X_2}{\partial w_{21}^2} \Big|_{\mathbf{w}_{\text{opt}}} \end{bmatrix}. \quad (11)$$

The functions $X_1(\cdot, \cdot)$ and $X_2(\cdot, \cdot)$ are different for each structure. They represent the deriving term for the weights' recursive equations and depend on the criterion used. Therefore (9) describes the change in time of the weights deriving terms from an initial $\Delta\underline{\mathbf{w}}$ away from

the optimal weights. If for example the eigenvalue of \mathbf{A} are all negative, the weight deriving terms $X_1(\cdot, \cdot)$ and $X_2(\cdot, \cdot)$ will decrease to zero and, hence, to the steady state point. etc.

The stability of equilibrium points depend on the eigenvalues of matrix \mathbf{A} . Considering the characteristics equation of the matrix \mathbf{A} from $|\lambda\mathbf{I} - \mathbf{A}| = 0$, we can find the eigenvalues of \mathbf{A} by solving

$$\lambda^2 - b\lambda + c = 0 \quad (12)$$

where

$$b = \left. \frac{\partial^2 X_1}{\partial w_{12}^2} \right|_{\mathbf{w}_{opt}} + \left. \frac{\partial^2 X_2}{\partial w_{21}^2} \right|_{\mathbf{w}_{opt}} \quad (13)$$

$$c = \left. \frac{\partial^2 X_1}{\partial w_{12}^2} \right|_{\mathbf{w}_{opt}} \left. \frac{\partial^2 X_2}{\partial w_{21}^2} \right|_{\mathbf{w}_{opt}} - \left. \frac{\partial^2 X_1}{\partial w_{12} \partial w_{21}} \right|_{\mathbf{w}_{opt}} \left. \frac{\partial^2 X_2}{\partial w_{12} \partial w_{21}} \right|_{\mathbf{w}_{opt}} \quad (14)$$

The nature of the eigenvalues of \mathbf{A} in the complex plane, or equivalently the relation between b and c defined in (13) and (14) determines the classification of the equilibrium point.

After finding the matrix \mathbf{A} for each of the three structures and for both equilibrium points of equations (3) to (8), we determine the condition for the stability of each structure. For the symmetric power-power and correlation-correlation, it was found that $\underline{\mathbf{w}}_{1opt}$ (Eq. (3) and (5)) are stable provided that the step size μ is positive and the effects of the discriminators are such that

$$\delta_{21}\delta_{12} > \delta_{11}\delta_{22} \quad (15)$$

That is, the discriminator always favors different signals at different outputs. Under the same conditions, the other equilibrium point $\underline{\mathbf{w}}_{1opt}$ is an unstable saddle point.

For the power-correlator scheme, a stable $\underline{\mathbf{w}}_{1opt}$ occurs if beside the condition in (15), the effect of the second discriminator is such that it is more pronounced than the cross-pol

power;

$$\frac{\delta_{21}}{\delta_{22}} = \frac{E\{I_2^2(n)\}}{E\{I_1^2(n)\}} \left(\frac{a_{12}}{a_{11}}\right)^2 \quad (16)$$

2.1.3 Dynamic Studies

With any adaptive algorithm, one first tries to define the performance criterion that the algorithm would seek to optimize, by finding the control weights which give the optimal performance. However, in a real time variable environment condition, the designer must devise a suitable algorithm which recursively would try to reach the steady state optimal weight values. For the bootstrapped algorithms, recursive procedures were suggested to simultaneously control the two weights. They require the knowledge of the gradient of the output power (when minimum power criterion is used) or the gradient of the square of correlation between outputs (when minimum correlation is sought). Instead, one may use an estimate of the gradient. Such possible estimation is the sample value of the gradient. Also a direct real-time measurement of a gradient of a random variable is difficult if not impossible, particularly if this is performed with signals of very high frequencies as in communication applications.

Without direct measurement, we present in this report two techniques for reaching the optimal weights with a recursive weight-updating procedure using estimates of the gradients. With the first technique, the estimates of the gradient are obtained by applying two orthogonal perturbation sequences to the two weights simultaneously. With the other technique, the weights are dithered with PN sequences. These are described next.

2.1.3.1 Dynamic Study of B/B Power-Power Scheme Using Orthogonal Perturbation Sequences

Figure (4) presents a Backward/Backward, power-power scheme of a cross-pol interference canceler controlled by orthogonal perturbation sequences. Two such sequences $p_1(n)$ and $p_2(n)$ with an amplitude Λ (taken to be sufficiently small) are added to the current

weight $w_{12}(i)$ and $w_{21}(i)$. The two outputs are then multiplied respectively by $p_1(n)$ and $p_2(n)$. The results give a measure of the corresponding change of the output powers.

In Appendix C of this report, we analyze the control structure of Figure (4). We show that the estimate of the gradient obtained with the perturbation method, converges in the mean to the actual gradient and, hence, the weights converge in the mean to the optimal weight. This is shown by examining the dynamic behavior of the error between the estimated weight and its desired optimal value. The condition for such convergence is also stated.

2.1.3.2 Dynamic Study of B/F Power-Correlator Scheme Using Weight Dithering with PN Sequences

Figure (5) presents a Backward/Forward, power-correlator bootstrap separator with weight dither control. A PN sequence is added to the weight w_i^f . The output power of $y_1(t)$ is multiplied after eliminating the DC component by the PN sequence. A low pass filter smoothed the control signal before it was added recursively to the previous weight value.

In Appendix D of this report we analyze this dither control structure. Since a possible candidate for such scheme is a microwave dual channel separator, we prefer to use the analog presentation of signal instead of the discrete one. Therefore, we examine the change in the weights versus time instead of checking convergence of the recursive equations. We found that under certain conditions which relate signals' power, system gain and dither magnitude the dither control results in weight steady state value that gives signal separation. However, depending on the magnitude of the dither, some small interference residue is retained. This sensitivity of the quality of separation to power of dither and other systems is also studied.

2.2 Bootstrapped Algorithms – Wideband Signal Separators

In [1], we performed some preliminary study in applying the bootstrapped algorithms to wideband signals. There, we suggested structure whose controlled elements are delays instead of complex weights or digital filters. Two applications of the wideband structures are considered in Appendices E and F of this report.

Figure (6) describes the scenario where two wideband sources are received by an array of two sensors. The outputs of the two sensors are:

$$\begin{aligned}z_1(t) &= s_1(t - D_1) + s_2(t - D_2) + e_1(t) \\z_2(t) &= s_1(t + D_1) + s_2(t + D_2) + e_2(t)\end{aligned}\tag{17}$$

where $s_1(t)$ and $s_2(t)$ are the signals radiated from the 1st and 2nd sources respectively. $e_1(t)$ and $e_2(t)$ are the additive noise processes in each of the sensors and D_i is given by,

$$D_i = \frac{d}{2c} \sin \theta_i \quad i = 1, 2\tag{18}$$

d is the distance between the sensors, c is the propagation velocity and θ_i , $i = 1, 2$ is the bearing of the i th source. We assume that the random signals $s_1(t)$, $s_2(t)$, $e_1(t)$ and $e_2(t)$ are mutually statistically uncorrelated.

In fact, the model described above is applicable to many communication and signal processing problems which can be put in three groups: bearing estimation (source localization), source separation, and interference cancellation.

- For bearing estimation (source localization) one is interested in \hat{D}_1 and \hat{D}_2 and not in the signal estimates $\hat{s}_1(t)$ and $\hat{s}_2(t)$. In this application, particularly in active radar or sonar, much is known about $s_1(t)$ and $s_2(t)$.
- For source separation, the objective is to get the cleanest possible version of $s_1(t)$ and $s_2(t)$, while D_1 and D_2 are nuisance parameters. Indeed, in some cases there is some

prior knowledge about θ_1 and θ_2 which can be used to restrict the possible values of D_1 and D_2 .

- For interference cancellation, the objective is to get a clean replica of one of the signals, say $s_1(t)$ while the other signal, as well as D_1 and D_2 are nuisance parameters which, in some applications, are partially known.

In Appendix E of this report we present adaptive systems which receive $z_1(t)$ and $z_2(t)$ at the input, deliver as output signals $y_1(t) = \hat{s}_1(t)$ and $y_2(t) = \hat{s}_2(t)$, and simultaneously provide the estimates of D_1 and D_2 from which the bearings of the two signals θ_1 and θ_2 can be derived. However, in order to enable separation, some information (statistical or physical) about the signals $s_1(t)$ and $s_2(t)$ is required. This information will be utilized in the design of the control loop. It enables initiation of the separation procedure. If further information is available (such as knowledge, for example, of one of the two bearings) the same algorithm, which then bootstraps itself, will result in performance improvement.

Two different configurations of the delay control bootstrapped separation algorithm are given in Figure (7); for the Backward/Backward and Forward/Forward structures. They are examined using small-error analysis of the outputs spectra. It is shown that while both configurations are able to perform simultaneous source separation and delay estimation (and hence direction-of-arrival), they exhibit several differences.

1. In the ideal situation, i.e., when $\psi_i = \tau_i - D_i = 0$, (τ_i is the controlled delays) is zero for $i=1,2$. The *BB* configuration provides undistorted versions of the signal waveforms, while the *FF* configuration decouples the signals but provides, at the outputs, a filtered version of them. Therefore, if an exact replica of the source signal is needed, the *BB* configuration is to be preferred.
2. In the presence of additive noise, the *FF* configuration provides better estimates of

the unknown source directions (the delays) than the *BB* configuration. Thus, the *FF* structure is better for direction of arrival (DOA) estimation.

3. The *BB* configuration exhibits difficulties in the presence of noise when applied at baseband.

In comparison to narrowband separators, we conclude the following:

1. The F/B configuration can only be implemented in the narrowband case.
2. Unlike the narrowband, when the B/B separator is used in the broadband case, decorrelation cannot be replaced by minimization of the output powers.
3. If the signals are only known to be uncorrelated and have the same spectrum, then, as in the narrowband case, a discriminator is needed to achieve separation. If, however, their spectrum is known to be different, then such a discriminator is unnecessary.

In Appendix F of this report we present a novel approach for rejecting a broadband interference from unknown direction when received by an array of two sensors. Two configurations of such an approach are presented. Both perform perfect interference cancellation when the input signal-to-noise ratio (SNR) is large enough, and do it much faster than LMS canceler. However, additive noise causes performance degradation to both. It is shown that no general claim can be made about the superiority of one configuration with respect to the other. The output signal to interference-plus-noise ratio (SNIR) depends on the spatial separation between the interference and the desired signal, as well as on the interference-to-noise ratio (INR), in a different manner for both configurations. This appendix also contains guidelines for the choice of one or the other configuration in different scenarios.

2.3 Bootstrapped Algorithms – Multi-Input, Multi-Output Separators

In many applications such as in neural networks and signal separation in multi channel CDMA, there is a need for multi-input, multi-output separators. Extension of the two-input/two-output separators was suggested and its performance was examined in [1 of Appendices E & F]. However, emphasis was put on nondispersive channel. In Appendix G of this report, we studied the performance of the backward/forward structure when it separates multi-signal composites in a multi-channel dispersive environment.

2.4 Bootstrapped Algorithm – Other Applications

Beside the application of the bootstrapped algorithm to cross-pol cancellation which was considered in [1] and in this report, initial work was performed in this phase of research in two important applications, blind equalization and spread-spectrum multi-signals separation. The first is important particularly in time division multiple access (TDMA) and the other in code division multiple access (CDMA) communications.

Decision feedback is generally preferred to other equalization methods, as it can compensate amplitude distortion with minimal noise enhancement. This structure is therefore used in our research. Decorrelation of the equalizer output is taken as performance criterion. As such this is a novel way of equalization which demonstrate many advantages (see Figure (8)).

In Appendix H, we show that this blind equalizer converges to correct equilibrium despite error propagation. It is also shown, using simulation, that the algorithm converges to the right equilibrium regardless of the initial condition, and, hence it is globally convergent.

In Appendix I of this report we included the results of research within adaptive scheme which separates two CDMA signals (see Figure (9)). Although it does not show bootstrapping phenomenon, it is rather different from the other canceler schemes in that it contains

decisions in the cancellation path. This study would be used in the future as a reference for comparison with the bootstrapped scheme for CDMA signal separators.

III. CONCLUSIONS AND RECOMMENDATIONS

In the period of this report we continue to analyze and evaluate performance of the bootstrapped algorithms. As in the previous year's work, reported in [1], we considered the three different structures, the B/B, the F/F and the B/F. It was then concluded that the algorithms result in signal separations. Also noted that unlike other algorithms, the speed of convergence does not depend on the signals' power ratios and hence does not depend on the eigenvalue spread of the input correlation matrix. Furthermore, the algorithms do not require supervisory reference and, hence, deserve to be named blind signal separator.

In this research, we first considered the question of bandwidth each of the three structures can tolerate. It was found that both the F/F and the B/F can be made reasonably wideband if it is used at microwave or higher band. The B/B is quite limited in bandwidth. The former structure require correlator in its implementation while the later require only power measurement. hence, we concluded on complexity - bandwidth trade-off.

The question of stability was also addressed. Linearization around the steady state point led us to conclude that values of optimal weights which result in signal separation are stable points.

Two schemes of real-time control of the weights were proposed, one uses orthogonal perturbation sequences and the other weight dithering with PN sequences. Both showed satisfactory results in that the weights converge in-the-mean to their optimal values.

Using delay control for wideband application, we show that the algorithms could perform simultaneous spatial separation of wideband signal while estimating their direction of arrival. It was also shown that delay controlled bootstrapped structure can successfully be used to cancel wideband interferences.

In the extension to multi-input/multi-output, we show that the algorithms can be used also to handle dispersive environment.

Finally we note that the previous research dealt mainly with dual channel cross-pol interference. This report contains some results related to the problem of blind equalization of digital data. It also contains an adaptive canceler for CDMA, which would become basis for comparison in future research when the bootstrapped algorithm is used.

The work carried out led us to make number of recommendations for further study.

a. Adaptive Algorithms - Implementation

We suggest further study of the theoretical and practical aspect of implementing the bootstrapped algorithms. In particular,

- Suggest specific control strategies to adjust the weights/time delays of the bootstrapped algorithm for the two-input/two-output configuration.
- Study convergence properties of the algorithm for the multi-input/multi-output case. Specifically, we will determine the feasibility of using dithering of the weights to improve convergence and enable time-multiplexing of control hardware.
- Continue to investigate the link between eigenvalue spread and performance of the bootstrapped algorithms.

b. Application to Signal Separation in Satellite Communication

- Further examine the possibility of cross-pol cancellation in satellite communication channels. Extend the current results obtained with M-ary QAM dual polarized signal to other possible modulation schemes.
- Develop the algorithm for handling signals such as frequency hoppers.

- Examine the effect of multipath of signal separation performance.
- Improve the quality of separation of code division multiplexed signals.

c. Adaptive Blind Equalizer

Bootstrapped algorithms in particular the minimum correlation schemes, are good candidates for blind equalizer, therefore, we propose

- Investigate the application of the bootstrapped algorithms to the blind decision feedback equalizer.
- Study the performance of such a scheme and compare it with the performance of the blind feed forward equalizer and the regular (non-blind) decision feedback equalizer.
- Examine convergence property of the algorithms.
- Examine ill convergence of the algorithm, and the conditions for false lock.

d. Application of Neural Networks

The bootstrapped algorithms can be applied as a new leaning paradigm and topological structure for designing a new class of neural networks. This new class of neural networks has the potential to revolutionize and improve various existing applications in the field of neural networks. Fast convergence, the self-organizing property, and reference signals which are no longer required, are among the major advantages that should be verified and investigated. In particular, we propose

- Map the bootstrapped algorithm onto a neural network framework. The resulting network, which belong to the class of recurrent self-organizing neural networks. will be referred to as the generalized.

- Demonstrate supremacy of the bootstrapped network in terms of performance, convergence and stability of the network will not only be demonstrated through simulations but we will attempt to prove it analytically as well.
- Explore potential applications including channel equalization, separation of superimposed signals, pattern recognition and prediction.

e. Demonstration Modules for Signal Processing Workstation It is important to try and build software modules for each of the applications suggested. It is proposed to use COMDISCO Signal Processing Workstation (SPW) version 3.0, for these modules.

IV. REFERENCES

- [1] Y. Bar-Ness, A. Dinç and H. Messer-Yaron, *The Bootstrapped Algorithm: Fast Algorithm for Blind Signal Separation*, Final Technical Report, RL-TR-93-24, April, 1993.
- [2] Y. Bar-Ness and J. Rokah, "Cross-Coupled Bootstrapped Interference Canceler." *1981 Antenna and Propagation Symposium*, conference proceedings, pp.292-295.
- [3] Y. Bar-Ness, J.W. Carlin and M.L. Steinberger, "Bootstrapping Adaptive Cross Pol Cancelers for Satellite Communication," *Proc. of IEEE Int. Conf. on Comm. (ICC)*, no. 4F5, June 1982, Philadelphia, PA.
- [4] J.W. Carlin, Y. Bar-Ness, S. Gross, M.L. Steinberger, and W.E. Studdiford, "An IF Cross-Pol Canceler for Microwave Radio Systems," *Journal on Selected Areas in Comm. - Advances in Digital Comm. by Radio*, Vol. 5 AC-5, No.3, pp. 502-514, April 1987.
- [5] Y. Bar-Ness and J.W. Carlin, "Cross-Pol Canceler Architecture for Microwave Radio Applications," *Int. Conf. on Comm., Conference Proceedings* paper No. 52.5, Seattle, WA, June 1987.
- [6] Y. Bar-Ness, "Effect of Number of Taps on Cross-Pol Canceler Performance for Digital Radio," *The IEEE Global Comm. Conf. (Globecom)* 1987, paper no. 31.7.

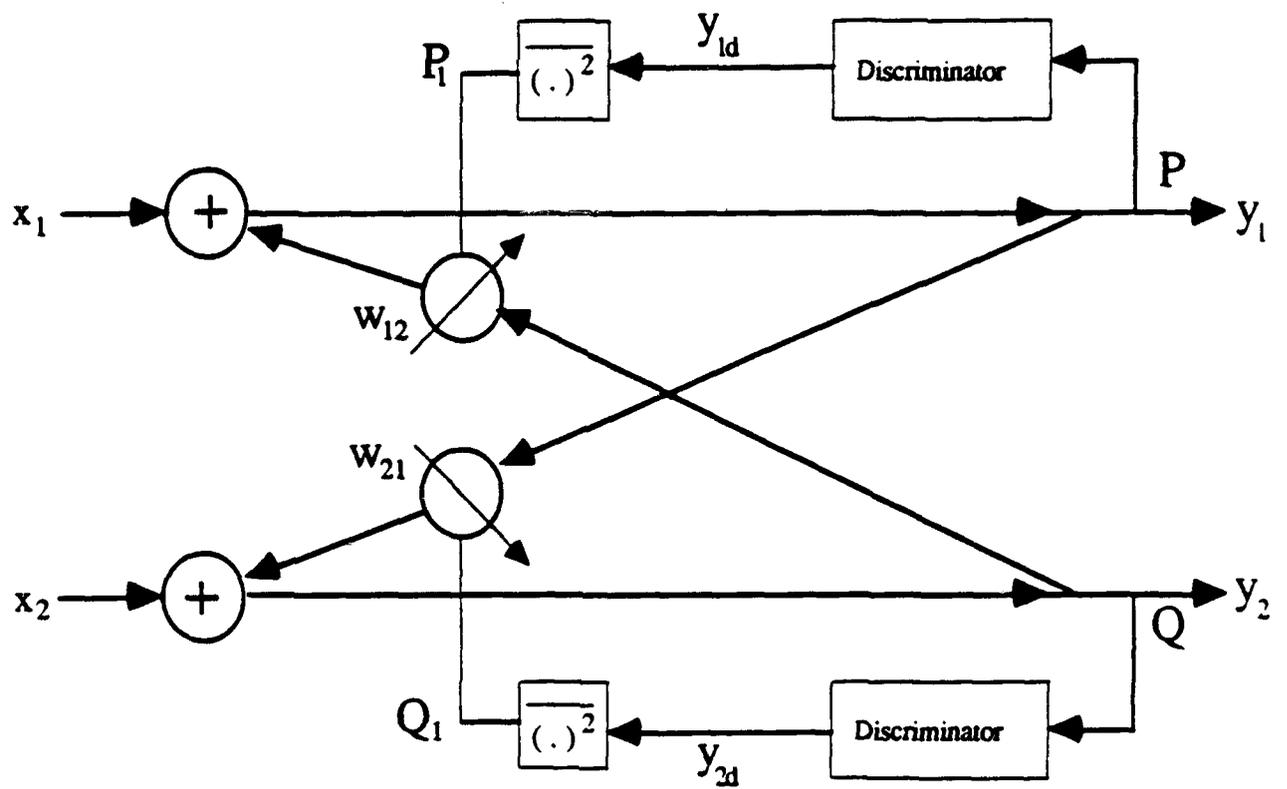


Figure 1 The Power-Power Cross-Pol Interference Canceler.

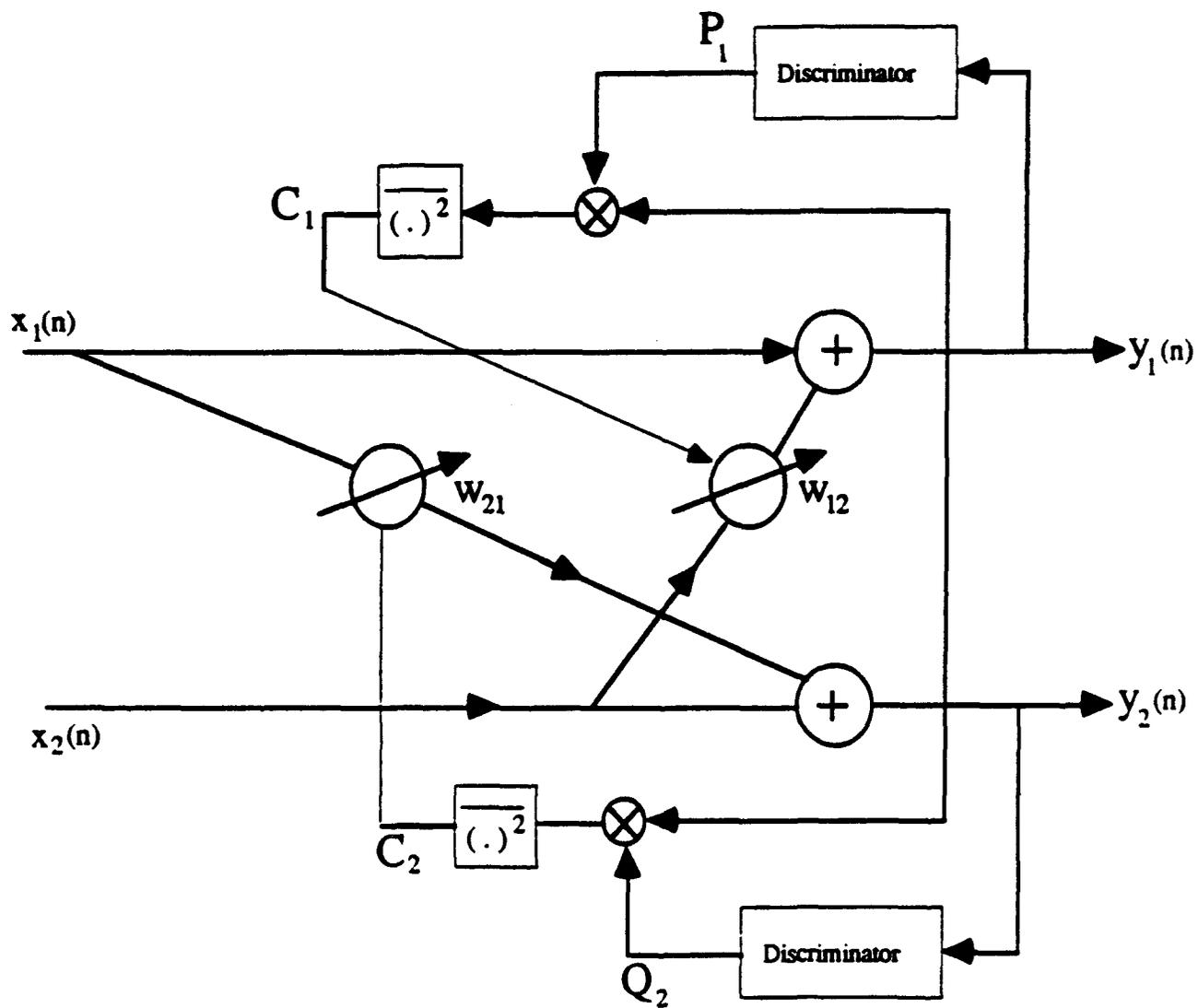


Figure 2 The Correlator-Correlator Cross-Pol Interference Canceler.

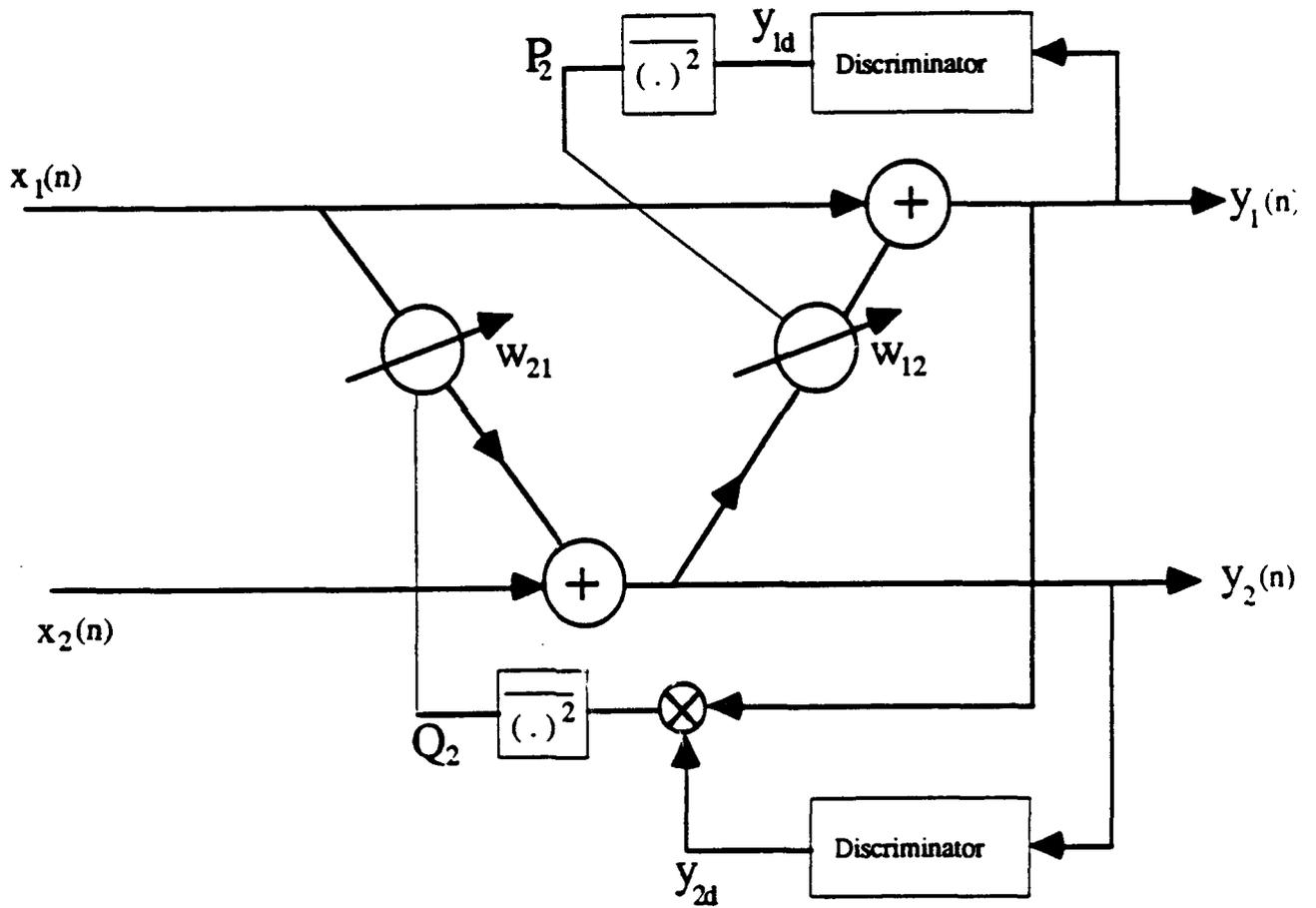


Figure 3 The Power-Correlator Cross-Pol Interference Canceler.

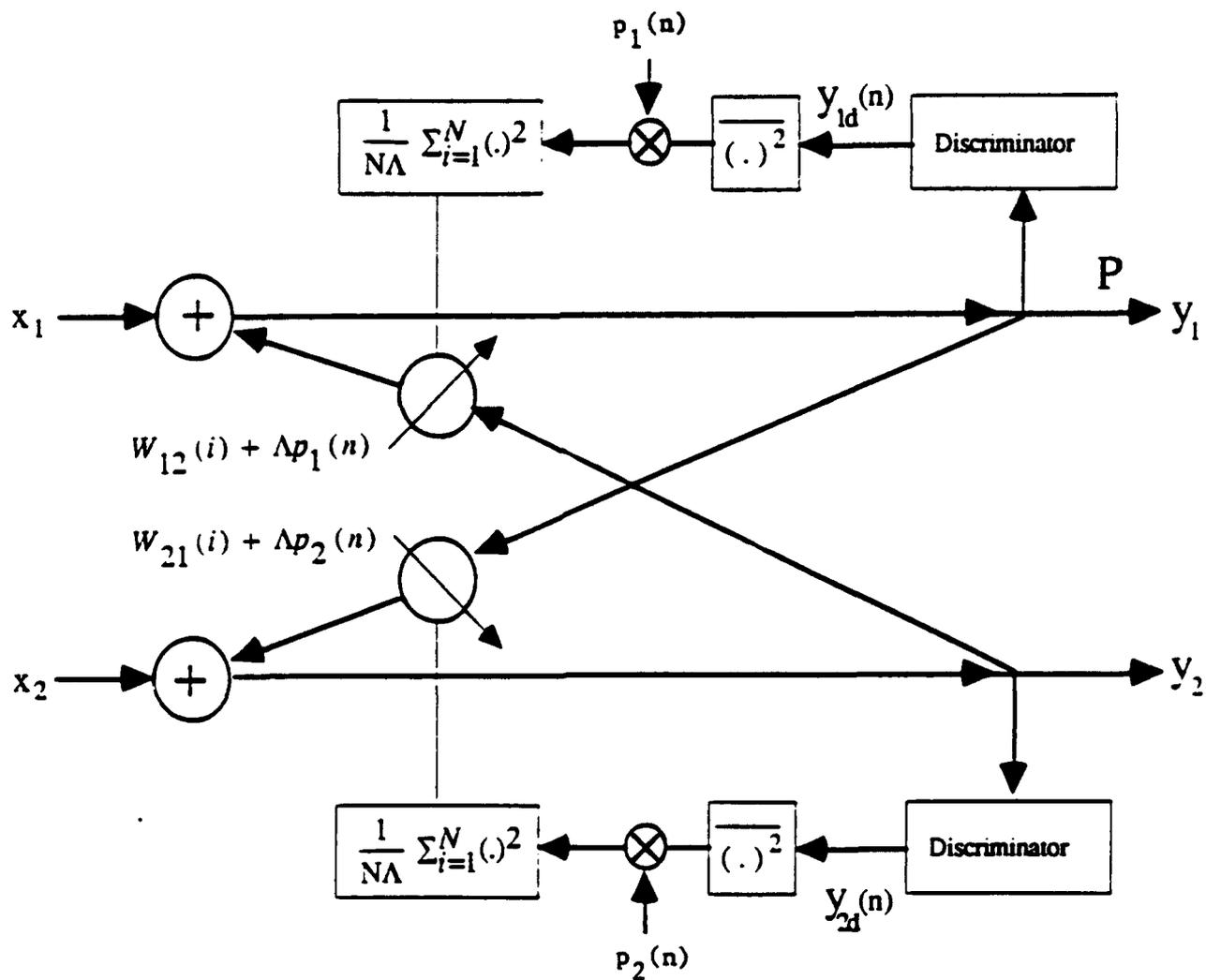


Figure 4 Power-Power Cross-Pol Interference Canceler controlled by orthogonal perturbation sequences.

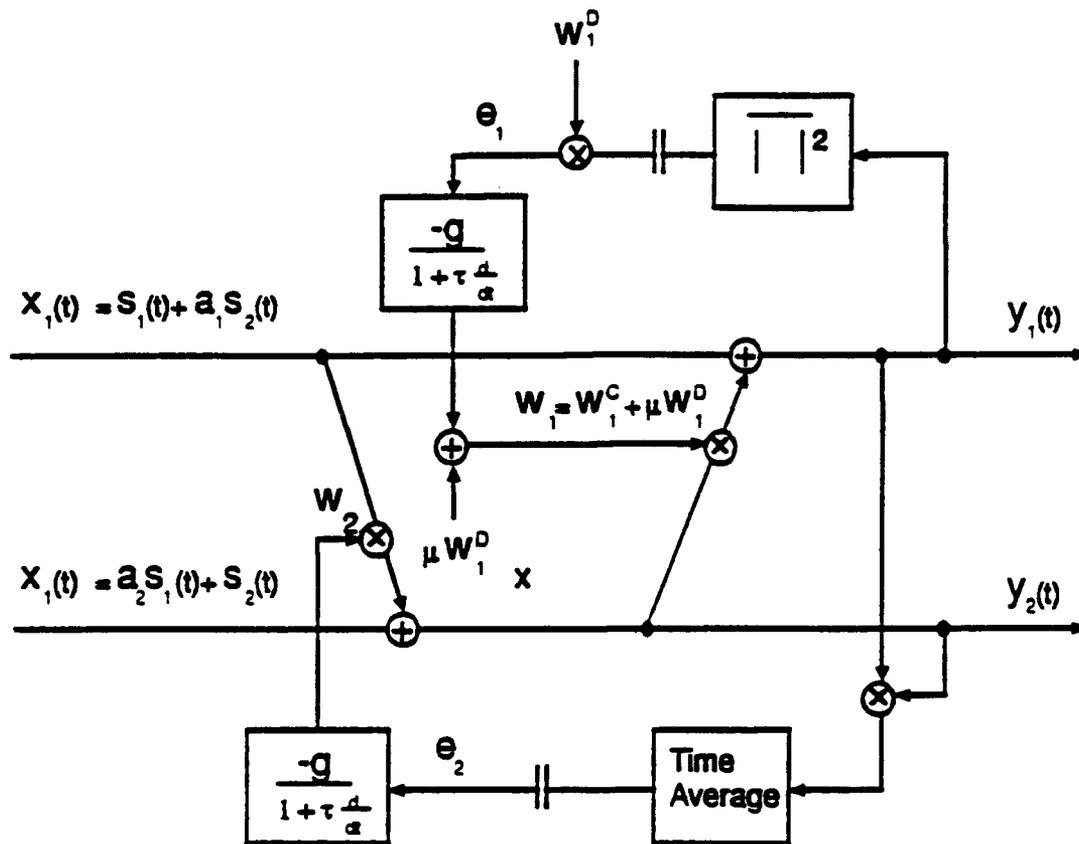


Figure 5 A diagram of the Forward/Backward bootstrap blind signal separator with dither control.

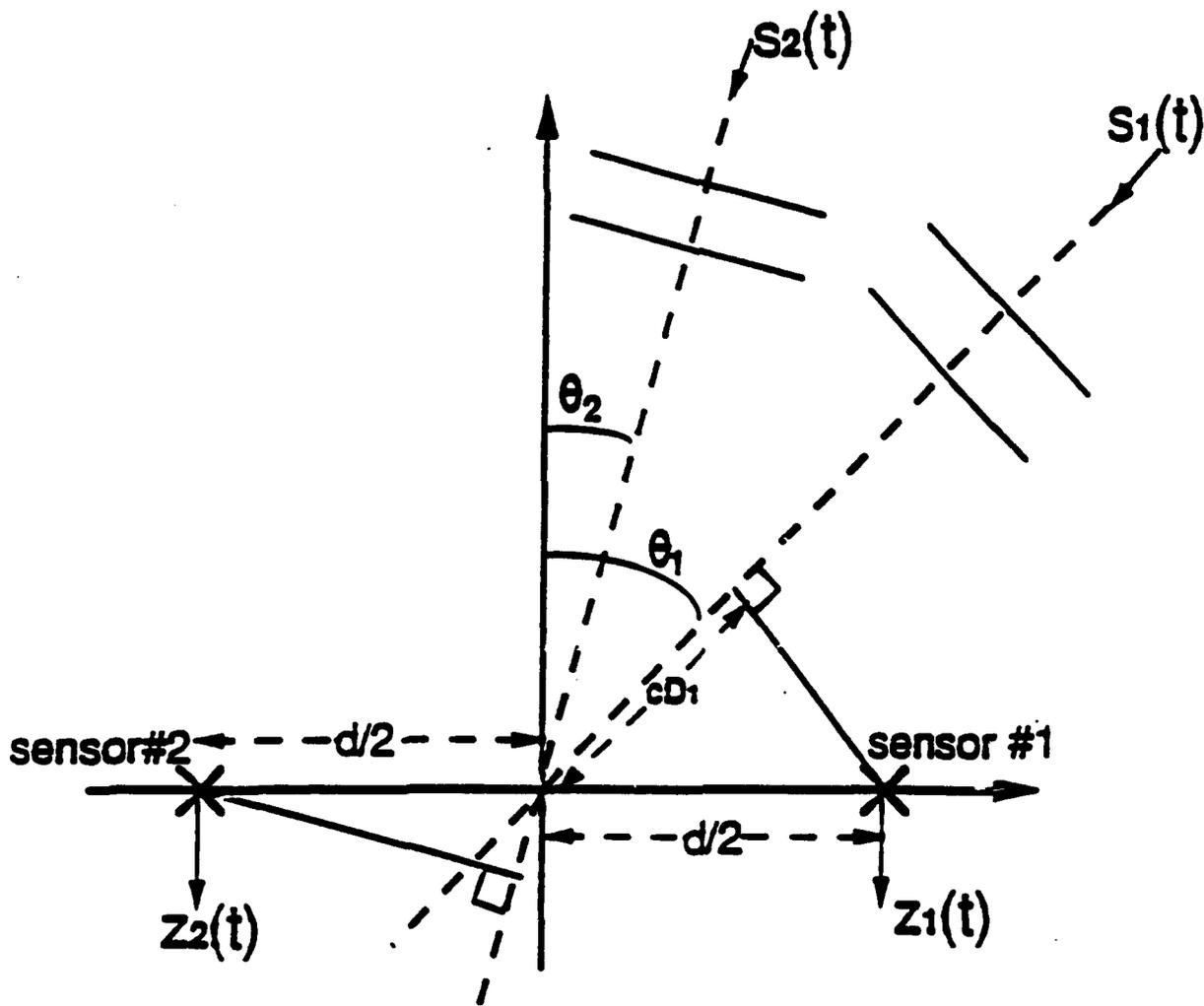


Figure 6 The model of the spatial signal mixture.

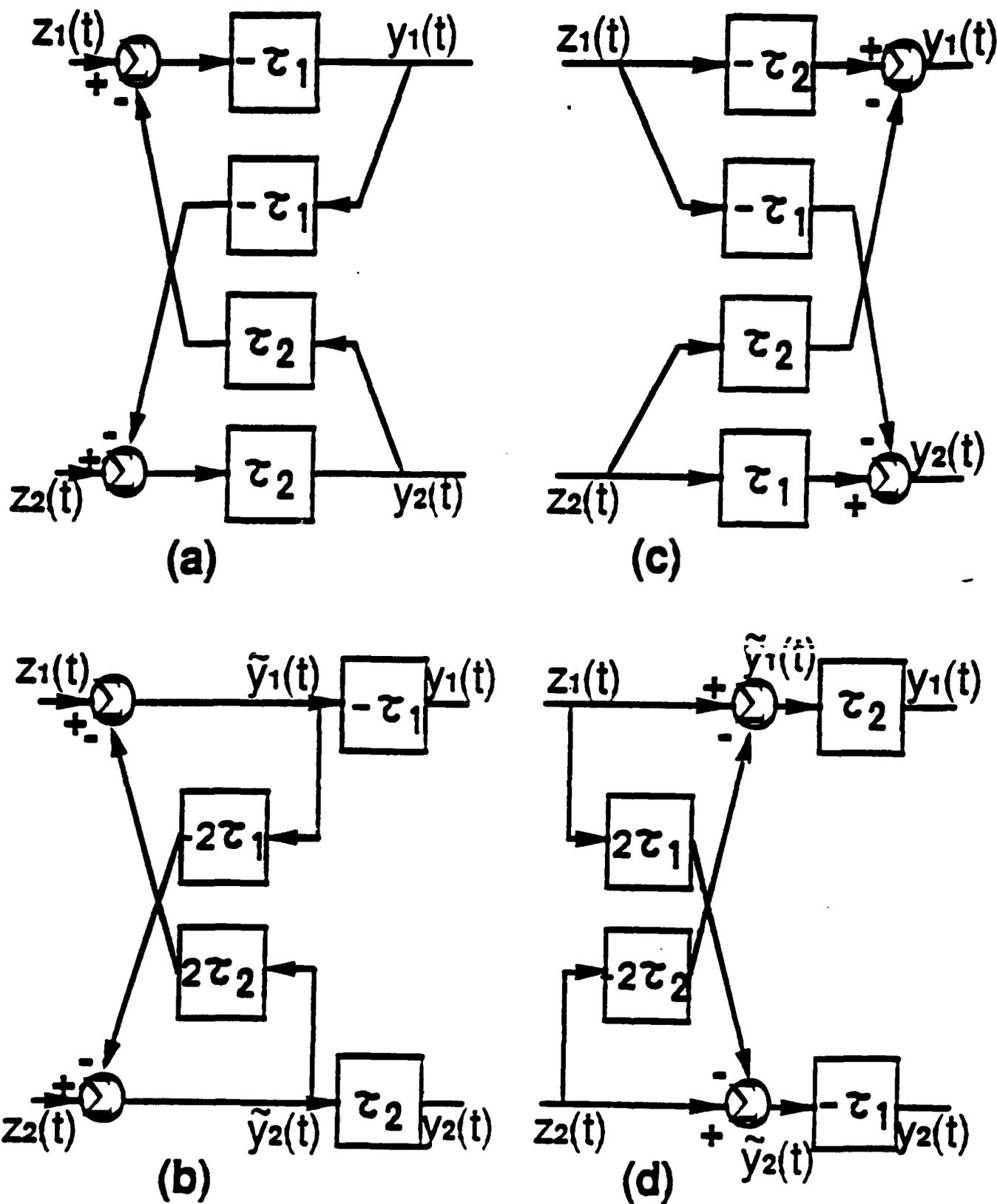


Figure 7 The different structures of the delay control separators - (a) the basic BB structure, (b) the modified BB structure, (c) the basic FF structure, (d) the modified FF structure.

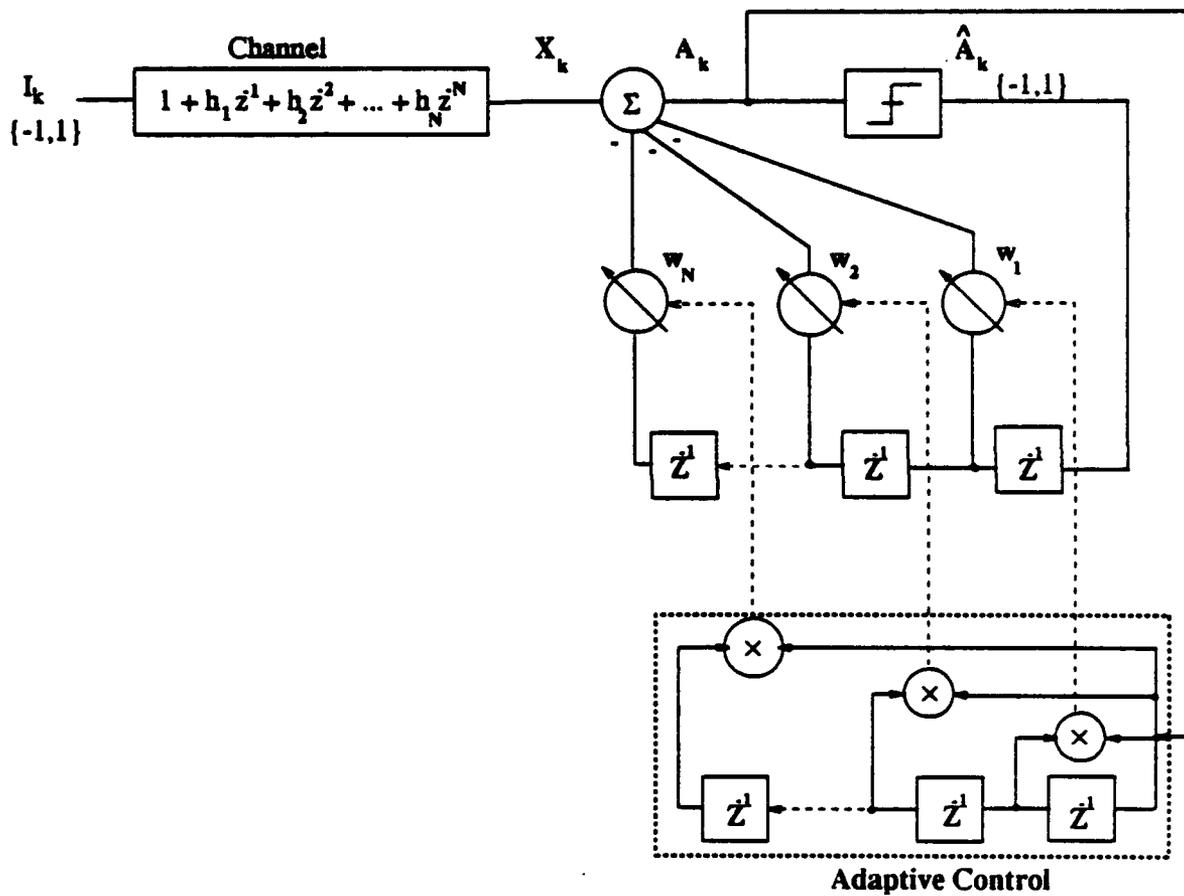


Figure 8 Channel and DFE model.

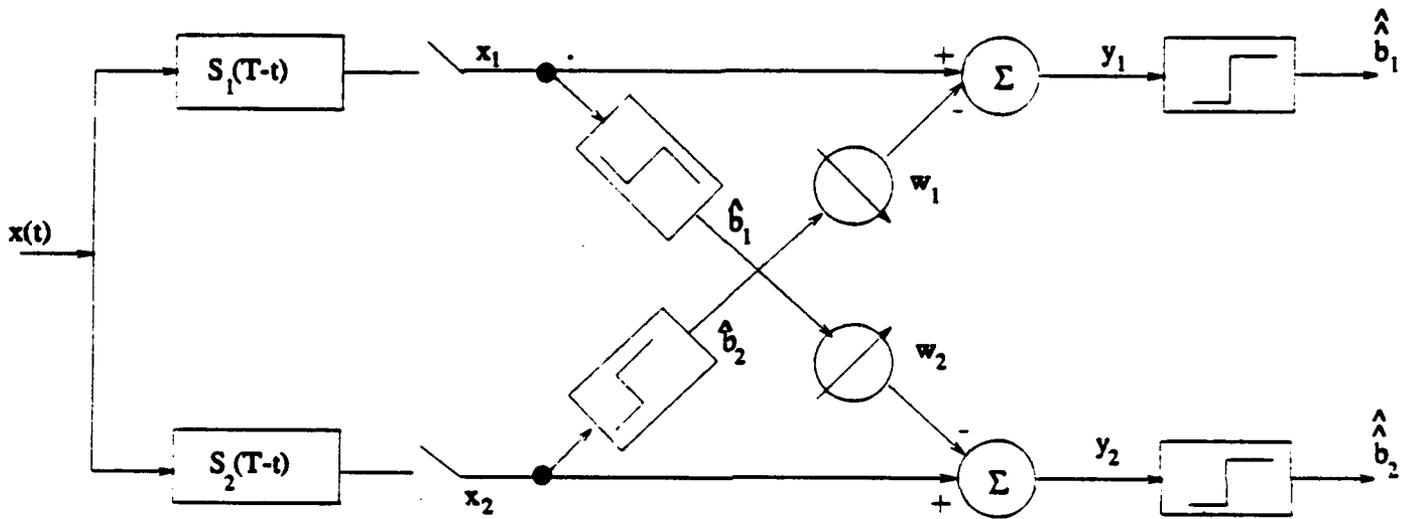


Figure 9 Two-stage receiver for two synchronous users.

APPENDIX A
BOOTSTRAPPED ADAPTIVE SEPARATION OF TWO
SUPERIMPOSED SIGNALS -
BANDWIDTH COMPLEXITY TRADE-OFF

by
Yeheskel Bar-Ness

I. INTRODUCTION

Three different bootstrapped systems were introduced in [1] for the purpose of cross-polarization (cross-pol) interference cancelation and signal separation. They differ by the criterion used to determine the performance index for realizing the optimum weight, gain and phase for each canceler module (e.g., minimization of each of the interfering signal powers at the output separately or the correlation between the two different outputs). The three systems: power-power, correlation-power and correlation-correlation cancelers are shown in Figures 1-3. A detailed steady state analysis of these separators is included in [2]. It was shown in [2] that each of these cancelation arrangements results in a power separation (a highly desired signal-to-interfering signal ratio at both output ports).

Since there is no need for a supervisory reference signal in the form of a training sequence, decision feedback, etc., these separators are in the category of blind separators. Comparing the steady state performance for the three systems showed that the symmetric power-power and correlation-correlation systems have identical output signals and signal-to-interference ratios, while the asymmetric power-correlation system has slightly different output signals and signal-to-interference ratios at one of its output ports. Nevertheless, they are all equivalent in that the desired signal separation is ideal provided certain conditions or assumptions are valid; specifically, when the effects of added noise, quantization error, and nonzero circuit delay are assumed to be negligible.

Since the introduction of the bootstrapped idea for signal separation (power-power) in [3], the three different bootstrapped scheme has been proposed and implemented for different applications. In particular, the power-correlation scheme was used in dual channel downlink cross-pol canceler operating over the COMSTAR satellite system [1]. Later, the same scheme was included in the design of the cross-pol canceler for dual-channel terrestrial digital microwave radio communication [4]. The power-power scheme, on the other hand, was proposed for tactical communications application in [5]. Other references related to the performance of the bootstrapped signal separator (interference cancelation) including extension to the multi-signals case are contained in [6]-[9].

Independently during the mid-eighties, a group of European researchers addressed the separation problem, in particular, for neural networks applications (see [10] for reference to these works). Recently, a third group of signal processing researchers applied a similar idea to speech signal separation [11].

Examination of Figures 1-3 illustrate that the three configurations require different levels of hardware complexity. The correlation-correlation scheme is expected to be the most complex since it requires correlators with zero offset, while the power-power scheme is the least complex. Also the three canceler systems have different signals' paths through the individual circuits and thus, when they are applied for separation of dually polarized signals, directly at microwave frequencies [1] or high I.F. frequencies as in dual polarized microwave terrestrial links [4], different system delays are expected.

The purpose of this paper is to examine the effect of these delay differences on the performance of each of the different cross-pol separator configurations. In particular, any bandwidth limitation due to these different delays is analyzed and discussed.

In the next three sections of this paper, we consider, respectively, the power-power, the power-correlator and correlator-correlator schemes. In each section we study the effect of these delays on output powers. Power ratios are considered next. In particular, the lower

bound on system bandwidth is estimated and used to calculate the bounds on input-output improvement in signal-to-interference ratios. Assuming uniform input signal spectra, the actual system bandwidth and the improvement signal-to-noise ratios are calculated. Finally, in section V the results are summarized and compared. Numerical examples are depicted to show the suitability of the different schemes to different applications. It is important to emphasize that in our analysis we will use the assumption that delays are quite small relative to the signals' correlation times.

II. THE POWER-POWER SCHEME

Figure 1 depicts the separator structure termed power-power. That is, the corresponding separator weights are chosen to minimize the powers at the two outputs respectively. In this section, we will examine the effect of system delays on the optimal weights, on the outputs' power and ratios and finally conclude with the level of bandwidth limitation as a result of these delays.

2.1 Effect of System Delays on the Optimal weights

Directly from Figure 1, we can write

$$\begin{aligned}
 v_p(t) = & s(t - \tau_1) + bn(t - \tau_1) + \beta cs(t - \tau_1 - \tau_2) + \beta n(t - 2\tau_1 - \tau_2) \\
 & + \alpha \beta v_p(t - 2\tau_1 - 2\tau_2)
 \end{aligned} \tag{1}$$

where we used, $v_1(t) = s(t) + bn(t)$, $v_2(t) = cs(t) + n(t)$ as the two inputs (in complex notation), b and c are complex valued with $|b|^2$ and $|c|^2$ as the input interference coupling ratios. τ_1 and τ_2 represent the path delays as they are depicted in Figure 1. α and β are the complex weights controlled by the two processors. The signal $s(t)$ and $n(t)$ are assumed to be zero-mean uncorrelated stationary complex processes. In cross-pol signals separation

applications $|b|$ and $|c|$ are much smaller than one. Using (1) we can easily find

$$\begin{aligned}\overline{|v_p(t) - \alpha\beta v_p(t - 2\tau_1 - 2\tau_2)|^2} &= (1 + |\alpha\beta|^2)P - 2\text{Real}[\alpha\beta]R_{v_p}(2\tau_1 + 2\tau_2) \\ &= (1 + |\beta c|^2)\overline{|s|^2} + 2\text{Real}[\beta c]R_s(\tau_1 + \tau_2) \\ &\quad (|b|^2 + |\beta|^2)\overline{|n|^2} + 2\text{Real}[\beta b^*]R_n(\tau_1 + \tau_2)\end{aligned}\quad (2)$$

where the overbar stands for the expected value and $R_{v_p}(\tau)$, $R_s(\tau)$ and $R_n(\tau)$ are the auto-correlation function of the output stochastic processes $v_p(t)$, the signal $s(t)$ and the signal $n(t)$, respectively and $P = \overline{|v_p(t)|^2}$.

Equation (2) can be rearranged to become,

$$\begin{aligned}|1 - \alpha\beta|^2 P + 2\text{Real}[\alpha\beta](P - R_{v_p}(2\tau_1 - 2\tau_2)) \\ = |1 + \beta c|^2 \overline{|s|^2} - 2\text{Real}[\beta c](\overline{|s|^2} - R_s(\tau_1 + \tau_2)) \\ + |\beta + b|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*](\overline{|n|^2} - R_n(\tau_1 + \tau_2)).\end{aligned}\quad (3)$$

Notice that when $\tau_1 = \tau_2 = 0$ (ideal case) then

$$P = \frac{|1 + \beta c|^2 \overline{|s|^2} + |\beta + b|^2 \overline{|n|^2}}{|1 - \alpha\beta|^2}.\quad (4)$$

The same equation was obtained in [2]. There, we showed that optimal weight, $\beta = -b$ and perfect cancelation of the n signal results. However, for τ_1 and $\tau_2 \neq 0$, if we assume $|\alpha| \ll 1$ and $|\beta| \ll 1^1$ then the second term on the left hand side of (3) can be neglected. For the same reason the second term on the right hand side might be neglected, particularly if $\tau_1 + \tau_2$ is sufficiently small, leaving

$$P \simeq \frac{|1 + \beta c|^2 \overline{|s|^2} + |\beta + b|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*](\lambda_n(\tau_1 + \tau_2))}{|1 - \alpha\beta|^2}\quad (5)$$

where

$$\lambda_n(\tau) = \overline{|n|^2} - R_n(\tau)\quad (6)$$

¹Such assumption on $|\alpha|$ and $|\beta|$ is reasonable in cross-pol applications where in $|b|$ and $|c| \ll 1$ and when the weights are close to their optimal weight

and the asterisk denotes complex conjugate.

To find the optimal value of the weight β we equate to zero the derivative² of the output power after processing by the discriminator; P_1 , with respect to β :

$$\frac{\partial P_1}{\partial \beta} = \frac{2(1 - \alpha\beta)[(1 + \beta c)(\alpha + c)^* \delta_{s1} \overline{|s|^2} + (b + \beta)^*(1 + \alpha\beta) \delta_{n1} \overline{|n|^2} - (b + \alpha^* \beta b^*) \delta_{n1} (\lambda_n(\tau_1 + \tau_2))]}{|1 - \alpha\beta|^4} \quad (7)$$

where δ_{s1} and δ_{n1} depict the effect of the discrimination network at output P and the derivative of a real function P with respect to a complex variable β is defined by $dP/d\beta = \partial P/\partial \beta_R + j \partial P/\partial \beta_I$. β_R and β_I are the real and imaginary part of β and $j = \sqrt{-1}$.

From $\partial P_1/\partial \beta = 0$ we get after certain manipulation

$$\begin{aligned} \beta_{op} &= - \frac{(\alpha + c)^* \delta_{s1} \overline{|s|^2} + b(1 + \alpha\beta)^* \delta_{n1} \overline{|n|^2} - b \delta_{n1} (\overline{|n|^2} - R_n(\tau_1 + \tau_2))}{c(\alpha + c)^* \delta_{s1} \overline{|s|^2} + (1 + \alpha\beta)^* \delta_{n1} \overline{|n|^2} - \alpha^* b^* \delta_{n1} (\overline{|n|^2} - R_n(\tau_1 + \tau_2))} \\ &= - \frac{(\alpha + c)^* \delta_{s1} \overline{|s|^2} + b \delta_{n1} (R_n(\tau_1 + \tau_2) + \alpha^* b^* \delta_{n1} \overline{|n|^2})}{c(\alpha + c)^* \delta_{s1} \overline{|s|^2} + \delta_{n1} \overline{|n|^2} + \alpha^* b^* \delta_{n1} R_n(\tau_1 + \tau_2)} \end{aligned} \quad (8)$$

Notice that in contrast to the ideal case we cannot easily conclude that perfect cancelation of the n signal take place.

Using similar steps we obtain, for the output at the other port,

$$Q = \overline{|v_q|^2} \simeq \frac{|1 + \alpha b|^2 \overline{|n|^2} + |c + \alpha|^2 \overline{|s|^2} - 2 \text{Real}[\alpha c] (\lambda_s(\tau_1 + \tau_2))}{|1 - \alpha\beta|^2} \quad (9)$$

where

$$\lambda_s(\tau) = \overline{|s|^2} - R_s(\tau) \quad (10)$$

When $\tau_1 = \tau_2 = 0$ (ideal case) then

$$Q = \frac{|1 + \beta c|^2 \overline{|n|^2} + |c + \alpha|^2 \overline{|s|^2}}{|1 - \alpha\beta|^2} \quad (11)$$

²In obtaining the complex derivative of a real function $A = |d + e\delta|^2$ with respect to the complex variable δ we used the simple general rule $dA/d\delta = 2e^*(d + e\delta)$

From equating to zero the derivative of the output power after processing by the second discriminator, Q_1 we have

$$\begin{aligned}\alpha_{op} &= -\frac{(b + \beta)^* \delta_{n2} \overline{|n|^2} + c(1 + c\beta)^* \delta_{s2} \overline{|s|^2} - c\delta_{s2}(\overline{|s|^2} - R_s(\tau_1 + \tau_2))}{b(b + \beta)^* \delta_{n2} \overline{|n|^2} + (1 + c\beta)^* \delta_{s2} \overline{|s|^2} - c^* \beta^* \delta_{s2}(\overline{|s|^2} - R_s(\tau_1 + \tau_2))} \\ &= -\frac{(b + \beta)^* \delta_{n2} \overline{|n|^2} + c\delta_{s2}(R_s(\tau_1 + \tau_2) + c^* \beta^* \delta_{s2} \overline{|s|^2})}{b(b + \beta)^* \delta_{n2} \overline{|n|^2} + \delta_{s2} \overline{|s|^2} + c^* \beta^* \delta_{s2} R_s(\tau_1 + \tau_2)}\end{aligned}\quad (12)$$

where δ_{n2} and δ_{s2} show the effect of the discriminator at output Q .

We notice in (8) and (12) that α_{op} and β_{op} depends on $\tau_1 + \tau_2$. To examine this dependency, we define

$$\beta_{op} = \beta_o + \epsilon_1 \quad (13)$$

$$\alpha_{op} = \alpha_o + \epsilon_2 \quad (14)$$

where β_o and α_o are the corresponding optimal weight when $\tau_1 = \tau_2 = 0$. It is easily noticeable from (8) and (12) that $\beta_o = -b$ and $\alpha_o = -c$. In which case a perfect cancelation of the interferences at both ports is obtained. Also, it was demonstrated in [2] that these values will be attainable by the system if a search algorithm is used, and appropriate signals' discrimination networks are implemented.

Using (13) and (14) will result in a set of nonlinear equation in ϵ_1 and ϵ_2 , which become linear if we neglect certain second order terms. Using the smallness condition on c and b , and the fact that $\tau_1 + \tau_2$ are quite small relative to signals' correlation times (but not zero), we finally find in Appendix A-1 the approximate solution for these linear equations, namely

$$\epsilon_1 = \frac{b(1 + bc)^* \delta_{n1} - c^*(1 + bc) \delta_{s1} (\lambda_s(\tau_1 + \tau_2) / \lambda_n(\tau_1 + \tau_2))}{(1 - cb)^* (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n(\tau_1 + \tau_2)} \quad (15)$$

$$\epsilon_2 = \frac{c(1 + bc)^* \delta_{s2} - b^*(1 + bc) \delta_{n2} (\lambda_n(\tau_1 + \tau_2) / \lambda_s(\tau_1 + \tau_2))}{(1 - cb)^* (\rho / \delta_{n1}) \overline{|s|^2} / \lambda_s(\tau_1 + \tau_2)} \quad (16)$$

where

$$\rho = \delta_{s2} \delta_{n1} - \delta_{s1} \delta_{n2} \quad (17)$$

It is obviously assumed that $\rho \neq 0$. Notice that in the case where $\tau_1 = \tau_2 = 0$, $\lambda_s(\tau_1 + \tau_2) = \lambda_n(\tau_1 + \tau_2) = 0$ and hence $\epsilon_1 = 0$ and $\epsilon_2 = 0$ is the unique solution for the linearized equation in ϵ_1 and ϵ_2 . This solution corresponds to the perfect cancelation case.

2.2 Effect of System Delays on Outputs' Powers and Ratios

Thus far we have considered the effect of the system delays on the optimal weights. To find the effect of this delay on the power outputs and the outputs power ratio and therefore on the depth of cancelation, we first notice from (5) together with (13) that the power of the s -signal,

$$P_s = \frac{|1 - cb|^2 |1 + c\epsilon_1 / (1 - cb)|^2 |s|^2}{|1 - \alpha\beta|^2}$$

Using (15) and the fact that $|c| \ll 1$ and $|b| \ll 1$ one can show that $c\epsilon_1 / (1 - cb)$ can be neglected leaving

$$P_s \simeq \frac{|1 - bc|^2}{|1 - \alpha\beta|^2} |s|^2 \quad (18)$$

From (5) the power of the n -signal is given by,

$$\begin{aligned} P_n &= |\beta + b|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*] \lambda_n(\tau_1 + \tau_2) \\ &\simeq |\beta + b|^2 \overline{|n|^2} + 2|b|^2 \lambda_n(\tau_1 + \tau_2) \end{aligned} \quad (19)$$

In the last step we applied the results established in Appendix (A-2) namely: $\text{Real}[\beta b^*] \simeq -|b|^2$. Using $\beta + b = \epsilon_1$ and (15) for ϵ_1 into the corresponding terms of (5) we end up with

$$P_n \simeq \frac{1}{|1 - \alpha\beta|^2} \left[\left| \frac{b(1 + bc)^* \delta_{n1} - c^*(1 + bc) \delta_{s1} (\lambda_s / \lambda_n)}{(1 - bc)^*(\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n} \right|^2 \overline{|n|^2} + \frac{2|b|^2}{\overline{|n|^2} / \lambda_n} \overline{|n|^2} \right] \quad (20)$$

Also, in the interest of brevity and when it is clearly understood we sometimes drop the dependence of the λ 's on the delay $\tau_1 + \tau_2$. Notice that the first term of (20) depends inversely on $(\overline{|n|^2} / \lambda_n)^2$ while the second one depends inversely on $(\overline{|n|^2} / \lambda_n)$. From the discussion in Appendix (A-3), we conclude that for almost all values of $\overline{|n|^2} / \lambda_n$, (20) is dominated by the second term of its left hand side.

The s-signal to n-signal power ratio is obtained from (18) and (20) with the aforementioned conclusion,

$$P_s/P_n \simeq \frac{|1 - bc|^2 \overline{|n|^2}/\lambda_n(\tau_1 + \tau_2)}{2|b|^2} \overline{|s|^2}/\overline{|n|^2}. \quad (21)$$

But $|1 - bc|^2 \simeq 1$ and if $\overline{|s|^2}/\overline{|n|^2} \simeq 1$ then

$$(P_s/P_n)_{dB} = (\overline{|n|^2}/\lambda_n(\tau_1 + \tau_2))_{dB} - 3_{dB} - |b|_{dB}^2 \quad (22)$$

That is, the input-output improvement in the s-signal to the n-signal is upper-bounded by $10 \log \left[(1/2) \overline{|n|^2}/\lambda_n(\tau_1 + \tau_2) \right]$.

Because of symmetry we conclude easily, by using (8) for the Q port, that

$$Q_n \simeq \frac{|1 - bc|^2 \overline{|n|^2}}{|1 - \alpha\beta|^2} \quad (23)$$

$$Q_s \simeq \frac{1}{|1 - \alpha\beta|^2} \left[\left| \frac{c(1 + bc)^* \delta_{s2} - b^*(1 + bc) \delta_{n2}(\lambda_n/\lambda_s)}{(1 - cb)^*(\rho/\delta_{n1}) \overline{|s|^2}/\lambda_s} \right|^2 \overline{|s|^2} + \frac{2|c|^2}{\overline{|s|^2}/\lambda_s} \overline{|s|^2} \right] \quad (24)$$

Using the same previous argument, we end up with

$$Q_n/Q_s \simeq \frac{|1 - bc|^2 \overline{|s|^2}/\lambda_s(\tau_1 + \tau_2)}{2|c|^2} \left[\overline{|n|^2}/\overline{|s|^2} \right] \quad (25)$$

Also since $|1 - bc|^2 \simeq 1$ and if $\overline{|n|^2}/\overline{|s|^2} \simeq 1$ then

$$(Q_n/Q_s)_{dB} \simeq (\overline{|s|^2}/\lambda_s(\tau_1 + \tau_2))_{dB} - 3_{dB} - |c|_{dB}^2 \quad (26)$$

That is an input-output improvement of $10 \log(1/2) \overline{|s|^2}/\lambda_s(\tau_1 + \tau_2)$ in the n-signal to the s-signal power ratio. Clearly when $\tau_1 = \tau_2 = 0$ (no delay) we have an infinitely large improvement which corresponds to the ideal perfect cancelation.

2.3 Bandwidth Limitation

Most of signals communicated via dual polarization are of very wideband nature. Therefore it is very important for these applications to estimate the separator bandwidth. To

obtain the system bandwidth limitation, we first notice using Taylor series expansion that.
for example:

$$\lambda_n(\tau_1 + \tau_2) = \overline{|n|^2} - R_n(0) - (\tau_1 + \tau_2) \dot{R}_n(0) - (1/2)(\tau_1 + \tau_2)^2 \ddot{R}_n(0) - \dots \quad (27)$$

where the dots stand for the derivatives with respect to $\tau_1 + \tau_2$. The autocorrelation function is an even function so that $\dot{R}_n(0)$ equal zero. Also $R_n(0) = \overline{|n|^2}$, and (27) reduces to

$$\lambda_n(\tau_1 + \tau_2) \simeq -(1/2)(\tau_1 + \tau_2)^2 \ddot{R}_n(0). \quad (28)$$

Using the Wiener-Khintchine relation [11] we get

$$\begin{aligned} \lambda_n(\tau_1 + \tau_2) &\simeq -(1/2)(\tau_1 + \tau_2)^2 \int_{-\infty}^{+\infty} \frac{d^2}{d\tau^2} [S_n(f) e^{j2\pi f\tau}] \Big|_{\tau=\tau_1+\tau_2} df \\ &= \frac{(\tau_1 + \tau_2)^2}{2} \int_{-W_n}^{W_n} (2\pi f)^2 S_n(f) df \end{aligned} \quad (29)$$

where $S_n(f)$ represents the two-sided spectral density of the n signal process and W_n is the bandwidth (without loss of generality the signals are assumed at baseband). Using (29), $\lambda_n(\tau_1 + \tau_2)$ can be upper-bounded as follows:

$$\begin{aligned} \lambda_n(\tau_1 + \tau_2) &\leq \frac{(\tau_1 + \tau_2)^2}{2} (2\pi W_n)^2 \int_{-W_n}^{W_n} S_n(f) df \\ &= \frac{(\tau_1 + \tau_2)^2}{2} (2\pi W_n)^2 \overline{|n|^2}. \end{aligned} \quad (30)$$

On the other hand, if the signal spectral density is approximately uniform in $(-W_n, W_n)$ then, $S_n(f) = \overline{|n|^2}/2W_n$. Substituting this value of $S_n(f)$ in (29), we obtain

$$\begin{aligned} \lambda_n(\tau_1 + \tau_2) &\simeq \frac{(\tau_1 + \tau_2)^2}{2} \frac{\overline{|n|^2}}{2W_n} \int_{-W_n}^{W_n} (2\pi f)^2 df \\ &= \frac{(\tau_1 + \tau_2)^2}{6} (2\pi W_n)^2 \overline{|n|^2}. \end{aligned} \quad (31)$$

Substituting (30) in (21) we get

$$P_s/P_n \geq \frac{|1 - bc|^2}{|b|^2} \left[\frac{1}{2\pi W_n(\tau_1 + \tau_2)} \right]^2 (\overline{|s|^2}/\overline{|n|^2}), \quad (32)$$

and for the uniform spectral density signals we obtain

$$(P_s/P_n)_u \simeq \frac{3|1 - bc|^2}{|b|^2} \left[\frac{1}{2\pi W_n(\tau_1 + \tau_2)} \right]^2 (\overline{|s|^2}/\overline{|n|^2}). \quad (33)$$

Clearly in order to minimize the effect of the systems' delays on the performance, we must require $\tau_1 + \tau_2 \ll \frac{1}{2\pi W_n}$.

Similar argument leads, for the other output, to

$$Q_n/Q_s \geq \frac{|1 - bc|^2}{|c|^2} \left[\frac{1}{2\pi W_s(\tau_1 + \tau_2)} \right]^2 (\overline{|n|^2}/\overline{|s|^2}) \quad (34)$$

for a signal $s(t)$ having any form of spectral density limited to a bandwidth W_s . If $s(t)$ has approximately uniform spectral density then

$$(Q_n/Q_s)_u \simeq \frac{3|1 - bc|^2}{|c|^2} \left[\frac{1}{2\pi W_s(\tau_1 + \tau_2)} \right]^2 (\overline{|n|^2}/\overline{|s|^2}). \quad (35)$$

Equations (32) and (34) might be written as

$$\gamma_p \triangleq \frac{P_s/P_n}{\overline{|s|^2}/\overline{|b|^2}\overline{|n|^2}} \geq \frac{|1 - bc|^2}{(2\pi W_n(\tau_1 + \tau_2))^2} \quad (36)$$

$$\gamma_q \triangleq \frac{Q_n/Q_s}{\overline{|n|^2}/\overline{|c|^2}\overline{|s|^2}} \geq \frac{|1 - bc|^2}{(2\pi W_s(\tau_1 + \tau_2))^2} \quad (37)$$

where γ_p and γ_q represent the input-output improvement in signal-to-interference ratio at P port and Q port respectively. For signals having uniform spectral density we have

$$\gamma_{pu} = \frac{3|1 - bc|^2}{[2\pi W_n(\tau_1 + \tau_2)]^2} \quad (38)$$

$$\gamma_{qu} = \frac{3|1 - bc|^2}{[2\pi W_s(\tau_1 + \tau_2)]^2}. \quad (39)$$

As an example, let $|1 - bc| \simeq 1$, $W_n \simeq W_s = 100 \text{ Mhz}$. Then, for an improvement of at least 20dB (corresponding to the 40dB signal-to-interference ratio at the output when $|b| \simeq |c| = 0.1$), $\tau_1 + \tau_2$ must be less than 0.16 nanosecond.

When the signals' spectral densities are approximately uniform, then $\tau_1 + \tau_2$ must be less than 0.27 nanosecond. These values of delay are obviously difficult to accomplish. On the other hand, if $\tau_1 + \tau_2 \simeq 12$ nanoseconds (values which are practically reasonable) then W_n would have to be less than 1.3 MHz or about 2.25 MHz if the signals' spectral densities are known to be approximately uniform.

III. THE POWER-CORRELATION SCHEME

Figure 2 depicts the separator structure termed power-correlation. That is, the corresponding separator weights are chosen to minimize the power at one output and the correlation between the two outputs. This section will follow the same steps performed for the power-power scheme in the previous section.

3.1 Effect of System Delays on the Optimal Weights

From Figure 2 we can write

$$\begin{aligned} v_p(t) = & s(t - \tau_1) + bn(t - \tau_1) + \beta cs(t - \tau_2 - \tau_1/2) + \beta n(t - \tau_2 - \tau_1/2) \\ & + \alpha\beta s(t - 2\tau_2) + \alpha\beta bn(t - 2\tau_2). \end{aligned} \quad (40)$$

Using the definition of P we obtain after rearranging terms,

$$\begin{aligned} P = \overline{|v_p(t)|^2} = & |1 + \beta(\alpha + c)|^2 \overline{|s|^2} - 2\text{Real}[\beta c](\overline{|s|^2} - R_s(\tau_1/2 - \tau_2)) \\ & - 2\text{Real}[\alpha\beta](\overline{|s|^2} - R_s(\tau_1 - 2\tau_2)) - 2|\beta|^2 \text{Real}[c\alpha^*](\overline{|s|^2} - R_s(\tau_1/2 - \tau_2)) \\ & + |b + \beta(1 + \alpha b)|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*](\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)) \\ & - 2|b|^2 \text{Real}[\alpha\beta](\overline{|n|^2} - R_n(\tau_1 - 2\tau_2)) - 2|\beta|^2 \text{Real}[\alpha\beta](\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)). \end{aligned} \quad (41)$$

For $\tau_1 = \tau_2 = 0$ (ideal case) we get

$$P = |1 + \beta(\alpha + c)|^2 \overline{|s|^2} + |b + \beta(1 + \alpha b)|^2 \overline{|n|^2} \quad (42)$$

in agreement with the result of [2]. However, if we assume that $|b|$ and $|c|$ are much smaller than unity, as is the case for signal mixtures obtained from depolarization of dually polarized signals, and $|\alpha|$ and $|\beta|$ are constrained to a value which is very small compared to unity, the second, third, and fourth terms of the left hand side of (41) can be neglected in comparison to the first term, particularly if $\tau_1 - 2\tau_2$ is sufficiently small, leaving

$$P = |1 + \beta(\alpha + c)|^2 |s|^2 + |b + \beta(1 + \alpha b)|^2 |n|^2 - 2\text{Real}[\beta b^*](|n|^2 - R_n(\tau_1/2 - \tau_2)) - 2|b|^2 \text{Real}[\alpha\beta](|n|^2 - R_n(\tau_1 - 2\tau_2)) - 2|\beta|^2 \text{Real}[\alpha b](|n|^2 - R_n(\tau_1/2 - \tau_2)). \quad (43)$$

To find the optimal value of the weight β we equate to zero the derivative of the power of this output after processing by the discriminator; P_1 , with respect to β

$$\begin{aligned} \frac{\partial P_1}{\partial \beta} = & (\alpha + c)^* \delta_{s1} |s|^2 + \beta |\alpha + c|^2 \delta_{s1} |s|^2 + \beta |1 + \alpha b|^2 \delta_{n1} |n|^2 + b(1 + \alpha b)^* \delta_{n1} |n|^2 \\ & - 2b \delta_{n1} (|n|^2 - R_n(\tau_1/2 - \tau_2)) - 2|b|^2 \alpha^* \delta_{n1} (|n|^2 - R_n(\tau_1 - 2\tau_2)) \\ & - 4\beta \text{Real}[\alpha b] \delta_{n1} (|n|^2 - R_n(\tau_1/2 - \tau_2)) \end{aligned} \quad (44)$$

where δ_{s1} and δ_{n1} are the effects of the discriminator on signal s and signal n , respectively. From $\partial P_1 / \partial \beta = 0$ we obtain

$$\begin{aligned} \beta_{op} = & \frac{b(1 + \alpha b)^* \delta_{n1} |n|^2 + (\alpha + c)^* \delta_{s1} |s|^2 - 2b \delta_{n1} (\lambda_n(\tau_1/2 - \tau_2)) - 2|b|^2 \alpha^* \delta_{n1} (\lambda_n(\tau_1 - 2\tau_2))}{|1 + \alpha b|^2 \delta_{n1} |n|^2 + |\alpha + c|^2 \delta_{s1} |s|^2 - 4\text{Real}[\alpha b] \delta_{n1} (\lambda_n(\tau_1/2 - \tau_2))} \\ = & \frac{b(1 + \alpha b)^* \delta_{n1} |n|^2 + (\alpha + c)^* \delta_{s1} |s|^2 - 2b(1 + 2\alpha b)^* \delta_{n1} (\lambda_n(\tau_1/2 - \tau_2))}{|1 + \alpha b|^2 \delta_{n1} |n|^2 + |\alpha + c|^2 \delta_{s1} |s|^2 - 4\text{Real}[\alpha b] \delta_{n1} (\lambda_n(\tau_1/2 - \tau_2))}. \end{aligned} \quad (45)$$

where we approximated $|n|^2 - R_n(\tau_1 - 2\tau_2)$ by $2(|n|^2 - R_n(\tau_1/2 - \tau_2))$.

For the Q port we notice that

$$v_q(t) = \alpha s(t - \tau_2 - \tau_1/2) + \alpha b n(t - \tau_2 - \tau_1/2) + c s(t - \tau_1) + n(t - \tau_1) \quad (46)$$

and

$$Q = \overline{|v_q(t)|^2} = |\alpha + c|^2 |s|^2 - 2\text{Real}[\alpha c^*] \lambda_s (\tau_1/2 - \tau_2) \\ + |1 + \alpha b|^2 \overline{|n|^2} - 2\text{Real}[\alpha b] \lambda_n (\tau_1/2 - \tau_2). \quad (47)$$

When $\tau_1 = \tau_2 = 0$ (ideal case) then

$$Q = |\alpha + c|^2 |s|^2 + |1 + \alpha b|^2 \overline{|n|^2} \quad (48)$$

is in agreement with the results in [2]. However, under the regular smallness condition assumed previously, we can neglect the last term in (47), particularly if $\tau_1/2 - \tau_2$ is sufficiently small, leaving

$$Q = |\alpha + c|^2 |s|^2 - 2\text{Real}[\alpha c^*] \lambda_s (\tau_1/2 - \tau_2) + |1 + \alpha b|^2 \overline{|n|^2} \quad (49)$$

Also, from Figure 2

$$Q_2 = \overline{|v_{q1}(t)v_p^*(t)|^2} = |A|^2 \quad (50)$$

v_{q1} contains the effect of the discriminator at output 2.

Where using (40) and (46) we find in Appendix A-4, (P-12):

$$A = (\alpha + c)\delta_{s2} \overline{|s|^2} - \alpha\delta_{s2} \lambda_s (\tau_1/2 - \tau_2) + \beta^* |\alpha + c|^2 \delta_{s2} R_s (\tau_1/2 - \tau_2) \\ + 2j\beta^* I_m[\alpha c^*] \delta_{s2} \lambda_s (\tau_1/2 - \tau_2) + \beta^* |1 + \alpha b|^2 \delta_{n2} R_n (\tau_1/2 - \tau_2) \\ + 2j\beta^* I_m[\alpha b] \delta_{n2} \lambda_n (\tau_1/2 - \tau_2) + b^* (1 + \alpha b) \delta_{n2} \overline{|n|^2} \\ - \alpha |b|^2 \delta_{n2} \lambda_n (\tau_1/2 - \tau_2). \quad (51)$$

Close to the optimal solution, $\alpha = -c$. it is plausible to assume that $\text{Real}[\beta|\alpha + c|^2] \ll \text{Real}[(\alpha + c)]$ and $\text{Imag}[\beta|\alpha + c|^2] \ll \text{Imag}[(\alpha + c)]$. Therefore we can neglect the third term of (51) in comparison to the first. Also since $\tau_1/2 - \tau_2$ is assumed to be very small. then one can neglect the last term of (51) in comparison to its preceding term. leaving

$$A = (\alpha + c)\delta_{s2} \overline{|s|^2} - \alpha\delta_{s2} \lambda_s (\tau_1/2 - \tau_2) + 2j\beta^* I_m \delta_{s2} [\alpha c^*] \lambda_s (\tau_1/2 - \tau_2) \\ + \beta^* |1 + \alpha b|^2 \delta_{n2} R_n (\tau_1/2 - \tau_2) + 2j\beta^* I_m [\alpha b] \delta_{n2} \lambda_n (\tau_1 - 2\tau_2) \\ + b^* (1 + \alpha b) \delta_{n2} \overline{|n|^2}. \quad (52)$$

Equating $A = 0$ gives us the optimal value α_{op} ³

$$\begin{aligned}\alpha_{op} &\simeq c - (1 + \alpha b)(b + \beta(1 + \alpha b))^* \delta_{n2} \overline{|n|^2} / \delta_{s2} \overline{|s|^2} \\ &\quad - 2j\beta^* I_m[\alpha c^*] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2) / \delta_{s2} \overline{|s|^2} \\ &\quad - 2j\beta^* I_m[\alpha b] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2) / \delta_{s2} \overline{|s|^2}\end{aligned}\quad (53)$$

where we used $\overline{|n|^2} \simeq R_n(\tau_1/2 - \tau_2)$ and $\overline{|s|^2} \simeq R_s(\tau_1/2 - \tau_2)$.

We notice from (45) and (53) that β_{op} and α_{op} depend on $\tau_1/2 - \tau_2$. To examine this dependency, we define,

$$\beta_{op} = \frac{\beta_0 + \epsilon_1}{1 + \alpha b} \quad (54)$$

$$\alpha_{op} = \alpha_0 + \epsilon_2 \quad (55)$$

where β_0 and α_0 are the corresponding optimal weights when $\tau_1/2 - \tau_2 = 0$. It can be shown, using (45) and (53) that in this case $\beta_0 = -b/(1 + \alpha b)$ and $\alpha_0 = -c$, and a perfect cancelation of interferences is obtained.

As in the case of the power-power canceler, by using (54) and (55) we get a set of nonlinear equations which can be linearized by neglecting the second order terms in ϵ_1 and ϵ_2 . In Appendix (A-5), we present the arguments and the analytic steps used to derive this linear set of equations. The solution of this linearized set is further simplified by using the regular smallness condition on $|b|$ and $|c|$ to obtain (see (P-20) and (P-21))

$$\epsilon_1 = \frac{2b(1 - bc)^* \delta_{s2} - 2j I_m[cb] \delta_{s1} (\delta_{n2} / \delta_{n1})}{(1 - bc)^* (\rho / \delta_{n1}) |n|^2 / \lambda_n(\tau_1/2 - \tau_2)} \quad (56)$$

$$\epsilon_2 = -\frac{2j I_m[cb] \delta_{n1} + 2b^*(1 - bc) \delta_{n1}}{(1 - bc)^* (\rho / \delta_{n2}) |s|^2 / \lambda_n(\tau_1/2 - \tau_2)} \quad (57)$$

where δ_{ni} and δ_{si} , $i = 1, 2$ represent the effect of the discriminator networks (ρ is assumed non-zero). Notice that if $\tau_1/2 = \tau_2$, $\lambda_n(\tau_1/2 - \tau_2) = 0$, then $\epsilon_1 = 0$ and $\epsilon_2 = 0$ is the unique

³In fact there might be other conditions other than $A = 0$ which lead to minimizing Q_2 of (50). However if the delay are sufficiently small then applying the arguments in [2] we showed that if certain smallness conditions apply then, $A = 0$ is the only condition that implies minimum Q_2 .

solution. This corresponds to the ideal, perfect cancelation case.

3.2 Effect of System Delays on Outputs' Powers and Ratios

To find the effect of the system delay on the power outputs and the outputs ratio and hence on the depth of cancelation, we first notice using (43) together with (54) and (55), that

$$P_s = \left| \frac{1 - bc + \epsilon_1 \epsilon_2}{1 - bc + b\epsilon_2} \right|^2 \overline{|s|^2} \simeq \overline{|s|^2} \quad (58)$$

and

$$\begin{aligned} P_n = & |b + \beta(1 + \alpha b)|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*](\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)) \\ & - 2|b|^2 \text{Real}[\alpha \beta](\overline{|n|^2} - R_n(\tau_1 - 2\tau_2)) - 2|\beta|^2 \text{Real}[\alpha b](\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)). \end{aligned}$$

Applying (P-22) of Appendix (A-6) we get

$$\begin{aligned} P_n = & |b + \beta(1 + \alpha b)|^2 \overline{|n|^2} + 2|\beta|^2(\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)) \\ & - 2\text{Real}[\beta^* \epsilon_1](\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)) - 2|b|^2 \text{Real}[\alpha \beta](\overline{|n|^2} - R_n(\tau_1 - 2\tau_2)). \end{aligned} \quad (59)$$

It is clear that because $|\beta| \approx |b|$ then $|b|^2 \text{Real}[\alpha \beta] \leq |b| |\alpha| |\beta|^2$. Therefore if $|\epsilon_1| \ll |\beta|$ then $\text{Real}[\beta^* \epsilon_1] \ll |\beta|^2$ the third and fourth terms of (59) might be neglected in comparison to the second term leaving

$$P_n = |b + \beta(1 + \alpha b)|^2 \overline{|n|^2} + 2|\beta|^2 \lambda_n(\tau_1/2 - \tau_2). \quad (60)$$

Using $b + \beta(1 + \alpha b) = \epsilon_1$, and (56) we end up with

$$P_n = \left| \frac{2b(1 - bc)^* \delta_{s2} - 2j I_m[cb](\delta_{n2}/\delta_{n1})}{|1 - bc|^2 (\rho/\delta_{n1}) \overline{|n|^2} / \lambda'_n} \right|^2 \overline{|n|^2} + \frac{2|b|^2}{\overline{|n|^2} / \lambda'_n} \overline{|n|^2} \quad (61)$$

where we defined

$$\lambda'_n = \lambda_n(\tau_1/2 - \tau_2) = \overline{|n|^2} - R_n(\tau_1/2 - \tau_2)$$

and used $|\beta|^2 \simeq |b|^2$.

In the interest of brevity we sometimes do as we did in (61), drop the dependence of the $\lambda_n(\tau_1/2 - \tau_2)$ and $\lambda_s(\tau_1/2 - \tau_2)$ on the delay $\tau_1/2 - \tau_2$. However, to distinguish these λ 's from the previously defined ones, we use primes. Notice that the first term of (61) depends inversely on $[\overline{|n|^2}/\lambda'_n]^2$ while the second depends inversely on $\overline{|n|^2}/\lambda'_n$. But λ'_n depends on $\tau_1/2 - \tau_2$ so that $\overline{|n|^2}/\lambda'_n$ can easily be made very large. From the discussion in Appendix (A-7) we conclude that for $\overline{|n|^2}/\lambda'_n$ moderately greater than unity (> 1.41), (61) is dominated by the second term.

Therefore s -signal to the n -signal power ratio is given by

$$P_s/P_n = \frac{\overline{|n|^2}/\lambda_n(\tau_1/2 - \tau_2)}{2|b|^2} (\overline{|s|^2}/\overline{|n|^2}) \quad (62)$$

When $\overline{|s|^2}/\overline{|n|^2} \simeq 1$, then

$$(P_s/P_n)_{dB} = (\overline{|n|^2}/\lambda_n(\tau_1/2 - \tau_2))_{dB} - 3_{dB} - |b|^2_{dB}. \quad (63)$$

That is, the input-output improvement in s -signal to the n -signal is upperbounded by $10 \log(1/2) \overline{|n|^2}/\lambda_n(\tau_1/2 - \tau_2)$.

For the other port we have from (49) together with (55)

$$Q_n \simeq |1 - bc|^2 \overline{|n|^2} \quad (64)$$

and

$$\begin{aligned} Q_s &= |\alpha + c|^2 \overline{|s|^2} - 2 \text{Real}[\alpha c^*] \lambda_s(\tau_1/2 - \tau_2) \\ &\simeq |\alpha + c|^2 \overline{|s|^2} + 2|c|^2 \lambda_s(\tau_1/2 - \tau_2) \end{aligned} \quad (65)$$

where we used $\text{Real}[\alpha c^*] = -|c|^2 + \text{Real}[c^* \epsilon_2] \simeq -|c|^2$ (see similar argument used in obtaining (19) in Appendix (A-2)). But by (55), $\alpha + c = \epsilon_2$ and together with (57) we write

$$Q_s = \left| \frac{2j I_m[cb] \delta_{n1} + 2b^*(1 - bc) \delta_{n1}}{(1 - bc)(\rho/\delta_{n2}) \overline{|s|^2}/\lambda'_n} \right|^2 \overline{|s|^2} + \frac{2|c|^2}{\overline{|s|^2}/\lambda'_s} \overline{|s|^2} \quad (66)$$

where we defined

$$\lambda'_s \triangleq \lambda_s(\tau_1/2 - \tau_2) = \overline{|s|^2} - R_s(\tau_1/2 - \tau_2). \quad (67)$$

From the discussion in Appendix (A-8) we conclude that for almost all $\overline{|s|^2}/\lambda'_s$, (66) is dominated by its second term. Therefore the n -signal to the s -signal power ratio is given by

$$Q_n/Q_s = \frac{\overline{|s|^2}/\lambda_s(\tau_1/2 - \tau_2)}{2|c|^2} (\overline{|n|^2}/\overline{|s|^2}). \quad (68)$$

When $\overline{|n|^2}/\overline{|s|^2} \simeq 1$, then

$$(Q_n/Q_s)_{dB} = (\overline{|s|^2}/\lambda_s(\tau_1/2 - \tau_2))_{dB} - 3_{dB} - |c|_{dB}^2. \quad (69)$$

That is, the input-output improvement in the n -signal to the s -signal is upperbounded by $10 \log(1/2)\overline{|s|^2}/\lambda_s(\tau_1/2 - \tau_2)$.

3.3 Bandwidth Limitation

By comparing (22) and (26) to (63) and (69) we conclude that despite the difference in hardware implementation between the power-power and power-correlation scheme, a similar relation governs the effect of the delays on the output power ratio. Nevertheless, while the first scheme depends on $\lambda(\tau_1 + \tau_2)$ the second depends on $\lambda(\tau_1/2 - \tau_2)$. Similar calculations which lead to (32) and (34) gives, for this case,

$$P_s/P_n \geq \frac{1}{|b|^2} \left[\frac{1}{2\pi W_n(\tau_1/2 - \tau_2)} \right]^2 (\overline{|s|^2}/\overline{|n|^2}) \quad (70)$$

$$Q_n/Q_s \geq \frac{1}{|c|^2} \left[\frac{1}{2\pi W_s(\tau_1/2 - \tau_2)} \right]^2 (\overline{|n|^2}/\overline{|s|^2}) \quad (71)$$

where W_n and W_s are defined as before. Equation (70) and (71) represent a lower bound on the performance for any kind of signals' spectra. For the uniform signals' spectra we get (see (33) and (35)),

$$(P_s/P_n)_u \simeq \frac{3}{|b|^2} \left[\frac{1}{2\pi W_n(\tau_1/2 - \tau_2)} \right]^2 (\overline{|s|^2}/\overline{|n|^2}) \quad (72)$$

$$(Q_n/Q_s)_u \simeq \frac{3}{|c|^2} \left[\frac{1}{2\pi W_s(\tau_1/2 - \tau_2)} \right]^2 (\overline{|n|^2}/\overline{|s|^2}). \quad (73)$$

As an example, let $W_n = W_s = 100\text{MHz}$. Then for an improvement of at least 20dB, $\tau_1/2 - \tau_2$ must be less than 0.16 nanoseconds. In practice, we can take $\tau_1 \simeq \tau_2 = 6$ nanosecond. If we can adjust the difference to 5% of each delay (0.03 nanosecond) then we can obtain at least a 20dB improvement for as much as a 500MHz bandwidth, or better than a 34dB improvement (corresponding to a 54 dB signal to interface ratio at the output when $|b| \simeq |c| = 0.1$) a for 100MHz bandwidth.

IV. THE CORRELATION-CORRELATION SCHEME

Figure 3 presents the separator structure termed correlation-correlation. That is, one separator weight is selected to minimize the correlation between one output after it is processed through a discriminator network and the other output and vice versa for controlling the other weight. In this section we will follow the same steps used for the other structures.

4.1 Effect of System Delays on the Optimal Weights

From Figure 3, we can write,

$$v_p(t) = s(t - \tau_1) + bn(t - \tau_1) + \beta cs(t - \tau_2 - \tau_2/2) + \beta n(t - \tau_2 - \tau_1/2). \quad (74)$$

Therefore,

$$\begin{aligned} P = \overline{|v_p(t)|^2} &= |b + \beta|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*] \lambda_n(\tau_1/2 - \tau_2) \\ &+ |1 + \beta c|^2 \overline{|s|^2} - 2\text{Real}[\beta c] \lambda_s(\tau_1/2 - \tau_2) \end{aligned} \quad (75)$$

where $\lambda_s(\tau) = \overline{|s|^2} - R_s(\tau)$ and $\lambda_n(\tau) = \overline{|n|^2} - R_n(\tau)$.

When $\tau_1 = \tau_2 = 0$ (ideal case) then

$$P = |b + \beta|^2 \overline{|n|^2} + |1 + \beta c|^2 \overline{|s|^2} \quad (76)$$

in agreement with the results in [2].

Also from Figure 3, we have

$$v_q(t) = n(t - \tau_1) + cs(t - \tau_1) + \alpha bn(t - \tau_2 - \tau_1/2) + \alpha s(t - \tau_2 - \tau_1/2) \quad (77)$$

so that

$$Q = |\overline{v_q(t)}|^2 = |c + \alpha|^2 \overline{|s|^2} + 2\text{Real}[\alpha c^*] \lambda_s(\tau_1/2 - \tau_2) + |1 + \alpha b|^2 \overline{|n|^2} - 2\text{Real}[\alpha b] \lambda_n(\tau_1/2 - \tau_2). \quad (78)$$

The control loops operate on the correlators' outputs P_1 and Q_2 respectively,

as shown in Figure 3,

$$P_1 = \overline{|v_{p1}(t)v_q^*(t)|^2} \triangleq |A_1|^2 \quad (79)$$

where one can easily show that

$$A_1 = (c + \alpha)^*(1 + c\beta)\delta_{s1}\overline{|s|^2} - (\alpha^* + \beta|c|^2)\lambda_s(\tau_1/2 - \tau_2)\delta_{s1} + (b + \beta)(1 + \alpha b)^*\delta_{n1}\overline{|n|^2} - (\beta + \alpha^*|b|^2)\lambda_n(\tau_1/2 - \tau_2)\delta_{n1} \quad (80)$$

with $\delta_{n1} = \overline{|n_1 n^*|}/\overline{|n|^2}$ and $\delta_{s1} = \overline{|s_1 s^*|}/\overline{|s|^2}$. n_1 and s_1 reflect the effect of the discriminator on output 1.

Similarly

$$Q_2 = \overline{|v_{q2}(t)v_p^*(t)|^2} = |A_2|^2 \quad (81)$$

with

$$A_2 = (b + \beta)^*(1 + \alpha b)\delta_{n2}\overline{|n|^2} - (\beta^* + \alpha|b|^2)\lambda_n(\tau_1/2 - \tau_2)\delta_{n2} + (c + \alpha)(1 + c\beta)^*\delta_{s2}\overline{|s|^2} - (\alpha + \beta^*|c|^2)\lambda_s(\tau_1/2 - \tau_2)\delta_{s2} \quad (82)$$

and $\delta_{s2} = \overline{|s_2 s^*|}/\overline{|s|^2}$ and $\delta_{n2} = \overline{|n_2 n^*|}/\overline{|n|^2}$. n_2 and s_2 reflect the effect of the discriminator on output 2.

To find β_{op} we must equate $\partial P_1/\partial\beta_R$ and $\partial P_1/\partial\beta_I$, simultaneously to zero. Again following the same arguments used in [2], and provided the delays are sufficiently small, one can show that $\partial P_1/\partial\beta_R = 0$ and $\partial P_1/\partial\beta_I = 0$ if and only if $A_1 = 0$. Similarly α_{op} is obtained by equating A_2 to zero. From $A_1 = 0$ we get

$$\beta_{op} = -\frac{b(1 + \alpha b)^* \delta_{n1} \overline{|n|^2} - \alpha^* |b|^2 \lambda_n (\tau_1/2 - \tau_2) \delta_{n1} + (c + \alpha)^* \delta_{s1} \overline{|s|^2} - \alpha^* \lambda_s (\tau_1/2 - \tau_2) \delta_{s1}}{(1 + \alpha b)^* \delta_{n1} \overline{|n|^2} - \lambda_n (\tau_1/2 - \tau_2) \delta_{n1} + c(c + \alpha)^* \delta_{s1} \overline{|s|^2} - |c|^2 \lambda_s (\tau_1/2 - \tau_2) \delta_{s1}} \quad (83)$$

From $A_2 = 0$ we obtain

$$\alpha_{op} = -\frac{c(1 + c\beta)^* \delta_{s2} \overline{|s|^2} - \beta^* |c|^2 \lambda_s (\tau_1/2 - \tau_2) \delta_{s2} + (b + \beta)^* \delta_{n2} \overline{|n|^2} - \beta^* \lambda_n (\tau_1/2 - \tau_2) \delta_{n2}}{(1 + c\beta)^* \delta_{s2} \overline{|s|^2} - \lambda_s (\tau_1/2 - \tau_2) \delta_{s2} + b(b + \beta)^* \delta_{n2} \overline{|n|^2} - |b|^2 \lambda_n (\tau_1/2 - \tau_2) \delta_{n2}} \quad (84)$$

To find the effect of delay $\tau_1/2 - \tau_2$ on the optimal weights we define

$$\beta_{op} = \beta_0 + \epsilon_1 \quad (85)$$

$$\alpha_{op} = \alpha_0 + \epsilon_2 \quad (86)$$

where $\beta_0 = -b$ and $\alpha_0 = -c$ are the optimal weights when $\tau_1/2 - \tau_2 = 0$.

Again, by using (85) and (86) in (83) and (84) we get two nonlinear equations in ϵ_1 and ϵ_2 . These can be linearized by neglecting second order terms in ϵ_1 and ϵ_2 . The resulted set of linear equation is then solved (see Appendix A-9) to finally obtain

$$\epsilon_1 = \frac{-b(1 + bc)^*}{(1 - cb)^* \overline{|n|^2} / \lambda_n (\tau_1/2 - \tau_2)} \quad (87)$$

$$\epsilon_2 = \frac{-c(1 + bc)^*}{(1 - cb)^* \overline{|s|^2} / \lambda_s (\tau_1/2 - \tau_2)} \quad (88)$$

Notice that if $\tau_1 = \tau_2$ then ϵ_1 and ϵ_2 are zero and we have the ideal perfect cancelation case.

4.2 Effect of System Delays on Outputs' Powers and Ratios

To estimate the effect of the system delay on the power outputs and the outputs ratio we first notice, using (75), that

$$P_s \simeq |1 + \beta c|^2 \overline{|s|^2} \quad (89)$$

where the last term in (75) was neglected due to the fact that $|c|$ is assumed less than unity. $|\beta|$ is restricted to less than unity, and, particularly, for $\tau_1/2 - \tau_2$ small, $\lambda_s(\tau_1/2 - \tau_2)$ is small in comparison to $\overline{|s|^2}$. Using (85) with $\beta_0 = -b$, (89) becomes

$$\begin{aligned} P_s &= |1 - bc + c\epsilon_1|^2 \overline{|s|^2} \\ &\simeq |1 - bc|^2 \overline{|s|^2}. \end{aligned} \quad (90)$$

Also from (75) by using (P-9),

$$\begin{aligned} P_n &= |b + \beta|^2 \overline{|n|^2} - 2\text{Real}[\beta b^*] \lambda_n(\tau_1/2 - \tau_2) \\ &\simeq |b + \beta|^2 \overline{|n|^2} + 2|b|^2 (\lambda_n(\tau_1/2 - \tau_2)). \end{aligned} \quad (91)$$

But $b + \beta = \epsilon_1$ and together with (87)

$$P_n = \left| \frac{b(1 + bc)^*}{(1 - bc)^* \overline{|n|^2} / \lambda_n(\tau_1/2 - \tau_2)} \right|^2 \overline{|n|^2} + \frac{2|b|^2}{\overline{|n|^2} / \lambda_n(\tau_1/2 - \tau_2)} \overline{|n|^2}. \quad (92)$$

It is easy to notice that if $|bc| \ll 1$ then (92) is dominated by its second term. Therefore, the s -signal to the n -signal ratio is given by

$$P_s/P_n \simeq \frac{|1 - bc|^2 \overline{|n|^2} / \lambda_n(\tau_1/2 - \tau_2)}{2|b|^2} (\overline{|s|^2} / \overline{|n|^2}). \quad (93)$$

Notice that this is exactly (21) obtained for the power-power scheme, except for the dependence of λ_n on $\tau_1/2 - \tau_2$ instead of on $\tau_1 + \tau_2$. Comparing it with (62) obtained for the power-correlation, we notice the small difference due to the inclusion of $|1 - bc|^2 \simeq 1$ in (93). Both in (62) and (93), λ_n depends on $\tau_1/2 - \tau_2$. Because of the symmetry we get by using (78) and implying similar steps and arguments;

$$Q_n \simeq |1 - bc|^2 \overline{|n|^2} \quad (94)$$

$$Q_s \simeq \left| \frac{c(1 + bc)^*}{(1 - bc)^* \overline{|s|^2} / \lambda_s(\tau_1/2 - \tau_2)} \right|^2 \overline{|s|^2} + \frac{2|c|^2}{\overline{|s|^2} / \lambda_s(\tau_1/2 - \tau_2)} \overline{|s|^2} \quad (95)$$

and

$$Q_n/Q_s \simeq \frac{|1 - bc|^2 \overline{|s|^2} / \lambda_s(\tau_1/2 - \tau_2)}{2|c|^2} (\overline{|n|^2} / \overline{|s|^2}). \quad (96)$$

Equation (96) is exactly the same as (25) obtained for the power-power scheme except for the dependence of λ_n . It differs slightly from (68) obtained for the power-correlation scheme.

3.4 Bandwidth Limitation

Disregarding the small differences between (93) or (96) and (62) or (68) respectively ($|1-bc| \simeq 1$), then (70) and (71) are lower bound limits on the output signal-to-interference ratio. In the case when the signals' spectra can be approximated by uniform spectra limited to bandwidths, W_n and W_s , for the n -signal and s -signal respectively, then (72) and (73) obtained for the power-correlation scheme are also in effect for the correlation-correlation case. Therefore the numerical examples used there are also valid for the correlation-correlation scheme.

V. SUMMARY OF RESULTS AND CONCLUSION

In this section the different results obtained in this paper are summarized. These are as follows:

1. The optimal weights β_{op} and α_{op} of the three schemes with delays τ_1 and τ_2 are shown in Table 1A.
2. The corresponding zero delay optimal weights β_0 and α_0 are given in Table 1B.
3. The effect of circuits delays on the optimal weights (Table 2) are represented through $\epsilon_1 = \beta_{op} - \beta_0$ and $\epsilon_2 = \alpha_{op} - \alpha_0$, except for the power-correlation scheme where $\epsilon_1 = (1 + \alpha b)(\beta_{op} - \beta_0)$.
4. The effect of circuits delays on the output power ratios are shown in Table 3.
5. Input-output improvement in signal-to interference power ratios are shown in Table 4.
6. Output power ratios as a function of signals' bandwidths and circuits delay are shown

in Table 5. For the case when the signals' spectral densities are assumed uniform, then the output power ratios are 5dB above the lower bound values shown in Table 5.

The effects of nonzero hardware delays on the performance of the three different cross-pol cancelers were examined in this paper. Firstly, the effects of these delays on the optimal weights were considered and the variations from the ideal zero delay case were obtained. Secondly, the output powers and power ratios at each of the canceler ports were exhibited as a function of the actual circuit delays. Depending on the different configuration, different forms of dependence resulted. In particular, we notice that while in the power-power scheme the performance depends inversely on $\tau_1 + \tau_2$, where τ_1 and τ_2 are delays of the summation and cancelation paths respectively, the performances of the power-correlation and correlation-correlation schemes depend inversely on $\tau_1/2 - \tau_2$. Finally, using the relation between the signals' autocorrelation and spectral density we could estimate the lower bound on the depth of cancelation as a function of both the signals' bandwidths and system delays and hence establish bandwidth limitation due to system delays. For the case when the signals' spectra are approximately uniform and limited to W_n and W_s , respectively for the n signal and the s signal, we show the output signal-to-interference ratio as a function of these bandwidths, the system delays, the input power ratio and the input cross-pol ratios. Using the power-power scheme, the output signal-to-interference ratio depends inversely on $\tau_1 + \tau_2$, for the other two schemes it depends inversely on $\tau_1/2 - \tau_2$. Since this delay difference can be made very small, these cancelers are preferable and yield good cancelation over a much wider signal bandwidths. The relative delay difference $\tau_1/2 - \tau_2$ can be held to under 0.3 ns for values of τ_1 and $\tau_2 \simeq 6nsec$ (as one may expect for a typical 14-18 GHz microwave circuitry). Using these values in the design of a power-correlation or correlation-correlation systems, one can obtain as much as 20dB input-output improvement in signal-to-interference ratio (corresponding to 40dB signal-to-interference ratio at the output when input cross-pol ratio $|b| \simeq |c| \simeq 0.1$ or 20dB) when the signals' bandwidth are of the order of 500MHz. On the

other hand, with the same assumption on accuracy of the delay timing we can achieve almost 34dB improvement (54dB output signal-to-interference ratio), if the signal bandwidths are only 100MHz. Similar examples show that the power-power scheme can handle only a small bandwidth if $\tau_1 = \tau_2$ is of the order of 6 nanoseconds. However, the power-power scheme might be applied with reasonable success in lower frequency low bandwidth application. For example, if the frequency band is in the 4-6 GHz range then τ_1 and τ_2 are of the order of 1 ns, so that 20dB input-output improvement in signal-to-interference can be accomplished with approximately 8 MHz. If the cancelation is performed in the 70 MHz IF frequency then the $\pi/4$ wavelength delay needed for the in-phase-quadrature weighting will restrict the delay τ to greater than 3.5 ns. If we assume $\tau_1 = \tau_2 = 5ns$ the bandwidth will be limited to approximately 2MHz. Notice, however, that the power-power scheme requires much simpler hardware than the other two schemes, so that the hardware complexity bandwidth trade-off is exhibited.

VI. ACKNOWLEDGEMENT

The author would like to acknowledge Dr.J.W. Carlin and Dr.M.L. Steinberger for their fruitful discussions, while some of the work was performed at AT&T Bell Labs.

VII. REFERENCES

- [1] Y.Bar-Ness, J.W.Carlin and M.L.Steinberger, "Bootstrapping Adaptive Cross Pol Cancelers for Satellite Communication," *Proc. of IEEE Int. Conf. on Comm. (ICC)*, no. 4F5, June 1982, Philadelphia, PA.
- [2] Y.Bar-Ness, "Bootstrapped Adaptive Cross-Pol Interference Cancelling-Steady State Analysis," submitted to signal processing, Elsevier Science Publisher, 1991.
- [3] Y.Bar-Ness and J.Rokah, "Cross-Coupled Bootstrapped Interference Canceler," 1981

Antenna and Propagation Symposium, conference proceedings, pp.292-295.

- [4] J.W. Carlin, Y.Bar-Ness, S.Gross, M.L.Steinberger, and W.E. Studdiford, "An IF Cross-Pol Canceler for Microwave Radio Systems," *Journal on Selected Areas in Comm.-Advances in Digital Comm. by Radio*, Vol. SAC-5, No.3, pp. 502-514, April 1987.
- [5] Y.Bar-Ness, "Bootstrapped Algorithms for Interference Cancelation." *AFCEA-IEEE Technical Conference on Tactical Comm.* pp. 111-118, May 1988.
- [6] Y.Bar-Ness, "Effect of Number of Taps on Cross-Pol Canceler Performance for Digital Radio," *The IEEE Global Comm. Conf. (Globecom)* 1987, paper no. 31.7.
- [7] Y.Bar-Ness and A.Dinç, "Performance Comparison of LMS, Diagonalizer and Bootstrapped Adaptive Cross-Pol Canceler Over Non-Dispersive Channel." *IEEE Military Comm. Conf. (Milcom)* 1990, paper 3.7.
- [8] A.Dinç and Y.Bar-Ness, "Error Probabilities of Bootstrapped Blind Adaptive Cross-pol Cancelers for Many QAM signals Over Non-Dispersive Fading Channel." *IEEE Int. Conf. on Comm. (ICC) 92*.
- [9] A. Dinç and Y. Bar-Ness, "Bootstrap: A Fast Unsupervised Learning Algorithm." *IEEE 1992 International Conference on Acoustic Speech and Signal Processing, ICASSP 92*. San Francisco, CA, paper no. 43.8.
- [10] C. Jatten and J. Herault, "Blind Separation of Sources. Part I. An Adaptive Algorithm Based on Neuromemetic Architecture," *Signal Processing*, Vol. 24, No. 1, July 1991.
- [11] E. Weinstein, M. Fedor and A. V. Oppenheim, "Multi Channel Signal Separation Based on Correlation," Submitted to *IEEE Trans. on ASSP*, 1992.
- [12] P.Z.Peebles, *Probability, Random Variables, and Random Signal Principles*. McGraw Hill, Inc., New York, NY.

VIII. APPENDICES

A-1

Substituting (13) in (8) and (14) in (12) we obtain, respectively;

$$\epsilon_1 = \frac{-(1-cb)\overline{|s|^2}\epsilon_2^*\delta_{s1} + b(1+bc-b\epsilon_2)^*\lambda_n(\tau_1+\tau_2)\delta_{n1}}{c\delta_{s1}\overline{|s|^2}\epsilon_2^* + (1-cb+b\epsilon_2)^*\delta_{n1}\overline{|n|^2} - b^*(\epsilon_2-c)^*\lambda_n(\tau_1+\tau_2)\delta_{n1}} \quad (\text{P-1})$$

and

$$\epsilon_2 = \frac{-(1-cb)\overline{|n|^2}\epsilon_1^*\delta_{n2} + b(1+bc-b\epsilon_1)^*\lambda_s(\tau_1+\tau_2)\delta_{s2}}{b\delta_{n2}\overline{|n|^2}\epsilon_1^* + (1-cb+b\epsilon_1)^*\delta_{n2}\overline{|s|^2} - c^*(\epsilon_1-b)^*\lambda_s(\tau_1+\tau_2)\delta_{s2}}. \quad (\text{P-2})$$

These two equations can be linearized by neglecting the second order terms that contain $\epsilon_1\epsilon_2^*$ and $\epsilon_2\epsilon_1^*$, namely;

$$\begin{aligned} & [(1-cb)^*\delta_{n1}\overline{|n|^2} + b^*c^*\delta_{n1}\lambda_n(\tau_1+\tau_2)]\epsilon_1 + [(1-cb)\delta_{s1}\overline{|s|^2} + \delta_{n1}\lambda_n(\tau_1+\tau_2)]\epsilon_2^* \\ & = b(1+bc)^*\delta_{n1}\lambda_n(\tau_1+\tau_2) \end{aligned} \quad (\text{P-3})$$

$$\begin{aligned} & [(1-cb)^*\delta_{s2}\overline{|s|^2} + b^*c^*\delta_{s2}\lambda_s(\tau_1+\tau_2)]\epsilon_2 + [(1-cb)\delta_{n2}\overline{|n|^2} + |c|^2\delta_{s2}\lambda_s(\tau_1+\tau_2)]\epsilon_1^* \\ & = c(1+bc)^*\delta_{s2}\lambda_s(\tau_1+\tau_2). \end{aligned} \quad (\text{P-4})$$

(P-3) and (P-4) also contain the effect of the two discriminator networks δ_{n1} , δ_{s1} , δ_{n2} and δ_{s2} . For simplicity of notation, we sometimes use $\overline{|n_1|^2} = \delta_{n1}\overline{|n|^2}$, $\overline{|n_2|^2} = \delta_{n2}\overline{|n|^2}$, $\overline{|s_1|^2} = \delta_{s1}\overline{|s|^2}$ and $\overline{|s_2|^2} = \delta_{s2}\overline{|s|^2}$. Equations (P-3) and (P-4) can be written as

$$\begin{aligned} & \begin{bmatrix} (1-cb)^*\overline{|n_1|^2} + b^*c^*\delta_{n1}\lambda_n(\tau_1+\tau_2) & (1-cb)\overline{|s_1|^2} + |b|^2\delta_{n1}\lambda_n(\tau_1+\tau_2) \\ (1-cb)^*\overline{|n_2|^2} + |c|^2\delta_{s2}\lambda_s(\tau_1+\tau_2) & (1-cb)\overline{|s_2|^2} + bc\delta_{s2}\lambda_s(\tau_1+\tau_2) \end{bmatrix} \\ & \times \begin{bmatrix} \epsilon_1 \\ \epsilon_2^* \end{bmatrix} = \begin{bmatrix} b(1+bc)^*\delta_{n1}\lambda_n(\tau_1+\tau_2) \\ c^*(1+bc)\delta_{s2}\lambda_s(\tau_1+\tau_2) \end{bmatrix}. \end{aligned} \quad (\text{P-5})$$

If $|b| \ll 1$ and $\tau_1 + \tau_2$ is small enough then it is reasonable to assume that

$|b|^2\delta_{n1}\lambda_n(\tau_1+\tau_2) \ll \text{Real}(1-cb)\overline{|s_1|^2}$. Similar argument implies that $|c|^2\delta_{s2}(\lambda_s(\tau_1+\tau_2)) \ll$

$Real(1 - cb)^* \overline{|n_2|^2}$. Also $\lambda_n(\tau_1 + \tau_2) \ll \overline{|n|^2}$, and $\lambda_s(\tau_1 + \tau_2) \ll \overline{|s|^2}$ so that the second terms in all the matrix elements can be neglected. Thus (P-5) becomes

$$\begin{bmatrix} (1 - cb)^* \overline{|n_1|^2} & (1 - cb) \overline{|s_1|^2} \\ (1 - cb)^* \overline{|n_2|^2} & (1 - cb) \overline{|s_2|^2} \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2^* \end{bmatrix} = \begin{bmatrix} b(1 + bc)^* \delta_{n1} \lambda_n(\tau_1 + \tau_2) \\ c^*(1 + bc) \delta_{s2} \lambda_s(\tau_1 + \tau_2) \end{bmatrix}. \quad (P-6)$$

The solution of (P-6) is given by

$$\epsilon_1 = \frac{b(1 + bc)^* \delta_{n1} - c^*(1 + bc) \delta_{s1} (\lambda_s(\tau_1 + \tau_2) / \lambda_n(\tau_1 + \tau_2))}{(1 - cb)^* (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n(\tau_1 + \tau_2)} \quad (P-7)$$

$$\epsilon_2 = \frac{c(1 + bc)^* \delta_{s2} - b^*(1 + bc) \delta_{n2} (\lambda_n(\tau_1 + \tau_2) / \lambda_s(\tau_1 + \tau_2))}{(1 - cb)^* (\rho / \delta_{n1}) \overline{|s|^2} / \lambda_s(\tau_1 + \tau_2)} \quad (P-8)$$

where

$$\rho = \delta_{s2} \delta_{n1} - \delta_{s1} \delta_{n2}$$

$$\lambda_s(\tau_1 + \tau_2) = \overline{|s|^2} - R_s(\tau_1 + \tau_2)$$

$$\lambda_n(\tau_1 + \tau_2) = \overline{|n|^2} - R_n(\tau_1 + \tau_2).$$

A-2

By definition of β ,

$$\beta b^* = -|b|^2 + \epsilon_1 b^*$$

and using (P-7)

$$\begin{aligned} \epsilon_1 b^* &= \frac{|b|^2(1 + bc)^* \delta_{n1} - (|bc|^2 + b^* c^*) \delta_{s1} (\lambda_s / \lambda_n)}{(1 - cb)^* (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n} \\ -|b|^2 + \epsilon_1 b^* &= \frac{-(1 - cb)^* |b|^2 (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n + |b|^2(1 + bc)^* \delta_{n1} - (|bc|^2 + b^* c^*) \delta_{s1} (\lambda_s / \lambda_n)}{(1 - cb)^* (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n} \\ &= \frac{-|b|^2 ((\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n - \delta_{n1} + |c|^2 \delta_{s1} (\lambda_s / \lambda_n)) + c^* b^* (|b|^2 (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n + |b|^2 \delta_{n1} - \delta_{s1} (\lambda_s / \lambda_n))}{(1 - cb)^* (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n} \end{aligned}$$

For simplicity of notation, we dropped the argument of λ_n and λ_s . If $\tau_1 + \tau_2$ is sufficiently small so that $\overline{|n|^2}/\lambda_n \gg 1$ then we can assume $|c|^2 \delta_{s1}(\lambda_s/\lambda_n) - \delta_{n1} \ll (\rho/\delta_{s2}) \overline{|n|^2}/\lambda_n$ and $|b|^2 \delta_{n1} \ll |b|^2 (\rho/\delta_{s2}) \overline{|n|^2}/\lambda_n$.

Therefore,

$$-|b|^2 + \epsilon_1 b^* \simeq -|b|^2 - \frac{b^* c^* \delta_{s1} \lambda_s / \lambda_n}{(1 - cb)^* (\rho / \delta_{s2}) \overline{|n|^2} / \lambda_n}.$$

Because $\overline{|n|^2}/\lambda_n \gg 1$ we conclude that,

$$\text{Real}[\beta b^*] \simeq \text{Real}[-|b|^2 + \epsilon_1 b] \simeq -|b|^2. \quad (\text{P} - 9)$$

A-3

To compare the magnitudes of the two terms in (20), we notice that both terms are monotonically decreasing with $\overline{|n|^2}/\lambda_n$. For $\overline{|n|^2}/\lambda_n$ very large the second term is larger than the first. Cross over occurs (i.e., the value for $\overline{|n|^2}/\lambda_n$ below which the second term becomes smaller than the first) for

$$\overline{|n|^2}/\lambda_n = \frac{|b(1 + bc)^* \delta_{n1} - c^*(1 + bc) \delta_{s1}(\lambda_s/\lambda_n)|^2}{|1 - cb|^2 (\rho/\delta_{s2})^2 |b|^2}. \quad (\text{P} - 10)$$

By its definition $\overline{|n|^2}/\lambda_n \geq 1$. Equality occurs only when τ_1 or τ_2 is sufficiently large. For $\tau_1 = \tau_2 = 0$ $\overline{|n|^2}/\lambda_n \rightarrow \infty$. It is easy to show that if

$$\frac{|1 + bc|(|b| \delta_{n1} + |c| \delta_{s1}(\lambda_s/\lambda_n))}{|1 - cb|(\rho/\delta_{s2})} < 1.41|b| \quad (\text{P} - 11)$$

then the right hand side of (P-10) is less than unity so that for almost all values of $\overline{|n|^2}/\lambda_n$, (20) is dominated by the second term. Since $\delta_{n1} < \delta_{s1} \simeq 1$, $\lambda_{s2}/\lambda_{n1} \simeq 1$, $\rho \simeq 1$ and $|b| \simeq |c| \ll 1$, then (P-11) is easily satisfied.

A-4

From (50)

$$A = \overline{v_{q_1}(t)v_p^*(t)}$$

and using (40) and (46) we obtain after some manipulation.

$$\begin{aligned} A = & (\alpha + c)\delta_{s_2}|s|^2 - \alpha\delta_{s_2}(\lambda_s(\tau_1/2 - \tau_2)) + \beta^*|\alpha + c|^2\delta_{s_2}R_s(\tau_1/2 - \tau_2) \\ & + \beta^*\alpha c^*\delta_{s_2}(\lambda_s(\tau_1/2 - \tau_2)) - \beta^*c\alpha^*\delta_{s_2}(R_s(\tau_1/2 - \tau_2) - R_s(\tau_1 - 2\tau_2)) \\ & + \beta^*|1 + \alpha b|^2\delta_{n_2}R_n(\tau_1/2 - \tau_2) + \beta^*\alpha b\delta_{n_2}(\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)) \\ & - \beta^*\alpha^*b^*\delta_{n_2}(R_n(\tau_1/2 - \tau_2) - R_n(\tau_1 - 2\tau_2)) \\ & + b^*(1 + \alpha b)\delta_{n_2}\overline{|n|^2} - \alpha|b|^2\delta_{n_2}(\overline{|n|^2} - R_n(\tau_1/2 - \tau_2)). \end{aligned}$$

Combining terms and using the approximation $\overline{|n|^2} - R_n(\tau_1/2 - \tau_2) \simeq R_n(\tau_1/2 - \tau_2) - R_n(\tau_1 - 2\tau_2)$ and $\overline{|s|^2} - R_s(\tau_1/2 - \tau_2) \simeq R_s(\tau_1/2 - \tau_2) - R_s(\tau_1 - 2\tau_2)$ we obtain.

$$\begin{aligned} A \simeq & (\alpha + c)\delta_{s_2}|s|^2 - \alpha\delta_{s_2}(\lambda_s(\tau_1/2 - \tau_2)) + \beta^*|\alpha + c|^2\delta_{s_2}R_s(\tau_1/2 - \tau_2) \\ & + 2j\beta^*I_m[\alpha c^*]\delta_{s_2}(\lambda_s(\tau_1/2 - \tau_2)) + \beta^*|1 + \alpha b|^2\delta_{n_2}(R_n(\tau_1/2 - \tau_2) \\ & + 2j\beta^*I_m[\alpha b]\delta_{n_2}(\lambda_n(\tau_1/2 - \tau_2)) + b^*(1 + \alpha b)\delta_{n_2}\overline{|n|^2} \\ & - \alpha|b|^2\delta_{n_2}(\lambda_n(\tau_1/2 - \tau_2))). \end{aligned} \quad (\text{P-12})$$

A-5

Using (54) in (53) we obtain,

$$(1 + \alpha b)^*\delta_{s_2}\overline{|s|^2}\epsilon_2 = -(|1 + \alpha b|^2\delta_{n_2}\overline{|n|^2} + v_s + v_n)\epsilon_1^* + b^*(v_s + v_n) \quad (\text{P-13})$$

where

$$v_s = 2jI_m[\alpha c^*]\delta_{s_2}(\lambda_s(\tau_1/2 - \tau_2))$$

and

$$v_n = 2jI_m[\alpha b]\delta_{n_2}(\lambda_n(\tau_1/2 - \tau_2)).$$

By using (55) and rearranging term, (P-13) becomes

$$\begin{aligned}
& (1 + \alpha b)^* \delta_{s2} \overline{|s|^2} \epsilon_2 - 2j b^* I_m[\epsilon_2 c^*] \delta_{s2}(\lambda_s(\tau_1/2 - \tau_2)) - 2j b^* I_m[\epsilon_2 b] \delta_{n2}(\lambda_n(\tau_1/2 - \tau_2)) \\
& + \{|1 + \alpha b|^2 \delta_{n2} \overline{|n|^2} + 2j I_m[\epsilon_2 c^*] \delta_{s2}(\lambda_s(\tau_1/2 - \tau_2)) - 2j I_m[(-c + \epsilon_2) b] \delta_{n2}(\lambda_n(\tau_1/2 - \tau_2))\} \epsilon_1^* \\
& = -2j I_m[cb] \delta_{n2}(\lambda_n(\tau_1/2 - \tau_2)). \tag{P-14}
\end{aligned}$$

Since $\lambda_s(\tau_1/2 - \tau_2)$ and $\lambda_n(\tau_1/2 - \tau_2)$ is much smaller than $\overline{|s|^2}$, it's reasonable to assume that both the real and imaginary parts of the second and third term of (P-14) are much smaller than the corresponding parts of the first term. Neglecting these terms as well the second order term in ϵ_1 and ϵ_2 we get;

$$\begin{aligned}
& (1 - bc)^* \delta_{s2} \overline{|s|^2} \epsilon_2 + \{|1 - bc|^2 \delta_{n2} \overline{|n|^2} + 2j I_m[cb] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2)\} \epsilon_1^* \\
& = -2j I_m[cb] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2). \tag{P-15}
\end{aligned}$$

Using (45) and recognizing that $\epsilon = b + \beta(1 + \alpha b)$ we get after some manipulation

$$\epsilon_1 = \frac{-(1 - bc)(\alpha + c)^* \delta_{s2} \overline{|s|^2} + 2b(1 + \alpha^* b^* + 2|\alpha b|^2) \delta_{n2} \lambda_n(\tau_1/2 - \tau_2)}{|1 + \alpha b|^2 \delta_{n2} \overline{|n|^2} + |\alpha + c|^2 \overline{|s|^2} - 4 \text{Real}[\alpha b] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2)}.$$

By neglecting $2|\alpha b|^2$ at the numerator, using $\alpha + c = \epsilon_2$ and neglecting the second order terms in ϵ_1 and ϵ_2 we obtain

$$\begin{aligned}
& |1 - bc|^2 \delta_{n1} \overline{|n|^2} \epsilon_1 + [(1 - bc) \overline{|s|^2} - 2|b|^2 \delta_{n1} \lambda_n(\tau_1/2 - \tau_2)] \epsilon_2^* \\
& = 2b(1 - bc)^* \delta_{n1} \lambda_n(\tau_1/2 - \tau_2). \tag{P-16}
\end{aligned}$$

Equations (P-15) and (P-16) are a set of linear equations in ϵ_1 and ϵ_2 , which can be written in matrix form;

$$\begin{aligned}
& \begin{bmatrix} |1 - cb|^2 \delta_{n1} \overline{|n|^2} & (1 - cb) \delta_{s1} \overline{|s|^2} - 2|b|^2 \delta_{n1} \lambda_n(\tau_1/2 - \tau_2) \\ |1 - cb|^2 \delta_{n2} \overline{|n|^2} - 2j I_m[cb] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2) & (1 - cb) \delta_{s2} \overline{|s|^2} \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2^* \end{bmatrix} \\
& = \begin{bmatrix} 2b(1 - bc)^* \delta_{n1} \lambda_n(\tau_1/2 - \tau_2) \\ 2j I_m[bc] \delta_{n2} \lambda_n(\tau_1/2 - \tau_2) \end{bmatrix}. \tag{P-17}
\end{aligned}$$

If $b \ll 1$ and $\tau_1/2 - \tau_2$ are sufficiently small then $2|b|^2 \delta_{n1} \lambda_n(\tau_1/2 - \tau_2) \ll \text{Real}(1 - bc) \delta_{s1} \overline{|s|^2}$.

With this in mind the solution of (P-17) is given by

$$\epsilon_1 = \frac{(1-bc)[2b(1-bc)^*\delta_{s2}\delta_{n1}\overline{|s|^2}\lambda_n(\tau_1/2 - \tau_2) - 2jI_m[cb]\delta_{s1}\delta_{n2}\overline{|s|^2}\lambda_n(\tau_1/2 - \tau_2)]}{(1-bc)|1-bc|^2(\rho|\overline{n|^2}|s|^2) + 2jI_m[cb]\overline{|s|^2}\lambda_n(\tau_1/2 - \tau_2)} \quad (\text{P-18})$$

$$\epsilon_2 = \frac{|1-bc|^2[-2jI_m[cb]\delta_{n1}\delta_{n2}\overline{|n|^2}\lambda_n(\tau_1/2 - \tau_2) - 2b^*(1-bc)\delta_{n1}\delta_{n2}\overline{|n|^2}\lambda_n(\tau_1/2 - \tau_2)]}{(1-bc)^*|1-bc|^2(\rho|\overline{n|^2}|s|^2) + 2jI_m[cb]\delta_{s1}\delta_{n2}\overline{|s|^2}\lambda_n(\tau_1/2 - \tau_2)} \quad (\text{P-19})$$

In obtaining (P-19) we used the assumption that the real of $2b(1-bc)^*\lambda_n(\tau_1/2 - \tau_2)$ is much smaller than $|1-bc|^2\overline{|n|^2}$ and the imaginary part of $2b(1-bc)^*I_m[cb]\lambda_n(\tau_1/2 - \tau_2)$ is much smaller than the real part of $b(1-bc)^*|1-bc|^2\overline{|n|^2}$. The second term at the denominator of (P-18) and (P-19) can easily be neglected in comparison to the first so that we can write

$$\epsilon_1 = \frac{2b(1-bc)^*\delta_{s2} - 2jI_m[cb]\delta_{s1}(\delta_{n2}/\delta_{n1})}{|1-bc|^2(\rho/\delta_{n1})\overline{|n|^2}/\lambda_n(\tau_1/2 - \tau_2)} \quad (\text{P-20})$$

$$\epsilon_2 = \frac{2jI_m[cb]\delta_{n1} + 2b^*(1-bc)\delta_{n1}}{|1-bc|^2(\rho/\delta_{n2})\overline{|s|^2}/\lambda_n(\tau_1/2 - \tau_2)} \quad (\text{P-21})$$

A-6

From (54) we have

$$\begin{aligned} \beta + \beta\alpha b &= -b + \epsilon_1 \\ \Rightarrow |\beta|^2 + |\beta|^2\alpha b &= \beta^*(-b + \epsilon_1) \\ |\beta|^2 \text{Real}[\alpha b] &= -|\beta|^2 - \text{Real}[\beta^*b] + \text{Real}[\beta^*\epsilon_1] \\ \Rightarrow |\beta|^2 \text{Real}[\alpha b] + \text{Real}[\beta^*b] &= -|\beta|^2 + \text{Real}[\beta^*\epsilon_1]. \end{aligned} \quad (\text{P-22})$$

A-7

For $\overline{|n|^2}/\lambda'_n$ very large the second term of (61) is larger than the first. The second term becomes smaller than the first only if

$$\left| \frac{2b(1-bc)^*\delta_{s2} - 2jI_m[cb]\delta_{s1}(\delta_{n2}/\delta_{n1})}{|1-bc|^2(\rho/\delta_{n1})\sqrt{2}|b|} \right|^2 > \overline{|n|^2}/\lambda'_n. \quad (\text{P-23})$$

Now if

$$\frac{|1 - bc||b|\delta_{s2} + I_m[cb]\delta_{s1}(\delta_{n2}/\delta_{n1})}{|1 - bc|^2(\rho/\delta_{n1})} < C|b|/1.41 \quad (\text{P} - 24)$$

then the left-hand side of (P-23) is less than C. But $\overline{|n|^2}/\lambda'_n$ is greater than unity, (in fact since λ_n is a function of $\tau_1/2 - \tau_2$; it can easily be made greater than some constant C). Therefore if (P-24) is satisfied then (P-23) cannot be satisfied, for all λ'_n such the $\overline{|n|^2}/\lambda'_n > C$, and the second term of (61) dominates the first. Since $\delta_{s1} < \delta_{s2}$, $\rho \simeq 1$ and $|b| \simeq |c| \ll 1$ then (P-24) is easily satisfied for C slightly greater than 1.41.

A-8

First we write (66)

$$Q_s = \left| \frac{2j I_m[cb]\delta_{n1} + 2b^*(1 - bc)\delta_{n1}}{(1 - bc)\rho(\lambda'_s/\lambda'_n)(|s|^2/\lambda'_s)} \right|^2 |s|^2 + \frac{2|c|^2}{|s|^2/\lambda'_s} |s|^2. \quad (\text{P} - 25)$$

If the two signals have approximately similar statistical properties and bandwidths, then with varying the delay, λ'_s/λ'_n remains approximately constant. Therefore, the first and second terms of (P-25) varies inversely proportional to $(|s|^2/\lambda'_s)^2$ and $|s|^2/\lambda'_s$ respectively. Thus applying the same argument of A-7 the second term of (P-25) dominates the first.

A-9

Using (85) and (86) in (83) and (84) we obtain respectively, after algebraic manipulation

$$\epsilon_1 = \frac{-(1 - bc)\overline{|s|^2}\delta_{s1}\epsilon_2^* - b(1 + bc - b\epsilon_2)^* \lambda_n(\tau_1/2 - \tau_2)\delta_{n1} - (|c|^2 + c^* - \epsilon_2^*)\lambda_s(\tau_1/2 - \tau_2)\delta_{s1}}{c\overline{|s|^2}\delta_{s1}\epsilon_2^* + (1 - cb + b\epsilon_2)^* \overline{|n|^2}\delta_{n1} - \lambda_n(\tau_1/2 - \tau_2)\delta_{n1} - |c|^2\lambda_s(\tau_1/2 - \tau_2)\delta_{s1}}$$

$$\epsilon_2 = \frac{-(1 - bc)\overline{|n|^2}\epsilon_1 - c(1 + bc - c\epsilon_1)^* \lambda_s(\tau_1/2 - \tau_2)\delta_{s2} - (|b|^2 + b^* - \epsilon_1^*)\lambda_n(\tau_1/2 - \tau_2)\delta_{n2}}{b\overline{|n|^2}\delta_{n2}\epsilon_1^* + (1 - cb + c\epsilon_1)^* \overline{|s|^2}\delta_{s2} - \lambda_s(\tau_1/2 - \tau_2)\delta_{s2} - |b|^2\lambda_n(\tau_1/2 - \tau_2)\delta_{n2}}$$

By neglecting the second order terms that contains $\epsilon_1\epsilon_2^*$ and $\epsilon_2\epsilon_1^*$, we can reduce these equations to the following linear set of equations,

$$\begin{aligned}
& [(1 - bc)^* \overline{|n|^2} \delta_{n1} - \lambda_n(\tau_1/2 - \tau_2) \delta_{n1} - |c|^2 \lambda_s(\tau_1/2 - \tau_2) \delta_{s1}] \epsilon_1 \\
& + [(1 - bc) \overline{|s|^2} \delta_{s1} - \lambda_s(\tau_1/2 - \tau_2) \delta_{s1} - |b|^2 \lambda_n(\tau_1/2 - \tau_2)] \epsilon_2^* \\
& = -b(1 + cb)^* \lambda_n(\tau_1/2 - \tau_2) \delta_{n1} - c^*(1 + cb) \lambda_s(\tau_1/2 - \tau_2) \delta_{s1}
\end{aligned}$$

$$\begin{aligned}
& [(1 - bc)^* \overline{|s|^2} \delta_{s2} - \lambda_s(\tau_1/2 - \tau_2) \delta_{s2} - |b|^2 \lambda_n(\tau_1/2 - \tau_2) \delta_{n2}] \epsilon_2 \\
& + [(1 - bc) \overline{|n|^2} \delta_{n2} - \lambda_n(\tau_1/2 - \tau_2) \delta_{n2} - |c|^2 \lambda_s(\tau_1/2 - \tau_2) \delta_{s2}] \epsilon_1^* \\
& = -c(1 + cb)^* \lambda_s(\tau_1/2 - \tau_2) \delta_{s2} - b^*(1 + cb) \lambda_n(\tau_1/2 - \tau_2) \delta_{n2}
\end{aligned}$$

or in matrix form;

$$\begin{aligned}
& \begin{bmatrix} ((1 - bc)^* \overline{|n|^2} - \lambda'_n) \delta_{n1} - |c|^2 \lambda'_s \delta_{s1} & ((1 - bc) \overline{|s|^2} - \lambda'_s) \delta_{s1} - |b|^2 \lambda'_n \delta_{n1} \\ ((1 - bc)^* \overline{|n|^2} - \lambda'_n) \delta_{n2} - |c|^2 \lambda'_s \delta_{s2} & ((1 - bc) \overline{|s|^2} - \lambda'_s) \delta_{s2} - |b|^2 \lambda'_n \delta_{n2} \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2^* \end{bmatrix} \\
& = \begin{bmatrix} -b(1 + bc)^* \lambda'_n \delta_{n1} - c^*(1 + bc) \lambda'_s \delta_{s1} \\ -b(1 + bc)^* \lambda'_n \delta_{n2} - c^*(1 + bc) \lambda'_s \delta_{s2} \end{bmatrix} \quad (\text{P-26})
\end{aligned}$$

where λ' stand for $\lambda(\tau_1/2 - \tau_2)$. If $|c| \ll 1$, $\delta_{s1} < \delta_{n1}$ then $|c|^2 \lambda'_s \delta_{s1} \ll |n|^2 \delta_{n1}$, particularly if $\overline{|s|^2} \simeq \overline{|n|^2}$ and $\tau_1/2 - \tau_2$ is very small in comparison to the signals' correlation time. Also, $\tau_1/2 - \tau_2$ small implies $\lambda'_n \ll \overline{|n|^2}$. This leaves $(1 - bc)^* \overline{|n|^2} \delta_{n1}$ and $(1 - bc) \overline{|s|^2} \delta_{s2}$ at the diagonal of the matrix in (P-26). By similar argument we can neglect in the off-diagonal terms the λ'_n and λ'_s term in comparison to $\overline{|n|^2}$ and $\overline{|s|^2}$ respectively. Thus (P-26) reduces

$$\begin{bmatrix} (1 - bc)^* \overline{|n|^2} \delta_{n1} & (1 - bc) \overline{|s|^2} \delta_{s1} \\ (1 - bc)^* \overline{|n|^2} \delta_{n2} & (1 - bc) \overline{|s|^2} \delta_{s2} \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2^* \end{bmatrix}$$

$$= \begin{bmatrix} -b(1+bc)^* \lambda'_n \delta_{n1} - c^*(1+bc) \lambda'_s \delta_{s1} \\ -b(1+bc)^* \lambda'_n \delta_{n2} - c^*(1+bc) \lambda'_s \delta_{s2} \end{bmatrix} \quad (\text{P-27})$$

whose solution is given by

$$\epsilon_1 = \frac{-b(1-bc)(1+bc)^* \overline{|s|^2} \lambda_n (\tau_1/2 - \tau_2)}{|1 - cb|^2 \overline{|s|^2} |n|^2} \quad (\text{P-28})$$

$$\epsilon_2 = \frac{-c(1-bc)(1+bc)^* \overline{|n|^2} \lambda_s (\tau_1/2 - \tau_2)}{|1 - cb|^2 \overline{|s|^2} |n|^2}. \quad (\text{P-29})$$

In the derivation of (P-28) we used; $\rho = \delta_{s2} \delta_{n1} - \delta_{s1} \delta_{n2}$ with $\delta_{s1} < \delta_{s2}$ and $\delta_{n2} < \delta_{n1}$. $|b|^2 \overline{|n|^2} \lambda'_n \delta_{n1} \delta_{n2}$ is of the same order as $|c|^2 \overline{|s|^2} \lambda'_s \delta_{s1} \delta_{s2}$, $c^* |b|^4 \lambda_n'^2 \delta_{n1} \delta_{n2} + c^* |b|^2 \lambda'_n \lambda'_s \delta_{n1} \delta_{s1} \ll c^* |b|^2 \overline{|s|^2} \lambda'_n \rho$, $b |cb|^2 \lambda'_n \lambda'_s \delta_{n1} \delta_{s2} \ll b |cb|^2 \overline{|s|^2} \lambda'_n \rho$ and $b |b|^2 \lambda_n'^2 \delta_{n1} \delta_{n2} \ll b \overline{|s|^2} \lambda'_n \rho$. Similar assumptions were applied in the derivation of (P-29).

	β_{op}	α_{op}
POWER-POWER SCHEME	$\frac{b(1 + \alpha b) \delta_{n1} n ^2 + (c + \alpha) \delta_{s1} s ^2 - b \lambda_n(\tau_1 + \tau_2) \delta_{n1}}{(1 + \alpha b) \delta_{n1} n ^2 + c(c + \alpha) \delta_{s1} s ^2 - \alpha \beta \lambda_n(\tau_1 + \tau_2) \delta_{n1}}$	$\frac{c(1 + \beta c) \delta_{s2} s ^2 + (b + \beta) \delta_{n2} n ^2 - c \lambda_n(\tau_1 + \tau_2) \delta_{s2}}{(1 + \beta c) \delta_{s2} s ^2 + b(b + \beta) \delta_{n2} n ^2 - \beta \alpha \lambda_n(\tau_1 + \tau_2) \delta_{s2}}$
POWER-CORR SCHEME	$\frac{b(1 + \alpha b) \delta_{n1} n ^2 + (c + \alpha) \delta_{s1} s ^2 - 2b(1 + \alpha b) \lambda_n(\tau_1/2 - \tau_2) \delta_{n1}}{ 1 + \alpha b ^2 \delta_{n1} n ^2 + \alpha + c ^2 \delta_{s1} s ^2 - 4 \operatorname{Re}\{(\alpha b) \lambda_n(\tau_1/2 - \tau_2) \delta_{n1}\}}$	$-c - (1 + \alpha b)(b + \beta(1 + \alpha b)) \delta_{n2} n ^2 / \delta_{s2} s ^2$ $- 2\beta \operatorname{Im}\{\alpha c\} \delta_{s2} \lambda_n(\tau_1/2 - \tau_2) + \alpha b \delta_{n2} \lambda_n(\tau_1/2 - \tau_2) / \delta_{s2} s ^2$
CORR-CORR SCHEME	$\frac{b(1 + \alpha b) \delta_{n1} n ^2 + (c + \alpha) \delta_{s1} s ^2 - \alpha \beta c ^2 \lambda_n(\tau_1/2 - \tau_2) \delta_{n1} - \alpha \lambda_n(\tau_1/2 - \tau_2) \delta_{s1}}{(1 + \alpha b) \delta_{n1} n ^2 + c(c + \alpha) \delta_{s1} s ^2 - \lambda_n(\tau_1/2 - \tau_2) \delta_{n1} - c ^2 \lambda_n(\tau_1/2 - \tau_2) \delta_{s1}}$	$\frac{c(1 + \beta c) \delta_{s2} s ^2 + (b + \beta) \delta_{n2} n ^2 - \beta c ^2 \lambda_n(\tau_1/2 - \tau_2) \delta_{s2} - \beta \lambda_n(\tau_1/2 - \tau_2) \delta_{n2}}{(1 + \beta c) \delta_{s2} s ^2 + b(b + \beta) \delta_{n2} n ^2 - \lambda_n(\tau_1/2 - \tau_2) \delta_{s2} - b ^2 \lambda_n(\tau_1/2 - \tau_2) \delta_{n2}}$

TABLE IA

	β_0	α_0
POWER-POWER SCHEME	$-b$	$-c$
POWER-CORR. SCHEME	$\frac{b}{1 + \alpha b}$	$-c$
CORR-CORR. SCHEME	$-b$	$-c$

TABLE IB

TABLE 1. The optimal weights β_{op} and α_{op} of the three different schemes (IA) and their corresponding zero-delay weights β_0 and α_0 . δ_{s1}, δ_{n1} represent the effect of the discriminator on output 1 and δ_{s2}, δ_{n2} the same effect on output 2. $\lambda_n = |n|^2 - R_n(\tau)$ and $\lambda_s(\tau) = |s|^2 - R_s(\tau)$ represent the effect of circuit delay with $R_n(\tau)$ and $R_s(\tau)$ are the autocorrelation function of $u(t)$ and $s(t)$ respectively.

	ϵ_1	ϵ_2
POWER-POWER SCHEME	$\frac{b(1+bc)^* \delta_{n1} - c^*(1+bc) \delta_{s1} (\lambda_s / \lambda_n)}{(1-bc)^* (\rho / \delta_{s2}) n ^2 / \lambda_n}$	$\frac{c(1+bc)^* \delta_{s2} - b^*(1+bc) \delta_{n2} (\lambda_n / \lambda_s)}{(1-bc)^* (\rho / \delta_{n1}) s ^2 / \lambda_s}$
POWER-CORR. SCHEME	$\frac{2b(1-bc)^* \delta_{s2} - 2j I_m [cb] \delta_{s1} (\delta_{n2} / \delta_{n1})}{(1-bc)^* (\rho / \delta_{n1}) n ^2 / \lambda_n'}$	$\frac{2j I_m [cb] \delta_{n1} + 2b^*(1-bc) \delta_{n1}}{(1-bc)^* (\rho / \delta_{n2}) s ^2 / \lambda_n'}$
CORR.-CORR. SCHEME	$\frac{-b(1+bc)^*}{(1-bc)^* \rho n ^2 / \lambda_n'}$	$\frac{-c(1+bc)^*}{(1-bc)^* \rho s ^2 / \lambda_s'}$

TABLE 2. Effect of circuits delay on the optimal weights $\epsilon_1 = \beta_{op} - \beta_0$, $\epsilon_2 = \alpha_{op} - \alpha_0$ with $\beta_{op}, \alpha_{op}, \beta_0$ and α_0 as in Table 1. [for power-correlator scheme $\epsilon_1 = (1 + \alpha b)(\beta_{op} - \beta_0)$, δ_{s1} and δ_{n1} , $i = 1, 2$ where defined in table 1. $\rho = \delta_{s2} \delta_{n1} - \delta_{n2} \delta_{s1}$] for simplicity we used $\lambda_n = \lambda_n(\tau_1 + \tau_2)$, $\lambda_s = \lambda_s(\tau_1 + \tau_2)$ and $\lambda_n' = \lambda_n(\tau_1/2 - \tau_2)$ and $\lambda_s' = \lambda_s(\tau_1/2 - \tau_2)$

		$\frac{(P_s/P_n)}{(s ^2/ n ^2)}$	$\frac{(Q_n/Q_s)}{(n ^2/ s ^2)}$
POWER-POWER SCHEME		$\frac{ 1 - bc ^2 n ^2 / \lambda_n(\tau_1 + \tau_2)}{2 b ^2}$	$\frac{ 1 - bc ^2 s ^2 / \lambda_s(\tau_1 + \tau_2)}{2 c ^2}$
POWER-CORR. SCHEME		$\frac{ n ^2 / \lambda_n(\tau_1/2 - \tau_2)}{2 b ^2}$	$\frac{ s ^2 / \lambda_s(\tau_1/2 - \tau_2)}{2 c ^2}$
CORR.-CORR. SCHEME		$\frac{ 1 - bc ^2 n ^2 / \lambda_n(\tau_1/2 - \tau_2)}{2 b ^2}$	$\frac{ 1 - bc ^2 s ^2 / \lambda_s(\tau_1/2 - \tau_2)}{2 c ^2}$

TABLE 3. Effect of circuits delay on P port and Q port outputs power ratios. $\lambda_n(\tau) = |n|^2 - R_n(\tau)$ and $\lambda_s(\tau) =$

$$|s|^2 - R_s(\tau)$$

		$(P_s/P_n)_{dB} - (s ^2/ b ^2 n ^2)_{dB}$	$(Q_n/Q_s)_{dB} - (n ^2/ c ^2 s ^2)_{dB}$
POWER-POWER SCHEME		$\geq (n ^2/\lambda_n(\tau_1 + \tau_2))_{dB} - 3_{dB}$	$\geq (s ^2/\lambda_s(\tau_1 + \tau_2))_{dB} - 3_{dB}$
POWER-CORR. SCHEME		$\geq (n ^2/\lambda_n(\tau_1/2 - \tau_2))_{dB} - 3_{dB}$	$\geq (s ^2/\lambda_s(\tau_1/2 - \tau_2))_{dB} - 3_{dB}$
CORR.-CORR. SCHEME		$\geq (n ^2/\lambda_n(\tau_1/2 - \tau_2))_{dB} - 3_{dB}$	$\geq (s ^2/\lambda_s(\tau_1/2 - \tau_2))_{dB} - 3_{dB}$

TABLE 4. Input-output improvement in signal-to-interference ratio.

	$\frac{(P_s/P_n)}{(s ^2/ n ^2)}$	$\frac{(Q_n/Q_s)}{(n ^2/ s ^2)}$
POWER-POWER SCHEME	$\geq \frac{ 1 - bc ^2}{ b ^2 [2\pi W_n (\tau_1 + \tau_2)]^2}$	$\geq \frac{ 1 - bc ^2}{ c ^2 [2\pi W_s (\tau_1 + \tau_2)]^2}$
POWER-CORR. SCHEME	$\geq \frac{1}{ b ^2 [2\pi W_n (\tau_1/2 - \tau_2)]^2}$	$\geq \frac{1}{ c ^2 [2\pi W_s (\tau_1/2 - \tau_2)]^2}$
CORR.-CORR. SCHEME	$\geq \frac{ 1 - bc ^2}{ b ^2 [2\pi W_n (\tau_1/2 - \tau_2)]^2}$	$\geq \frac{ 1 - bc ^2}{ c ^2 [2\pi W_s (\tau_1/2 - \tau_2)]^2}$

TABLE 5. Output power ratio as a function of singals' bandwidth, W_s and W_n , and circuits delay τ_1, τ_2

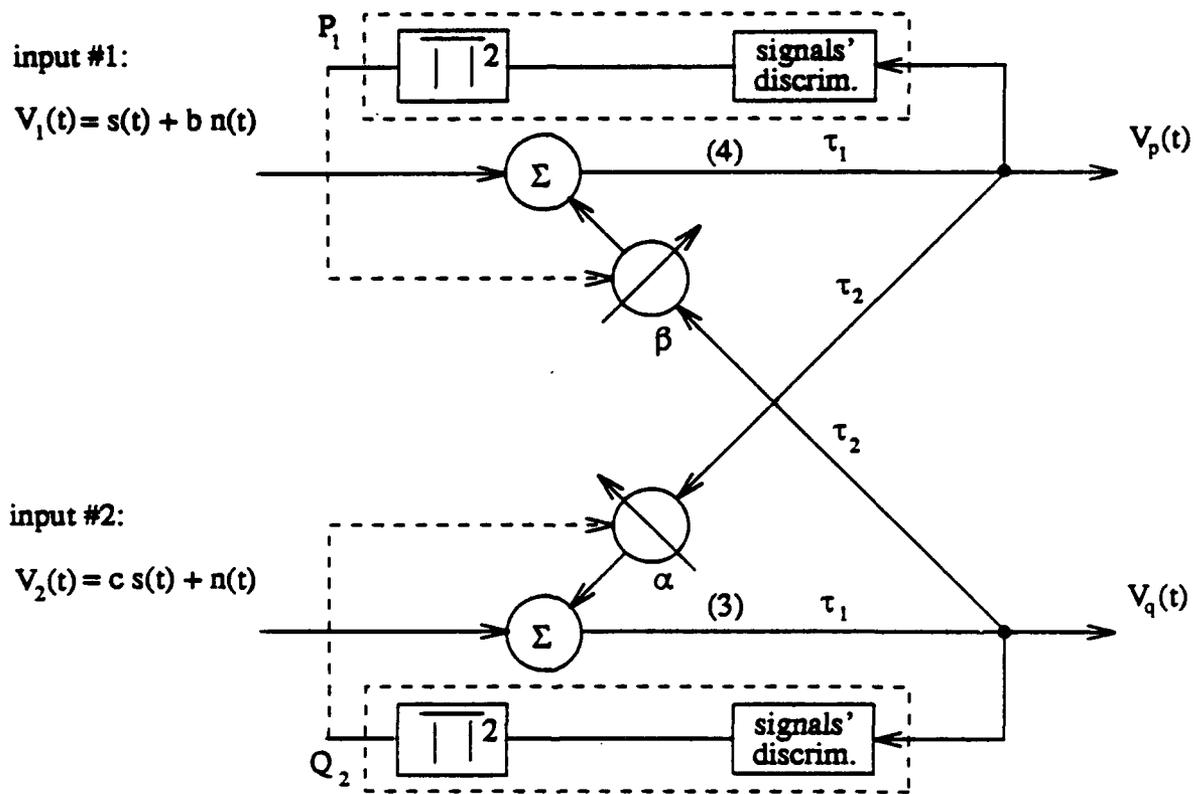


Figure 1. Power-Power Bootstrapped Cross-Pol Canceler.

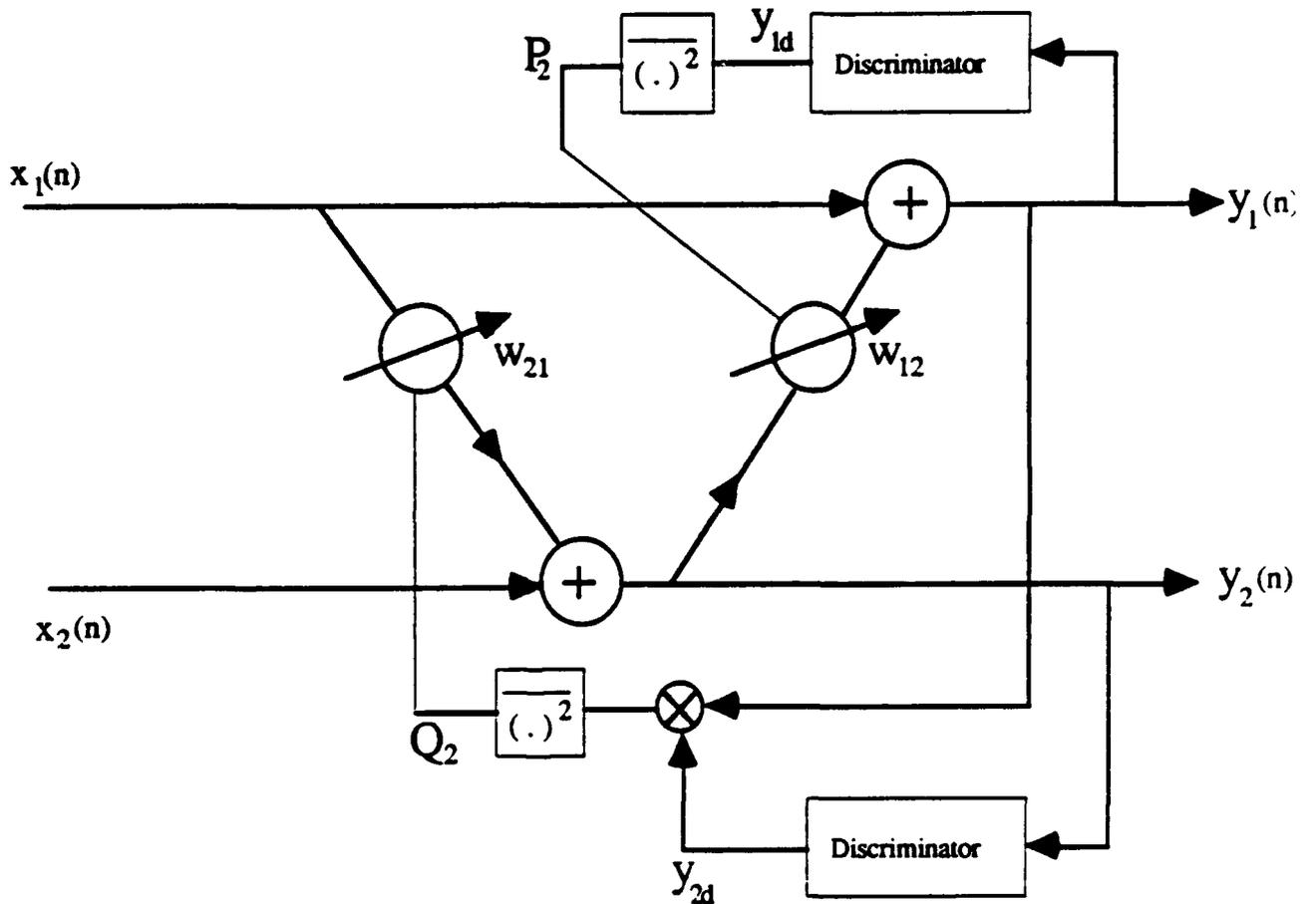


Figure 2. Power-Correlator Bootstrapped Cross-Pol Canceler.

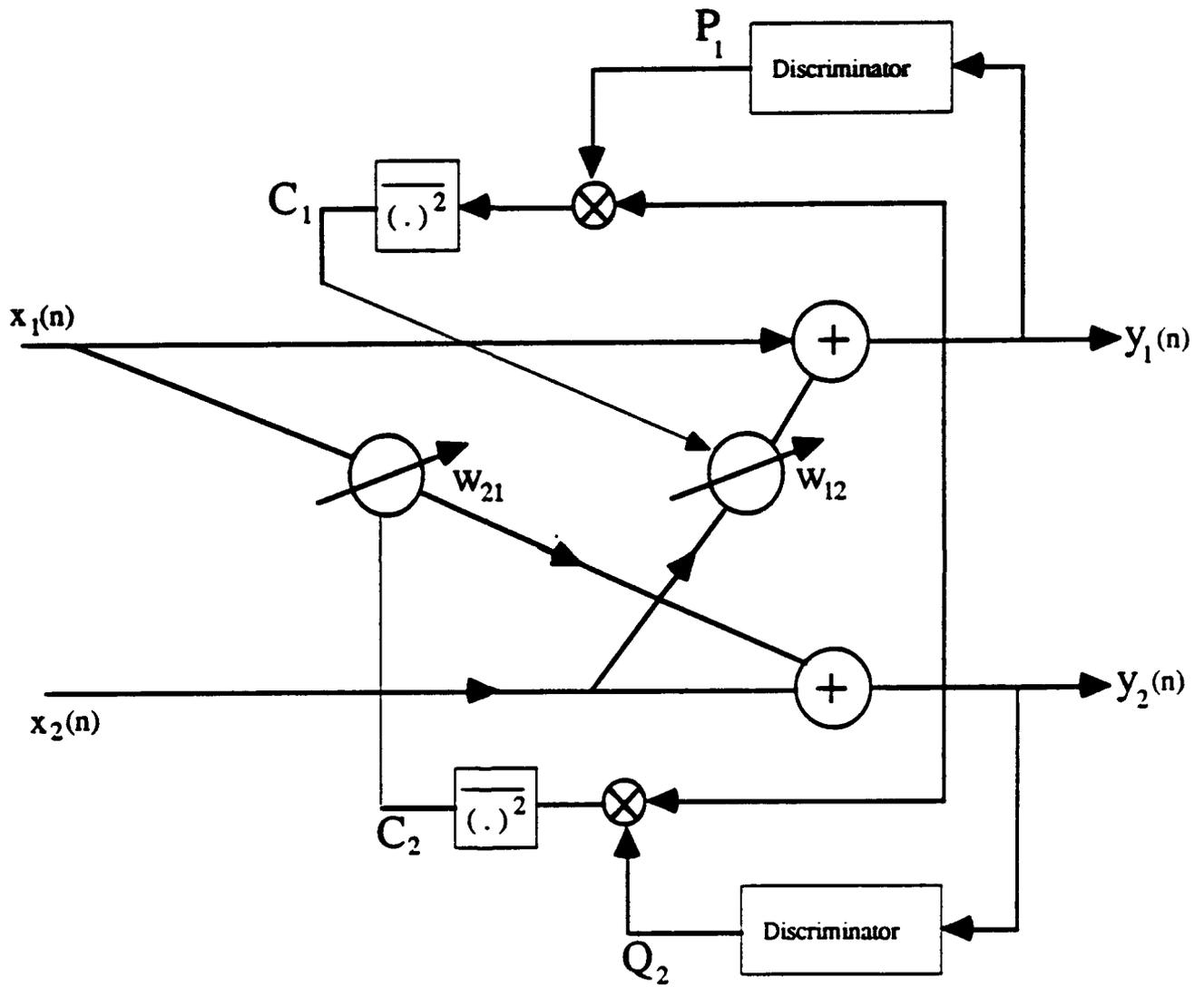


Figure 3. Correlator-Correlator Bootstrapped Cross-Pol Canceler.

APPENDIX B
BOOTSTRAPPED ADAPTIVE SEPARATION OF TWO
SUPERIMPOSED SIGNALS -
STABILITY CONSIDERATIONS

by

Abdulkadir Dinç and Yeheskel Bar-Ness

I. INTRODUCTIONS

In a previous study [1], we found the equilibrium points for the weights of the bootstrapped algorithm. The question which arises is whether these equilibriums are stable steady state points. We will answer this question for the three schemes of the bootstrapped algorithm, the power-power, correlator-correlator and power-correlator, separately. We will restrict our discussion to the case of no noise; that is the dual channel noises $E\{n_1^2(n)\}$ and $E\{n_2^2(n)\}$ are zero. Also, for the sake of simplicity, we will consider the signal to be real. This is the case, for example, when the transmitted data is an M-ary signal. From Fig.1. the channel response in the no noise case is,

$$\begin{aligned} x_1(n) &= a_{11}I_1(n) + a_{12}I_2(n) \\ x_2(n) &= a_{21}I_1(n) + a_{22}I_2(n) \end{aligned} \quad (1)$$

where $x_1(n)$ and $x_2(n)$ are the sampled received signals at the first and second channels respectively. $I_i(n)$ $i = 1, 2$ are the inputs of the channel, which are taken to be real, equally likely distributed from the set $\{\pm 1c, \pm 3c, \dots, \pm(\sqrt{M}-1)c\}$ and c is a constant which determines the distance to the decision boundary from each signal location. Also, we will assume in this study the channel co-pol and cross-pol responses to be real, that is,

$$\frac{a_{12}}{a_{22}} = r_1, \quad \frac{a_{21}}{a_{11}} = r_2 \quad (2)$$

r_1, r_2 denote the magnitude of the normalized cross-pol interference (XPI) constants of one channel onto the other.

II. EQUILIBRIUM POINTS

2.1 The Power-Power Scheme

With the arrangement of the power-power cross-pol canceler (XPC) of Fig. 2, the canceler outputs, $y_1(n)$ and $y_2(n)$ are given by,

$$y_1(n) = \frac{x_1(n) + x_2(n)w_{12}}{1 - w_{12}w_{21}} \quad (3)$$

$$y_2(n) = \frac{x_2(n) + x_1(n)w_{21}}{1 - w_{12}w_{21}} \quad (4)$$

Substituting for $x_1(n)$ and $x_2(n)$ from (1), we get

$$y_1(n) = \frac{I_1(n)(a_{11} + w_{12}a_{21}) + I_2(n)(a_{12} + w_{12}a_{22})}{1 - w_{12}w_{21}} \quad (5)$$

$$y_2(n) = \frac{I_1(n)(a_{21} + w_{21}a_{11}) + I_2(n)(a_{22} + w_{21}a_{12})}{1 - w_{12}w_{21}} \quad (6)$$

For this scheme the control algorithm simultaneously minimizes the output powers P and Q . In fact it simultaneously searches for $\partial E\{y_{1d}(n)^2\}/\partial w_{12} = 0$ and $\partial E\{y_{2d}(n)^2\}/\partial w_{21} = 0$, where $E\{\cdot\}$ is the expected value and y_{1d}, y_{2d} are the outputs of the discriminators. Finding the zero derivatives can be done by successive use of the following recursive equations, provided that $1 - w_{12}w_{21} \neq 0$.

$$w_{12}^{i+1} = w_{12}^i + \mu_1 \frac{\partial}{\partial w_{12}^i} E\{y_{1d}^i(n)^2\} \quad (7)$$

$$w_{21}^{i+1} = w_{21}^i + \mu_2 \frac{\partial}{\partial w_{21}^i} E\{y_{2d}^i(n)^2\} \quad (8)$$

where μ_1 and μ_2 are the constants which determine the stability of convergence. The discriminators enforce a change in the powers of the I_1 and I_2 signals by δ_{ij} where $i=1,2$ refers

to the two outputs y_1 and y_2 , respectively, while $j=1$ refers to signal I_1 and I_2 , respectively. Due to the assumption that the channel responses and the signals are real, w_{12} and w_{21} are also real. Clearly, the equilibrium points must simultaneously satisfy the following equations

$$\mu_1 \frac{\partial P_1(w_{12}, w_{21})}{\partial w_{12}} = \mu_1 \frac{\partial}{\partial w_{12}} E\{y_{1d}(n)^2\} = 0 \quad (9)$$

$$\mu_2 \frac{\partial Q_1(w_{12}, w_{21})}{\partial w_{21}} = \mu_2 \frac{\partial}{\partial w_{21}} E\{y_{2d}(n)^2\} = 0, \quad (10)$$

where by using (5) and (6), we get the power at the output of the discriminators:

$$P_1(w_{12}, w_{21}) = \frac{\delta_{11}E\{I_1^2(n)\}(a_{11} + w_{12}a_{21})^2 + \delta_{12}E\{I_2^2(n)\}(a_{12} + w_{12}a_{22})^2}{(1 - w_{12}w_{21})^2} \quad (11)$$

$$Q_1(w_{12}, w_{21}) = \frac{\delta_{21}E\{I_1^2(n)\}(a_{21} + w_{21}a_{11})^2 + \delta_{22}E\{I_2^2(n)\}(a_{22} + w_{21}a_{12})^2}{(1 - w_{12}w_{21})^2} \quad (12)$$

w_{12opt} and w_{21opt} are taken to be the steady state value of these weights from (7) and (8), respectively. That is, they satisfy (9) and (10). Taking the derivative of (11) and (12) with respect to w_{12} and w_{21} and multiplying with the convergence constants respectively we get:

$$\begin{aligned} \mu_1 \frac{\partial P_1(w_{12}, w_{21})}{\partial w_{12}} &= \frac{2\mu_1}{(1 - w_{12}w_{21})^3} \left[\delta_{11}E\{I_1^2(n)\}(a_{11} + w_{12}a_{21})(a_{21} + w_{21}a_{11}) \right. \\ &\quad \left. + \delta_{12}E\{I_2^2(n)\}(a_{12} + w_{12}a_{22})(a_{22} + w_{21}a_{12}) \right] \end{aligned} \quad (13)$$

$$\begin{aligned} \mu_2 \frac{\partial Q_1(w_{12}, w_{21})}{\partial w_{21}} &= \frac{2\mu_2}{(1 - w_{12}w_{21})^3} \left[\delta_{21}E\{I_1^2(n)\}(a_{11} + w_{12}a_{21})(a_{21} + w_{21}a_{11}) \right. \\ &\quad \left. + \delta_{22}E\{I_2^2(n)\}(a_{12} + w_{12}a_{22})(a_{22} + w_{21}a_{12}) \right] \end{aligned} \quad (14)$$

Note that without the inclusion of the discriminator the two equations in (13) and (14) are dependent. If the discriminator is chosen such that $\delta_{11}\delta_{22} \neq \delta_{21}\delta_{12}$ then (13) and (14) are independent. These solutions determine the equilibrium points for (7) and (8). Writing these equations in matrix form

$$\begin{bmatrix} \delta_{11}E\{I_1^2(n)\} & \delta_{12}E\{I_2^2(n)\} \\ \delta_{21}E\{I_1^2(n)\} & \delta_{22}E\{I_2^2(n)\} \end{bmatrix} \begin{bmatrix} (a_{11} + w_{12}a_{21})(w_{21}a_{11} + a_{21}) \\ (a_{12} + w_{12}a_{22})(w_{21}a_{12} + a_{22}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (15)$$

we notice that if $\delta_{11}\delta_{22} \neq \delta_{12}\delta_{21}$ these equations are independent. Therefore (13) and (14) become zero if and only if

$$(a_{11} + w_{12}a_{21})(w_{21}a_{11} + a_{21}) = 0 \quad (16)$$

and

$$(a_{12} + w_{12}a_{22})(w_{21}a_{12} + a_{22}) = 0. \quad (17)$$

There are two solutions for w_{12} and w_{21} (assuming $\frac{a_{21}}{a_{11}} \neq \frac{a_{22}}{a_{12}}$).

$$(1) \quad w_{12\text{opt1}} = -\frac{a_{12}}{a_{22}} = -r_1 \quad w_{21\text{opt1}} = -\frac{a_{21}}{a_{11}} = -r_2 \quad (18)$$

$$(2) \quad w_{12\text{opt2}} = -\frac{a_{11}}{a_{21}} = -\frac{1}{r_2} \quad w_{21\text{opt2}} = -\frac{a_{22}}{a_{12}} = -\frac{1}{r_1} \quad (19)$$

2.2 The Correlator-Correlator Scheme

From Fig. 3, the outputs of the canceler can be written as,

$$\begin{aligned} y_1(n) &= x_1(n) + w_{12}x_2(n) \\ y_2(n) &= x_2(n) + w_{21}x_1(n). \end{aligned} \quad (20)$$

Substituting (1) in (20), we get

$$\begin{aligned} y_1(n) &= I_1(n)[a_{11} + w_{12}a_{21}] + I_2(n)[a_{12} + w_{12}a_{22}] \\ y_2(n) &= I_1(n)[a_{21} + w_{21}a_{11}] + I_2(n)[a_{22} + w_{21}a_{12}]. \end{aligned} \quad (21)$$

The control algorithm simultaneously minimizes the output correlation powers $C_1 = A_1^2(w_{12}, w_{21}) = [E\{y_{1d}(n)y_2(n)\}]^2$ and $C_2 = B_1^2(w_{12}, w_{21}) = [E\{y_1(n)y_{2d}(n)\}]^2$. It simultaneously searches for $\partial C_1(w_{12}, w_{21})/\partial w_{12} = 0$ and $\partial C_2(w_{12}, w_{21})/\partial w_{21} = 0$, where y_{1d} and y_{2d} are the discriminator outputs. This can be performed by successive use of the following recursive equations.

$$w_{12}^{i+1} = w_{12}^i + \mu_1 \frac{\partial}{\partial w_{12}^i} [A_1(w_{12}^i, w_{21}^i)]^2 \quad (22)$$

$$w_{21}^{i+1} = w_{21}^i + \mu_2 \frac{\partial}{\partial w_{21}^i} [B_1(w_{12}^i, w_{21}^i)]^2 \quad (23)$$

where $A_1(w_{12}, w_{21})$ and $B_1(w_{12}, w_{21})$ are the correlation between the output of channel 1. and the output of channel 2 after discrimination, and vice versa, respectively. That is:

$$A_1(w_{12}, w_{21}) = E\{y_{1d}(n)y_2(n)\} = \delta_{11}E\{I_1^2(n)\}(a_{11} + w_{12}a_{21})(w_{21}a_{11} + a_{21}) \\ + \delta_{12}E\{I_2^2(n)\}(a_{12} + w_{12}a_{22})(w_{21}a_{12} + a_{22}) \quad (24)$$

$$B_1(w_{12}, w_{21}) = E\{y_1(n)y_{2d}(n)\} = \delta_{21}E\{I_1^2(n)\}(a_{11} + w_{12}a_{21})(w_{21}a_{11} + a_{21}) \\ + \delta_{22}E\{I_2^2(n)\}(a_{12} + w_{12}a_{22})(w_{21}a_{12} + a_{22}). \quad (25)$$

Notice again that for equations (21) and (22) to be independent, discrimination effects $\delta_{i,j}$ are necessary.

The optimum weights can be obtained from

$$\mu_1 \frac{\partial C_1(w_{12}, w_{21})}{\partial w_{12}} = 2\mu_1 A_1(w_{12}, w_{21}) \frac{\partial}{\partial w_{12}} A_1(w_{12}, w_{21}) = 0 \quad (26)$$

$$\mu_2 \frac{\partial C_2(w_{12}, w_{21})}{\partial w_{21}} = 2\mu_2 B_1(w_{12}, w_{21}) \frac{\partial}{\partial w_{21}} B_1(w_{12}, w_{21}) = 0. \quad (27)$$

It was shown in [1] that for (24) and (25) to equal zero simultaneously, it is necessary to have $A_1(w_{12}, w_{21}) = B_1(w_{12}, w_{21}) = 0$, simultaneously. Therefore, equations (24) and (25) are simultaneously zero if and only if $A_1(w_{12}, w_{21})$ and $B_1(w_{12}, w_{21})$ equal zero. The optimum weights w_{12opt} and w_{21opt} are found by equating (24) and (25) to zero, respectively. Except for the normalization (assumed non zero) these equations are the same as (13) and (14) so that the two possible optimum weights are the same as in the power-power scheme.

2.3 The Power-Correlator Scheme

From Fig. 4, the canceler outputs are

$$y_1(n) = x_1(n)(1 + w_{12}w_{21}) + w_{21}x_2(n) \\ y_2(n) = x_1(n)w_{21} + x_2(n). \quad (28)$$

Substituting (1) in (28), we get

$$y_1(n) = I_1(n)[a_{11} + w_{12}(a_{21} + w_{21}a_{11})] + I_2(n)[a_{12} + w_{12}(a_{22} + w_{21}a_{12})] \quad (29)$$

$$y_2(n) = I_1(n)[a_{21} + w_{21}a_{11}] + I_2(n)[a_{22} + w_{21}a_{12}]. \quad (30)$$

In this case, the control algorithm simultaneously minimizes the output power $P_2 = E\{y_{1d}(n)^2\}$ and the correlation power $Q_2 = B_2^2(w_{12}, w_{21}) = \{E\{y_1(n)y_{2d}(n)\}\}^2$. It simultaneously searches for $\partial E\{y_{1d}(n)^2\}/\partial w_{12} = 0$ and $\partial B_2^2(w_{12}, w_{21})/\partial w_{21} = 0$, where y_{1d} , y_{2d} are the discriminator outputs. This can be performed by successive use of the following recursive equations.

$$w_{12}^{i+1} = w_{12}^i + \mu_1 \frac{\partial}{\partial w_{12}^i} P_2(w_{12}, w_{21}) \quad (31)$$

$$w_{21}^{i+1} = w_{21}^i + \mu_2 \frac{\partial}{\partial w_{21}^i} Q_2(w_{12}, w_{21}) \quad (32)$$

where μ_1 and μ_2 are the constants which determine the stability of convergence. From (29), we write the power at the output of the channel 1 discriminator

$$P_2(w_{12}, w_{21}) = \delta_{11} E\{I_1^2(n)\} [a_{11} + w_{12}(a_{21} + w_{21}a_{11})]^2 + \delta_{12} E\{I_2^2(n)\} [a_{12} + w_{12}(a_{22} + w_{21}a_{12})]^2 \quad (33)$$

and the correlation between the output of channel 1 and that of channel 2 after discrimination;

$$B_2(w_{12}, w_{21}) = \delta_{21} E\{I_1^2(n)\} [a_{11} + w_{12}(a_{21} + w_{21}a_{11})][a_{21} + w_{21}a_{11}] + \delta_{22} E\{I_2^2(n)\} [a_{12} + w_{12}(a_{22} + w_{21}a_{12})][a_{22} + w_{21}a_{12}]. \quad (34)$$

Taking the derivative of (33) with respect to w_{12} and multiplying by the convergence constant μ_1 , we get;

$$\mu_1 \frac{\partial P_2(w_{12}, w_{21})}{\partial w_{12}} = 2\mu_1 \left[\delta_{11} E\{I_1^2(n)\} [a_{11} + w_{12}(a_{21} + w_{21}a_{11})](a_{21} + w_{21}a_{11}) + \delta_{12} E\{I_2^2(n)\} [a_{12} + w_{12}(a_{22} + w_{21}a_{12})][a_{22} + w_{21}a_{12}] \right]. \quad (35)$$

From the definition of $B_2(w_{12}, w_{21})$, we have $Q_2(w_{12}, w_{21}) = (B_2(w_{12}, w_{21}))^2$, therefore,

$$\mu_2 \frac{\partial Q_2(w_{12}, w_{21})}{\partial w_{21}} = 2\mu_2 B_2(w_{12}, w_{21}) \frac{\partial}{\partial w_{21}} B_2(w_{12}, w_{21}) \quad (36)$$

Similar to the correlator-correlator scheme, we showed in [1] that (36) becomes zero if and only if $B_2(w_{12}, w_{21}) = 0$. Writing (34) and (35) in matrix form and equating the result to zero;

$$\begin{bmatrix} \delta_{11}E\{I_1^2(n)\} & \delta_{12}E\{I_2^2(n)\} \\ \delta_{21}E\{I_1^2(n)\} & \delta_{22}E\{I_2^2(n)\} \end{bmatrix} \begin{bmatrix} (a_{11} + w_{12}(a_{21} + w_{21}a_{11}))(w_{21}a_{11} + a_{21}) \\ (a_{12} + w_{12}(a_{22} + w_{21}a_{12}))(w_{21}a_{12} + a_{22}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (37)$$

We notice that if $\delta_{11}\delta_{22} \neq \delta_{12}\delta_{21}$ these equations are independent, so that (34) and (35) become zero if and only if

$$(a_{11} + w_{12}(a_{21} + w_{21}a_{11}))(w_{21}a_{11} + a_{21}) = 0,$$

and

$$(a_{12} + w_{12}(a_{22} + w_{21}a_{12}))(w_{21}a_{12} + a_{22}) = 0.$$

There are two solutions for w_{12} and w_{21} (assuming $\frac{a_{21}}{a_{11}} \neq \frac{a_{22}}{a_{12}}$);

$$w_{12opt1} = -\frac{a_{12}}{a_{22}(1 - \frac{a_{21}a_{12}}{a_{11}a_{22}})} \quad w_{21opt1} = -\frac{a_{21}}{a_{11}} = -r_2 \quad (38)$$

and;

$$w_{12opt2} = -\frac{a_{11}}{a_{21}(1 - \frac{a_{22}a_{11}}{a_{21}a_{12}})} \quad w_{21opt2} = -\frac{a_{22}}{a_{12}} = \frac{1}{r_1}. \quad (39)$$

III. STABILITY PARAMETERS

Equations (7) (8), (22) ,(23) and (31), (32) are all nonlinear in w_{12} and w_{21} . Therefore to classify the equilibrium points of these equations , we will consider a small deviation from the equilibrium points, i.e, by varying w_{12} and w_{21} to $w_{12opt} + \Delta w_{12}$ and $w_{21opt} + \Delta w_{21}$.

respectively with Δw_{12} and Δw_{21} very small. For Δw_{12} and Δw_{21} small, any differentiable function $X(w_{12}, w_{21})$ satisfying $\frac{\partial X(w_{12}, w_{21})}{\partial w_{12}} = 0$ can be approximated by

$$\begin{aligned} \frac{\partial X_1(w_{12opt} + \Delta w_{12}, w_{21opt} + \Delta w_{21})}{\partial w_{12}} &= \frac{\partial^2 X_1}{\partial w_{12}^2} \Big|_{\mathbf{w}_{opt}} \Delta w_{12} \\ &+ \frac{\partial^2 X_1}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{opt}} \Delta w_{21} \end{aligned} \quad (40)$$

$$\begin{aligned} \frac{\partial X_2(w_{12} + \Delta w_{12}, w_{21} + \Delta w_{21})}{\partial w_{21}} &= \frac{\partial^2 X_2}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{opt}} \Delta w_{12} \\ &+ \frac{\partial^2 X_2}{\partial w_{21}^2} \Big|_{\mathbf{w}_{opt}} \Delta w_{21} \end{aligned} \quad (41)$$

where

$$\mathbf{w}_{opt} = [w_{12opt}, w_{21opt}]^T. \quad (42)$$

In matrix notation

$$\dot{\mathbf{X}} = \mathbf{A} \Delta \mathbf{w} \quad (43)$$

where $\Delta \mathbf{w} = [\Delta w_{12}, \Delta w_{21}]^T$ and

$$\dot{\mathbf{X}} = \left[\frac{\partial X_1(w_{12} + \Delta w_{12}, w_{21} + \Delta w_{21})}{\partial w_{12}}, \frac{\partial X_2(w_{12} + \Delta w_{12}, w_{21} + \Delta w_{21})}{\partial w_{21}} \right]^T \quad (44)$$

where

$$\mathbf{A} = \begin{bmatrix} \frac{\partial^2 X_1}{\partial w_{12}^2} \Big|_{\mathbf{w}_{opt}} & \frac{\partial^2 X_1}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{opt}} \\ \frac{\partial^2 X_2}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{opt}} & \frac{\partial^2 X_2}{\partial w_{21}^2} \Big|_{\mathbf{w}_{opt}} \end{bmatrix}. \quad (45)$$

The stability of equilibrium points depend on the eigenvalues of matrix \mathbf{A} . Considering the characteristics equation of the matrix \mathbf{A} from $|\lambda \mathbf{I} - \mathbf{A}| = 0$, we can find the eigenvalues of \mathbf{A} by solving

$$\lambda^2 - b\lambda + c = 0 \quad (46)$$

where

$$\mathbf{b} = \frac{\partial^2 X_1}{\partial w_{12}^2} \Big|_{\mathbf{w}_{\text{opt}}} + \frac{\partial^2 X_2}{\partial w_{21}^2} \Big|_{\mathbf{w}_{\text{opt}}} \quad (47)$$

$$\mathbf{c} = \frac{\partial^2 X_1}{\partial w_{12}^2} \Big|_{\mathbf{w}_{\text{opt}}} \frac{\partial^2 X_2}{\partial w_{21}^2} \Big|_{\mathbf{w}_{\text{opt}}} - \frac{\partial^2 X_1}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} \frac{\partial^2 X_2}{\partial w_{12} \partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}}. \quad (48)$$

The nature of the eigenvalues of \mathbf{A} in the complex plane, or equivalently the relation between \mathbf{b} and \mathbf{c} defined in (47) and (48) determines the classification of the equilibrium point. Next, we intend to find the different entries of \mathbf{A} for the different bootstrapped schemes.

3.1 Stability Parameters of The Power-Power Scheme

Here, $X_1(w_{12}, w_{21})$ and $X_2(w_{12}, w_{21})$ are given by $\mu_1 P_1(w_{12}, w_{21})$ and $\mu_2 Q_1(w_{12}, w_{21})$, respectively from (13) and (14). First notice that for any rational function in x , $f(x) = \frac{N(x)}{D(x)}$

$$\begin{aligned} \frac{df(x)}{dx} \Big|_{N(x)=0} &= \frac{D(x) \frac{dN(x)}{dx} - N(x) \frac{dD(x)}{dx}}{D^2(x)} \Big|_{N(x)=0} \\ &= \frac{dN(x)}{dx} \frac{1}{D(x)}. \end{aligned} \quad (49)$$

Using this relation in (13) and (14), we get

$$\begin{aligned} \mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} \Big|_{\mathbf{w}_{\text{opt}}} &= \frac{2\mu_1}{(1 - w_{12\text{opt}} w_{21\text{opt}})^3} \left[\delta_{11} E\{I_1^2(n)\} \right. \\ &\quad \left. (a_{21} + w_{21\text{opt}} a_{11}) a_{21} + \delta_{12} E\{I_2^2(n)\} (a_{22} + w_{21\text{opt}} a_{12}) a_{22} \right] \end{aligned} \quad (50)$$

$$\begin{aligned} \mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21}^2} \Big|_{\mathbf{w}_{\text{opt}}} &= \frac{2\mu_2}{(1 - w_{12\text{opt}} w_{21\text{opt}})^3} \left[\delta_{21} E\{I_1^2(n)\} \right. \\ &\quad \left. (a_{11} + w_{12\text{opt}} a_{21}) a_{11} + \delta_{22} E\{I_2^2(n)\} (a_{12} + w_{12\text{opt}} a_{22}) a_{12} \right]. \end{aligned} \quad (51)$$

Substituting for \mathbf{w}_{opt} from (18) and (19) in (50), we obtain respectively

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} \Big|_{\mathbf{w}_{\text{opt}1}} = 2\mu_1 \frac{\delta_{12} E\{I_2^2(n)\} a_{22}^2}{(1 - r_1 r_2)^2} \quad (52)$$

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} \Big|_{\mathbf{w}_{\text{opt}2}} = 2\mu_1 \frac{\delta_{11} E\{I_1^2(n)\} a_{21}^2}{\left(1 - \frac{1}{r_1 r_2}\right)^2} \quad (53)$$

and substituting $\underline{w}_{\text{opt}}$ from (18) and (19) in (51), we obtain respectively

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21}^2} \Big|_{\underline{w}_{\text{opt}1}} = 2\mu_2 \frac{\delta_{21} E\{I_1^2(n)\} a_{11}^2}{(1 - r_1 r_2)^2} \quad (54)$$

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21}^2} \Big|_{\underline{w}_{\text{opt}2}} = 2\mu_2 \frac{\delta_{22} E\{I_2^2(n)\} a_{12}^2}{\left(1 - \frac{1}{r_1 r_2}\right)^2} \quad (55)$$

Similarly by applying (49) to (13) and (14) with respect to w_{21} and w_{12} respectively

$$\begin{aligned} \mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{\underline{w}_{\text{opt}}} &= \frac{2\mu_1}{(1 - w_{12\text{opt}} w_{21\text{opt}})^3} \left[\delta_{11} E\{I_1^2(n)\} \right. \\ &\quad \left. \cdot (a_{11} + w_{12\text{opt}} a_{21}) a_{11} + \delta_{12} E\{I_2^2(n)\} (a_{12} + w_{12\text{opt}} a_{22}) a_{12} \right] \end{aligned} \quad (56)$$

$$\begin{aligned} \mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{\underline{w}_{\text{opt}}} &= \frac{2\mu_2}{(1 - w_{12\text{opt}} w_{21\text{opt}})^3} \left[\delta_{21} E\{I_1^2(n)\} \right. \\ &\quad \left. \cdot (a_{21} + w_{21\text{opt}} a_{11}) a_{21} + \delta_{22} E\{I_2^2(n)\} (a_{22} + w_{21\text{opt}} a_{12}) a_{22} \right]. \end{aligned} \quad (57)$$

Also, substituting (18) and (19) in (56), we obtain respectively

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{\underline{w}_{\text{opt}1}} = 2\mu_1 \frac{\delta_{11} E\{I_1^2(n)\} a_{11}^2}{(1 - r_1 r_2)^2} \quad (58)$$

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{\underline{w}_{\text{opt}2}} = 2\mu_1 \frac{\delta_{12} E\{I_2^2(n)\} a_{12}^2}{\left(1 - \frac{1}{r_1 r_2}\right)^2} \quad (59)$$

and using (18) and (19) in (57), we obtain respectively

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{\underline{w}_{\text{opt}1}} = 2\mu_2 \frac{\delta_{22} E\{I_2^2(n)\} a_{22}^2}{(1 - r_1 r_2)^2} \quad (60)$$

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{\underline{w}_{\text{opt}2}} = 2\mu_2 \frac{\delta_{21} E\{I_1^2(n)\} a_{21}^2}{\left(1 - \frac{1}{r_1 r_2}\right)^2} \quad (61)$$

Using (52) and (54) in (47), we can calculate $\underline{b}_{\text{opt}1}$ and by using these equations together with (58) and (60) in (48) we can calculate $\underline{c}_{\text{opt}1}$.

Let $\mu_1 = \mu_2 = \mu$, then for the first equilibrium point $(w_{12\text{opt}1}, w_{21\text{opt}1})$, we have.

$$\underline{b}_{\text{opt}1} = \frac{2\mu}{(1 - r_1 r_2)^2} \left[\delta_{21} a_{11}^2 E\{I_1^2(n)\} + \delta_{12} a_{22}^2 E\{I_2^2(n)\} \right] \quad (62)$$

$$\underline{c}_{\text{opt}1} = \frac{4\mu^2}{(1 - r_1 r_2)^4} a_{22}^2 a_{11}^2 E\{I_2^2(n)\} E\{I_1^2(n)\} (\delta_{21} \delta_{12} - \delta_{22} \delta_{11}). \quad (63)$$

For the second equilibrium point, by using (53) and (55) in (47), we calculate $b_{\text{opt}2}$ and by using these equations together with (59) and (61) in (48), we calculate $c_{\text{opt}2}$;

$$b_{\text{opt}2} = \frac{2\mu}{\left(1 - \frac{1}{r_1 r_2}\right)^2} \left[\delta_{22} a_{12}^2 E\{I_2^2(n)\} + \delta_{11} a_{21}^2 E\{I_1^2(n)\} \right] \quad (64)$$

$$c_{\text{opt}2} = \frac{-4\mu^2}{\left(1 - \frac{1}{r_1 r_2}\right)^4} a_{12}^2 a_{21}^2 E\{I_2^2(n)\} E\{I_1^2(n)\} (\delta_{21} \delta_{12} - \delta_{11} \delta_{22}). \quad (65)$$

3.2. Stability Parameters of Correlator-Correlator Scheme

For this scheme, $X_1(w_{12}, w_{21})$ and $X_2(w_{12}, w_{21})$ are respectively, $\mu_1 P_1(w_{12}, w_{21}) = \mu_1 A_1^2(w_{12}, w_{21})$ and $\mu_2 Q_1(w_{12}, w_{21}) = \mu_1 B_1^2(w_{12}, w_{21})$.

Now,

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} = 2\mu_1 \left[\left(\frac{\partial A_1(w_{12}, w_{21})}{\partial w_{12}} \right)^2 + A_1(w_{12}, w_{21}) \frac{\partial^2 A_1(w_{12}, w_{21})}{\partial w_{12}^2} \right]. \quad (66)$$

In section 2, we concluded that for the optimum weight, we must have

$A_1(w_{12}, w_{21})|_{w_{\text{opt}}} = 0$. Therefore, (66) results in

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} \Big|_{w_{\text{opt}}} = 2\mu_1 \left(\frac{\partial A_1(w_{12}, w_{21})}{\partial w_{12}} \Big|_{w_{\text{opt}}} \right)^2 \quad (67)$$

and

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{w_{\text{opt}}} = 2\mu_1 \frac{\partial A_1(w_{12}, w_{21})}{\partial w_{21}} \Big|_{w_{\text{opt}}} \frac{\partial A_1(w_{12}, w_{21})}{\partial w_{12}} \Big|_{w_{\text{opt}}}. \quad (68)$$

Similarly for $Q_1(w_{12}, w_{21})$

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21}^2} \Big|_{w_{\text{opt}}} = 2\mu_2 \left(\frac{\partial B_1(w_{12}, w_{21})}{\partial w_{21}} \right)^2 \Big|_{w_{\text{opt}}}. \quad (69)$$

and

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{w_{\text{opt}}} = 2\mu_2 \frac{\partial B_1(w_{12}, w_{21})}{\partial w_{21}} \Big|_{w_{\text{opt}}} \frac{\partial B_1(w_{12}, w_{21})}{\partial w_{12}} \Big|_{w_{\text{opt}}}. \quad (70)$$

Also using (24) and (25), we can write,

$$\frac{\partial A_1(w_{12}, w_{21})}{\partial w_{12}} = \delta_{11} E\{I_1(n)\} a_{21} (w_{21} a_{11} + a_{21})$$

$$+\delta_{12}E\{I_2(n)\}a_{22}(w_{21}a_{12} + a_{22}) \quad (71)$$

$$\begin{aligned} \frac{\partial A_1(w_{12}, w_{21})}{\partial w_{21}} &= \delta_{11}E\{I_1^2(n)\}a_{11}(w_{12}a_{21} + a_{11}) \\ &+ \delta_{12}E\{I_2^2(n)\}a_{12}(w_{12}a_{22} + a_{12}) \end{aligned} \quad (72)$$

$$\begin{aligned} \frac{\partial B_1(w_{12}, w_{21})}{\partial w_{12}} &= \delta_{21}E\{I_1^2(n)\}a_{21}(a_{21} + w_{21}a_{11}) \\ &+ \delta_{22}E\{I_2^2(n)\}a_{22}(a_{22} + w_{21}a_{12}) \end{aligned} \quad (73)$$

$$\begin{aligned} \frac{\partial B_1(w_{12}, w_{21})}{\partial w_{21}} &= \delta_{21}E\{I_1^2(n)\}a_{11}(w_{12}a_{21} + a_{11}) \\ &+ \delta_{22}E\{I_2^2(n)\}a_{12}(w_{12}a_{22} + a_{12}). \end{aligned} \quad (74)$$

Using (67) together with (71) and substituting (18), (19) we get respectively, for the two equilibrium points,

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} \Big|_{w_{\text{opt1}}} = 2\mu_1 \left[\delta_{12}E\{I_2^2(n)\}a_{22}^2(1 - r_1r_2) \right]^2 \quad (75)$$

$$\mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12}^2} \Big|_{w_{\text{opt2}}} = 2\mu_1 \left[\delta_{11}E\{I_1^2(n)\}a_{21}^2(1 - \frac{1}{r_1r_2}) \right]^2 \quad (76)$$

Similarly using (69) together with (72), we get

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21}^2} \Big|_{w_{\text{opt1}}} = 2\mu_2 \left[\delta_{21}E\{I_1^2(n)\}a_{11}^2(1 - r_1r_2) \right]^2 \quad (77)$$

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{21}^2} \Big|_{w_{\text{opt2}}} = 2\mu_2 \left[\delta_{22}E\{I_2^2(n)\}a_{12}^2(1 - \frac{1}{r_1r_2}) \right]^2 \quad (78)$$

Also, using (68) and (70), we obtain respectively, for the two equilibrium points.

$$\begin{aligned} \mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{w_{\text{opt1}}} &= 2\mu_1 [E\{I_1^2(n)\}E\{I_2^2(n)\}a_{11}^2a_{22}^2 \\ &\cdot (1 - r_1r_2)^2 \delta_{12}\delta_{11}] \end{aligned} \quad (79)$$

$$\begin{aligned} \mu_1 \frac{\partial^2 P_1(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{w_{\text{opt2}}} &= 2\mu_1 [E\{I_1^2(n)\}E\{I_2^2(n)\}a_{12}^2a_{21}^2 \\ &\cdot (1 - \frac{1}{r_1r_2})^2 \delta_{12}\delta_{11}] \end{aligned} \quad (80)$$

$$\begin{aligned} \mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{w_{\text{opt1}}} &= 2\mu_2 [E\{I_1^2(n)\}E\{I_2^2(n)\}a_{11}^2a_{22}^2 \\ &\cdot (1 - r_1r_2)^2 \delta_{21}\delta_{22}] \end{aligned} \quad (81)$$

$$\mu_2 \frac{\partial^2 Q_1(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{w_{\text{opt}2}} = 2\mu_2 [E\{I_1^2(n)\} E\{I_2^2(n)\} a_{12}^2 a_{21}^2 \cdot (1 - \frac{1}{r_1 r_2})^2 \delta_{21} \delta_{22}. \quad (82)$$

Finally, using (75) and (77) in (47), we calculate $b_{1\text{opt}1}$ and using these equations together with (79) and (81) in (48), we can calculate $c_{1\text{opt}1}$.

Let $\mu_1 = \mu_2 = \mu$, then for the first equilibrium point,

$$b_{1\text{opt}1} = 2\mu(1 - r_1 r_2)^2 \left[(\delta_{12} E\{I_2^2(n)\} a_{22}^2)^2 + (\delta_{21} E\{I_1^2(n)\} a_{11}^2)^2 \right] \quad (83)$$

$$c_{1\text{opt}1} = 4\mu^2 \left[E\{I_2^2(n)\} E\{I_1^2(n)\} a_{11}^2 a_{22}^2 (1 - r_1 r_2) \right]^2 \delta_{21} \delta_{12} (\delta_{12} \delta_{21} - \delta_{11} \delta_{22}). \quad (84)$$

Similarly, for the second equilibrium, using (76) and (78) in (47) and (80), (82) in (48), we can calculate $b_{1\text{opt}2}$ and $c_{1\text{opt}2}$, respectively;

$$b_{1\text{opt}2} = 2\mu \left(1 - \frac{a_{11} a_{22}}{a_{12} a_{21}}\right)^2 \left[(\delta_{11} E\{I_1^2(n)\} a_{21}^2)^2 + (\delta_{22} E\{I_2^2(n)\} a_{12}^2)^2 \right] \quad (85)$$

$$c_{1\text{opt}1} = -4\mu^2 \left[E\{I_2^2(n)\} E\{I_1^2(n)\} a_{12}^2 a_{21}^2 \left(1 - \frac{a_{11} a_{22}}{a_{12} a_{21}}\right) \right]^2 \delta_{22} \delta_{11} (\delta_{12} \delta_{21} - \delta_{11} \delta_{22}). \quad (86)$$

3.3. Stability Parameters of The Power-Correlator Scheme

For this scheme, $X_1(w_{12}, w_{21})$ and $X_2(w_{12}, w_{21})$ are given by $\mu_1 P_2(w_{12}, w_{21})$ from (33) and $\mu_2 Q_2(w_{12}, w_{21}) = \mu_1 B_2^2(w_{12}, w_{21})$ from (34), respectively.

From (35)

$$\mu_1 \frac{\partial^2 P_2(w_{12}, w_{21})}{\partial w_{12}^2} = 2\mu_1 \left[\delta_{11} E\{I_1^2(n)\} (a_{21} + w_{21\text{opt}} a_{11})^2 + \delta_{12} E\{I_2^2(n)\} (a_{22} + w_{21\text{opt}} a_{12})^2 \right]. \quad (87)$$

and from (36)

$$\mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{21}^2} = 2\mu_2 \left[\left(\frac{\partial B_2(w_{12}, w_{21})}{\partial w_{21}} \right)^2 + B_2(w_{12}, w_{21}) \frac{\partial^2 B_2(w_{12}, w_{21})}{\partial w_{21}^2} \right]. \quad (88)$$

But in section 2, we concluded that at the optimum weights, we must have

$B_2(w_{12}, w_{21})|_{\mathbf{w}_{\text{opt}}} = 0$. Therefore, (88) becomes;

$$\mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{21}^2} |_{\mathbf{w}_{\text{opt}1}} = 2\mu_2 \left(\frac{\partial B_2(w_{12}, w_{21})}{\partial w_{21}} |_{\mathbf{w}_{\text{opt}1}} \right)^2. \quad (89)$$

From (34) we find that

$$\begin{aligned} \frac{\partial B_2(w_{12}, w_{21})}{\partial w_{21}} &= \delta_{21} E\{I_1^2(n)\} [a_{11}^2 + w_{12} a_{11} (a_{12} + w_{21} a_{11})] \\ &\quad \delta_{22} E\{I_2^2(n)\} [a_{12}^2 + w_{12} a_{12} (a_{22} + w_{21} a_{12})] \end{aligned} \quad (90)$$

$$\frac{\partial B_2(w_{12}, w_{21})}{\partial w_{12}} = \delta_{21} E\{I_1^2(n)\} (a_{21} + w_{21} a_{11})^2 + \delta_{22} E\{I_2^2(n)\} (a_{22} + w_{21} a_{12})^2. \quad (91)$$

Substituting (38) and (39) in (87), we get, for the two equilibrium points,

$$\mu_1 \frac{\partial^2 P_2(w_{12}, w_{21})}{\partial w_{12}^2} |_{\mathbf{w}_{\text{opt}1}} = 2\mu_1 \delta_{12} E\{I_2^2(n)\} a_{22}^2 (1 - r_1 r_2)^2 \quad (92)$$

$$\mu_1 \frac{\partial^2 P_2(w_{12}, w_{21})}{\partial w_{12}^2} |_{\mathbf{w}_{\text{opt}2}} = 2\mu_1 \delta_{11} E\{I_1^2(n)\} a_{21}^2 \left(1 - \frac{1}{r_1 r_2}\right)^2. \quad (93)$$

Similarly (90) in (89) together and using (38) and (39), we have for the two equilibrium points,

$$\mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{21}^2} |_{\mathbf{w}_{\text{opt}1}} = 2\mu_2 [\delta_{21} E\{I_1^2(n)\} a_{11}^2 - \delta_{22} E\{I_2^2(n)\} a_{12}^2]^2 \quad (94)$$

$$\mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{21}^2} |_{\mathbf{w}_{\text{opt}2}} = 2\mu_2 [-\delta_{21} E\{I_1^2(n)\} a_{11}^2 + \delta_{22} E\{I_2^2(n)\} a_{12}^2]^2. \quad (95)$$

Also taking the derivative of (35) with respect to w_{21} , we get

$$\begin{aligned} \mu_1 \frac{\partial^2 P_2(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} &= 2\mu_1 \left[\delta_{11} E\{I_1^2(n)\} [a_{11}^2 + 2w_{12} (a_{21} + w_{21} a_{11}) a_{11}] \right. \\ &\quad \left. + \delta_{12} E\{I_2^2(n)\} [a_{12}^2 + 2w_{12} (a_{22} + w_{21} a_{12}) a_{12}] \right]. \end{aligned} \quad (96)$$

Substituting (38), (39) in (96), we get respectively,

$$\mu_1 \frac{\partial^2 P_2(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} |_{\mathbf{w}_{\text{opt}1}} = 2\mu_1 [\delta_{11} E\{I_1^2(n)\} a_{11}^2 - \delta_{12} E\{I_2^2(n)\} a_{12}^2] \quad (97)$$

$$\mu_1 \frac{\partial^2 P_2(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} |_{\mathbf{w}_{\text{opt}2}} = -2\mu_1 [\delta_{11} E\{I_1^2(n)\} a_{11}^2 - \delta_{12} E\{I_2^2(n)\} a_{12}^2]. \quad (98)$$

From the definition of $Q_2(w_{12}, w_{21})$ and (36), we have,

$$\mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} = 2\mu_2 \left[\frac{\partial B_2(w_{12}, w_{21})}{\partial w_{21}} \frac{\partial B_2(w_{12}, w_{21})}{\partial w_{12}} + B_2(w_{12}, w_{21}) \frac{\partial B_2(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \right]. \quad (99)$$

But at equilibrium point, $B_2(w_{12}, w_{21})$ is equal to zero and we get,

$$\mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{21} \partial w_{12}} \Big|_{w_{\text{opt}1}} = 2\mu_2 \frac{\partial B_2(w_{12}, w_{21})}{\partial w_{21}} \Big|_{w_{\text{opt}1}} \frac{\partial B_2(w_{12}, w_{21})}{\partial w_{12}} \Big|_{w_{\text{opt}1}}. \quad (100)$$

Using (90) and (91) in (100) together with (38) and (39), we get respectively, for the two equilibrium points

$$\begin{aligned} \mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{w_{\text{opt}1}} &= 2\mu_2 [\delta_{21} E\{I_1^2(n)\} a_{11}^2 - \delta_{22} E\{I_2^2(n)\} a_{12}^2] \\ &\quad \cdot \delta_{22} E\{I_2^2(n)\} a_{22}^2 (1 - r_1 r_2)^2 \end{aligned} \quad (101)$$

$$\begin{aligned} \mu_2 \frac{\partial^2 Q_2(w_{12}, w_{21})}{\partial w_{12} \partial w_{21}} \Big|_{w_{\text{opt}2}} &= -2\mu_2 [\delta_{21} E\{I_1^2(n)\} a_{11}^2 - \delta_{22} E\{I_2^2(n)\} a_{12}^2] \\ &\quad \cdot \delta_{21} E\{I_1^2(n)\} a_{21}^2 \left(1 - \frac{1}{r_1 r_2}\right)^2. \end{aligned} \quad (102)$$

Finally using (92), (94) in (47), together with the definition of X_1 and X_2 for this scheme, we calculate $b_{2\text{opt}1}$ and by using these equations together with (97) and (101) in (48), we calculate $c_{2\text{opt}1}$.

Let $\mu_1 = \mu_2 = \mu$, then

$$\begin{aligned} b_{2\text{opt}1} &= 2\mu \left[\delta_{12} a_{22}^2 E\{I_2^2(n)\} (1 - r_1 r_2)^2 + \left[\delta_{21} E\{I_1^2(n)\} a_{11}^2 - \delta_{22} E\{I_2^2(n)\} a_{12}^2 \right]^2 \right] \\ c_{2\text{opt}1} &= 4\mu^2 E\{I_2^2(n)\} E\{I_1^2(n)\} a_{11}^2 a_{22}^2 (1 - r_1 r_2)^2 \\ &\quad \cdot \left[\delta_{21} E\{I_1^2(n)\} a_{11}^2 - \delta_{22} E\{I_2^2(n)\} a_{12}^2 \right] (\delta_{21} \delta_{12} - \delta_{11} \delta_{22}). \end{aligned} \quad (103)$$

For a second equilibrium point of (39) we calculate $b_{2\text{opt}2}$ by using (93) and (95) in (47) and by using these equations together with (98) and (102) in (48), we calculate $c_{2\text{opt}1}$. If $\mu_1 = \mu_2 = \mu$, then;

$$\begin{aligned} b_{2\text{opt}2} &= 2\mu \left[\delta_{11} a_{21}^2 E\{I_1^2(n)\} \left(1 - \frac{a_{22} a_{11}}{a_{21} a_{12}}\right)^2 + \left[-\delta_{21} E\{I_1^2(n)\} a_{11}^2 + \delta_{22} E\{I_2^2(n)\} a_{12}^2 \right]^2 \right] \\ c_{2\text{opt}2} &= 4\mu^2 E\{I_2^2(n)\} E\{I_1^2(n)\} a_{21}^2 a_{12}^2 \left(1 - \frac{a_{22} a_{11}}{a_{21} a_{12}}\right)^2 \\ &\quad \left[\delta_{21} E\{I_1^2(n)\} a_{11}^2 - \delta_{22} E\{I_2^2(n)\} a_{12}^2 \right] (\delta_{21} \delta_{12} - \delta_{11} \delta_{22}). \end{aligned} \quad (104)$$

IV. STABILITY CONDITIONS

From the characteristics equation in (46), the two eigenvalues λ_1 and λ_2 are related to b and c as they are defined at the two equilibrium points;

The stability condition can be summarized as follows,

- If $c < 0$, the eigenvalues of (46), λ_1 and λ_2 are both real and $\lambda_1\lambda_2 < 0$. Since one of the eigenvalue is positive, then the equilibrium is unstable.
- If $c > 0$, the two eigenvalues are either both real or complex-conjugate pair. and $\lambda_1\lambda_2 > 0$. Both eigenvalues (if they are real) or their real part (if they are complex) are negative if $b < 0$, positive if $b > 0$. Therefore, the system is stable if $c > 0$ and $b < 0$ and unstable if $c > 0$ and $b > 0$.
- If $c = 0$, then one of the eigenvalues is zero and the other one is equal to b and the system is unstable [2].

4.1 Stability Conditions for The Power-Power Scheme

The equilibrium points for this scheme are given by (18) and (19). First, for c at the first equilibrium point to be positive, we must have from (63); $\delta_{21}\delta_{12} > \delta_{11}\delta_{22}$ and for b to be negative, we must have from (62) $\mu < 0$. These two conditions will result in convergence of the algorithm at point (18). However, from (64) and (65) these conditions will result in divergence at the second equilibrium point in (19) (saddle point).

4.2 Stability Conditions for The Correlator-Correlator Scheme

The equilibrium for this scheme is the same as the previous scheme and given by (18) and (19). Again at the first equilibrium point it is convergent if $\delta_{21}\delta_{12} > \delta_{11}\delta_{22}$ and μ_1, μ_2 or μ are negative. This results from the signs of c_1 and b_1 given in (83) and (84), respectively. The same condition leads to divergence as a result of the signs of c_1 and b_1 given in (85) and (86).

4.3 Stability Conditions for Power-Correlator Scheme

The equilibrium points for this scheme are given by (38) and (39). For c_2 at the first equilibrium point in (103) to be positive, we must have from (38) that;

1. $\delta_{21}\delta_{12} > \delta_{11}\delta_{22}$, and
2. $\frac{\delta_{21}}{\delta_{22}} > \frac{E\{I_2^2(n)\}}{E\{I_1^2(n)\}} \left(\frac{a_{12}}{a_{11}}\right)^2$,

and for $b_2 < 0$ to be negative, we must further have μ_1 , μ_2 or μ be less than zero.

For the second equilibrium point from (104), the same conditions lead to divergence.

V. REFERENCES

- [1] Y. Bar-ness, A. Dinç and H. Messer-Yaron, *The Bootstrapped Algorithm: Fast Algorithm for Blind Signal Separation*, Final Technical Report, RL-TR-93-24, April, 1993.
- [2] Blaguiera, A. *Nonlinear System Analysis* Academic Press, 1966.

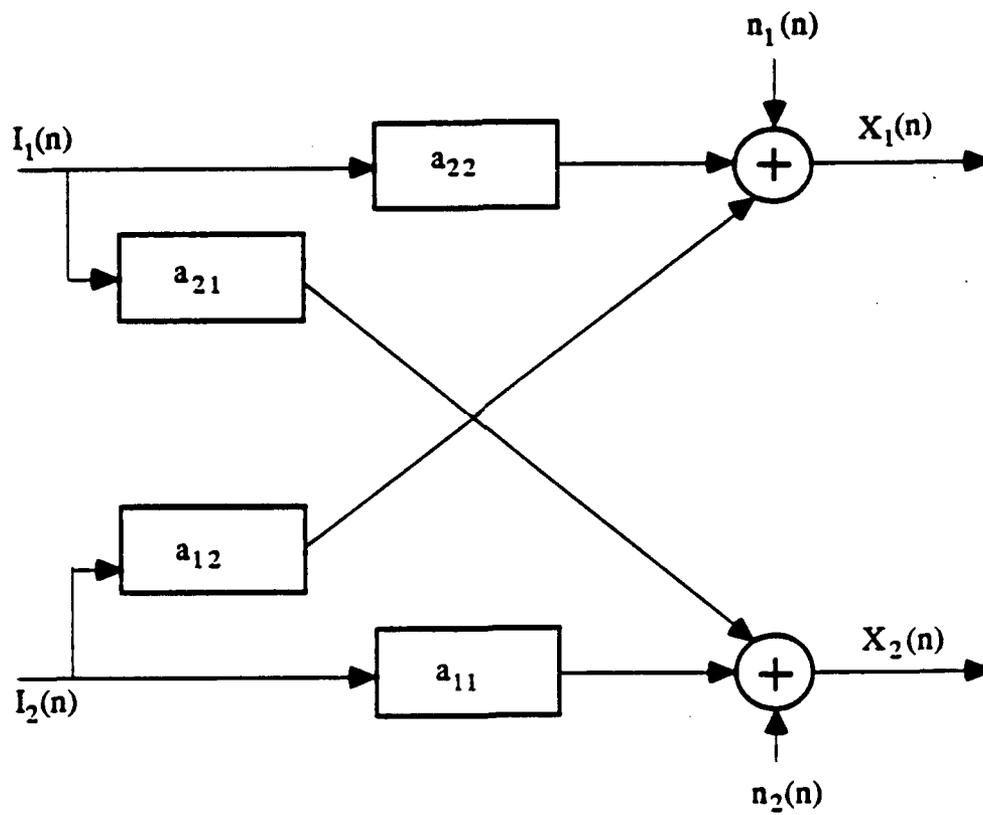


Figure 1 No noise dually polarized channel Model.

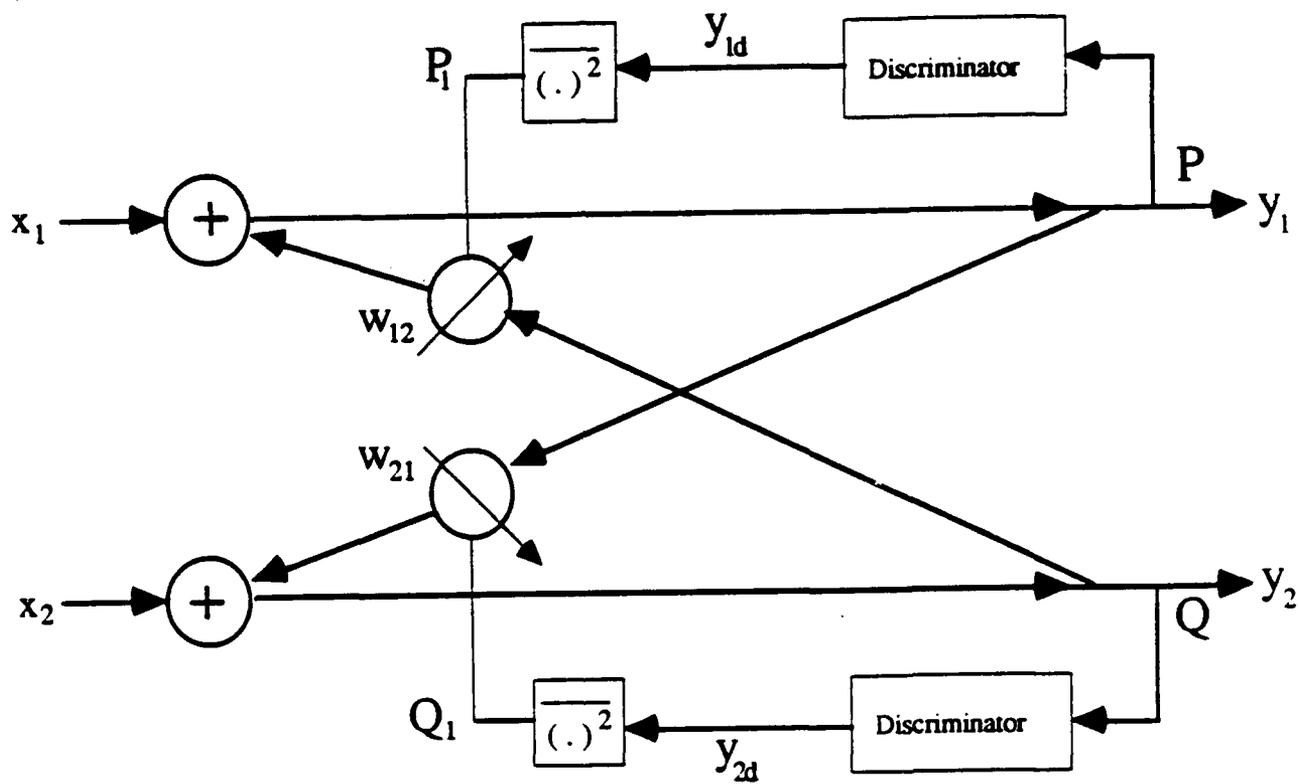


Figure 2 The Power-Power Cross-Pol Interference Canceller.

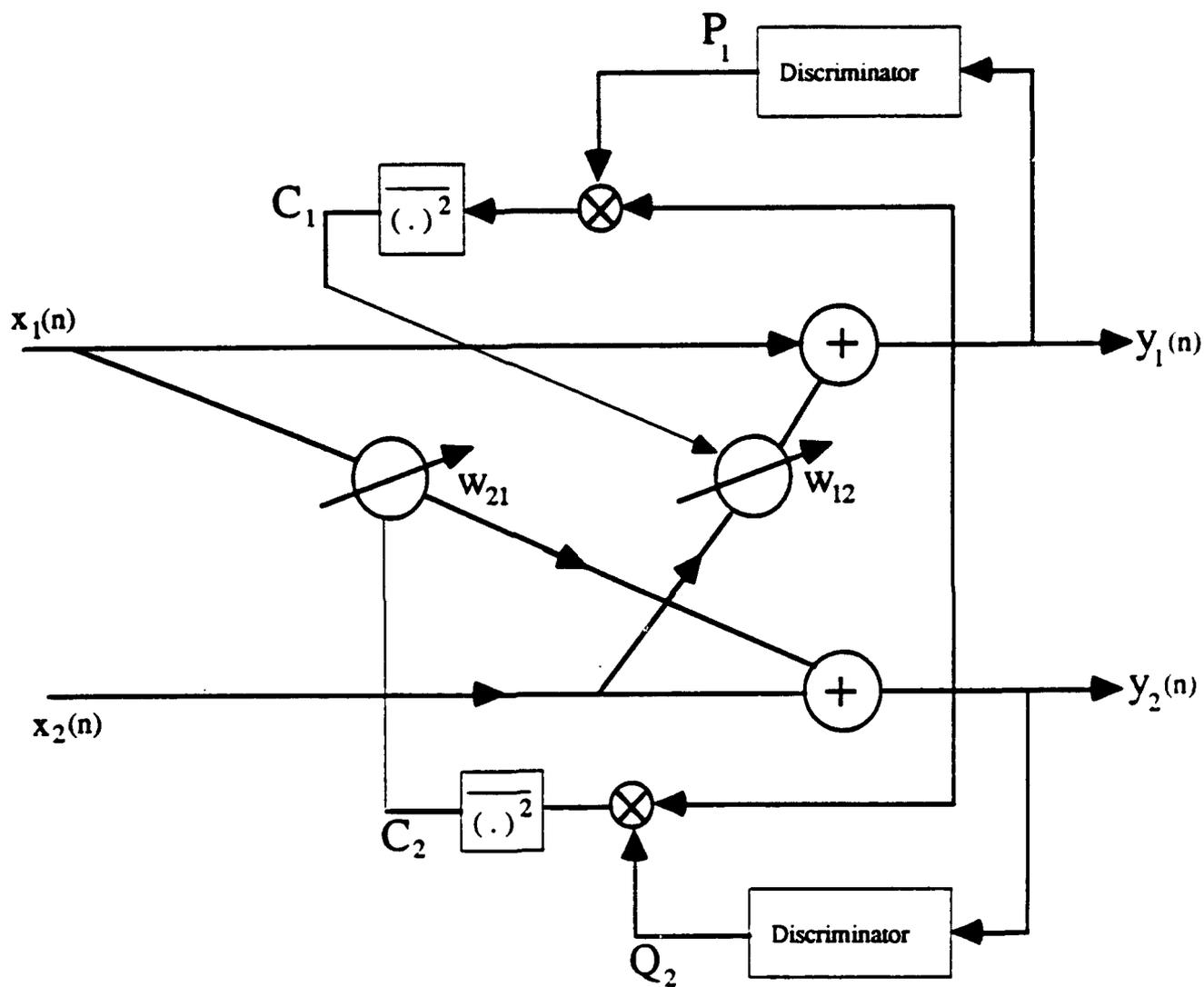


Figure 3 The Correlator-Correlator Cross-Pol Interference Canceler.

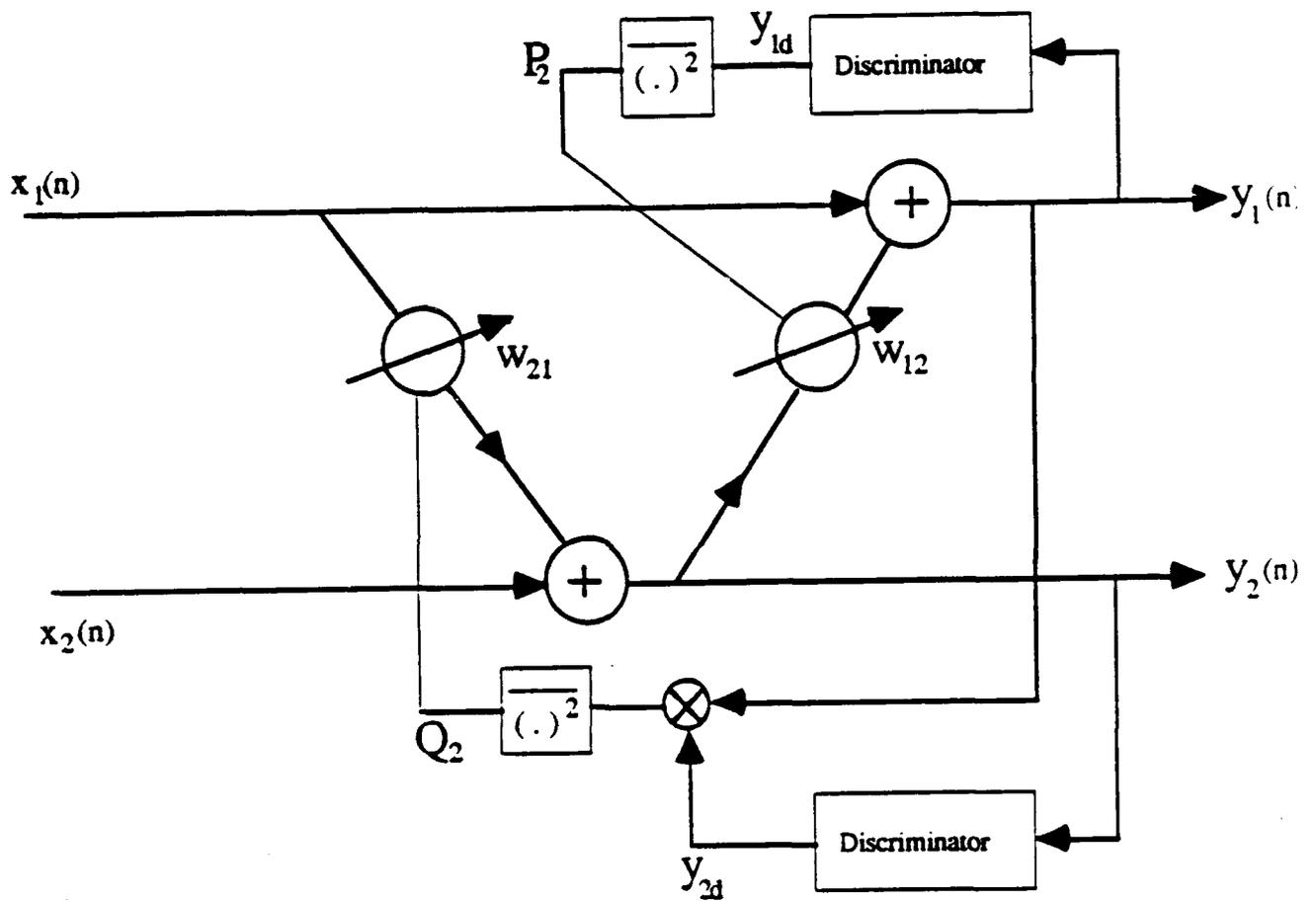


Figure 4 The Power-Correlator Cross-Pol Interference Canceler.

APPENDIX C

BOOTSTRAPPED ADAPTIVE SEPARATION OF TWO SUPERIMPOSED SIGNALS - DYNAMIC STUDY OF POWER-POWER SCHEME USING ORTHOGONAL PERTURBATION SEQUENCES

by

Abdulkadir Dinç and Yeheskel Bar-Ness

I. INTRODUCTIONS

In a previous report [1], we studied the steady state behaviour of the different configurations of the bootstrapped schemes. In each scheme, the computation of the optimal weights require the knowledge of the channel model as well as the signal and the noise powers.

Alternative procedures to find these optimal weights are to use recursive relations. These recursive algorithms were presented in [1] Appendix A for power-power, correlator-correlator and for power-correlator schemes. All these recursive procedures require knowledge of the gradient of square of the output powers or the gradient of correlation between the outputs.

In this report, we will present a dynamic analysis of the power-power scheme. That is a backward/backward structure whose weight is controlled via minimizing the output powers, respectively. We will present a technique for reaching the optimal weights with a recursive weight-updating procedure using estimates of the gradients. The estimate of the gradients will be obtained by applying two orthogonal perturbation sequences to the two weights simultaneously, and measuring the corresponding changes at the two output powers P and Q of the power-power scheme.

II. THE SEPARATOR'S SYSTEM AND PARAMETERS

2.1 The Channel

The received signals sampled after matched filters, are taken to be;

$$\begin{aligned}x_1(n) &= a_{11}I_1(n) + a_{12}I_2(n) + n_1(n) \\x_2(n) &= a_{21}I_1(n) + a_{22}I_2(n) + n_2(n)\end{aligned}\quad (1)$$

where $x_1(n)$ and $x_2(n)$ are the sampled received signals at the first and second channels respectively. $I_i(n)$ and $n_i(n)$ are the corresponding signals and noises at these outputs such that I_i , $i = 1, 2$ are M-ary signals from set $\{\pm 1, \pm 3, \pm 5, \dots, \pm(\sqrt{M} - 1)\}$ and a_{ij} , $i, j = 1, 2$ are real channel couplings. $n_1(n)$ and $n_2(n)$ are assumed independent samples of zero mean Gaussian with;

$$E\{n_i^2(n)\} = \sigma_i^2, \quad i = 1, 2. \quad (2)$$

2.2 The Canceler Outputs

From Fig. 1, the outputs $y_1(n)$ and $y_2(n)$ are as follows

$$\begin{aligned}y_1(n) &= \frac{x_1(n) + x_2(n)w_{12}(n)}{1 - w_{12}(n)w_{21}(n)} \\y_2(n) &= \frac{x_2(n) + x_1(n)w_{21}(n)}{1 - w_{12}(n)w_{21}(n)}.\end{aligned}\quad (3)$$

Substituting for $x_1(n)$ and $x_2(n)$ from (1) we get the outputs after the discrimination at the n th instant of time, respectively,

$$y_{1d}(n) = \frac{\sqrt{\delta_{11}}I_1(n)(a_{11} + w_{12}(n)a_{21}) + \sqrt{\delta_{12}}I_2(n)(a_{12} + w_{12}(n)a_{22}) + n_1(n) + n_2(n)w_{12}(n)}{1 - w_{12}(n)w_{21}(n)} \quad (4)$$

$$y_{2d}(n) = \frac{\sqrt{\delta_{21}}I_1(n)(a_{21} + w_{21}(n)a_{11}) + \sqrt{\delta_{22}}I_2(n)(a_{22} + w_{21}(n)a_{12}) + n_1(n)w_{21}(n) + n_2(n)}{1 - w_{12}(n)w_{21}(n)} \quad (5)$$

where δ_{ij} , $i, j=1,2$ is the effect of the discriminator at output i on the power of signal j .

Also, from Fig. 1, we can write the weights $w_{12}(n)$ and $w_{21}(n)$ as a sum of a nominal value $w_{12}(i)$ and $w_{21}(i)$ plus perturbation sequences $p_1(n)$ and $p_2(n)$ whose magnitudes are Λ . That is

$$w_{12}(n) = w_{12}(i) + \Lambda p_1(n) \quad (6)$$

$$w_{21}(n) = w_{21}(i) + \Lambda p_2(n). \quad (7)$$

2.3 Approximate Canceler Outputs

In ref([1] Appendix A), we found that, in the no noise environment the optimal weights are; $w_{21\text{opt}} = -\frac{a_{21}}{a_{11}}$ and $w_{12\text{opt}} = -\frac{a_{12}}{a_{22}}$.

Given that the cross-couplings $\frac{a_{21}}{a_{11}} \frac{a_{12}}{a_{22}} \ll 1$, then.

$$w_{12}(n)w_{21}(n) \ll 1. \quad (8)$$

The assumption in (8) will be used to simplify the analysis of dynamic study. With this approximation, we can write (4) and (5) as

$$y_{1d}(n) \approx \left[\sqrt{\delta_{11}} I_1(n) [a_{11} + w_{12}(n)a_{21}] + \sqrt{\delta_{12}} I_2(n) [a_{12} + w_{12}(n)a_{22}] \right. \\ \left. + n_1(n) + n_2(n)w_{12}(n) \right] [1 + w_{12}(n)w_{21}(n)] \quad (9)$$

$$y_{2d}(n) \approx \left[\sqrt{\delta_{21}} I_1(n) [a_{21} + w_{21}(n)a_{11}] + \sqrt{\delta_{22}} I_2(n) [a_{22} + w_{21}(n)a_{12}] \right. \\ \left. + n_2(n) + n_1(n)w_{21}(n) \right] [1 + w_{12}(n)w_{21}(n)]. \quad (10)$$

2.4 Mean Output Powers

In the steady state, using (9) and (10) respectively, we can write the mean output powers with weights fixed at w_{12} and w_{21} ; $P(w_{12}, w_{21}) = E\{y_{1d}^2(n)\}$ and $Q(w_{12}, w_{21}) = E\{y_{2d}^2(n)\}$, respectively,

$$P(w_{12}, w_{21}) \approx \left[\delta_{11} E\{I_1^2(n)\} (a_{11} + w_{12}a_{21})^2 + \delta_{12} E\{I_2^2(n)\} (a_{12} + w_{12}a_{22})^2 \right. \\ \left. + E\{n_1^2(n)\} + E\{n_2^2(n)\} w_{12}^2 \right] (1 + w_{12}w_{21})^2. \quad (11)$$

$$Q(w_{12}, w_{21}) \approx \left[\delta_{21} E\{I_1^2(n)\} (a_{21} + w_{21}a_{11})^2 + \delta_{22} E\{I_2^2(n)\} (a_{22} + w_{21}a_{12})^2 \right. \\ \left. + E\{n_2^2(n)\} + E\{n_1^2(n)\} w_{21}^2 \right] (1 + w_{12}w_{21})^2 \quad (12)$$

where $E\{(\cdot)\}$ denotes the expected value of (\cdot) .

2.5 Optimum Weights

The optimal weight vector; $\mathbf{w}_{\text{opt}} = [w_{12\text{opt}}, w_{21\text{opt}}]^T$ which minimizes the mean output powers P and Q are found by taking the derivative of P and Q with respect to w_{12} and w_{21} and equating the result to zero, respectively.

From (11) and (12), we get respectively,

$$\begin{aligned} \frac{\partial P(w_{12}, w_{21})}{\partial w_{12}} = & 2 \left[\delta_{11} E\{I_1^2(n)\} \left((1 + w_{12}w_{21})^2 (a_{11} + w_{12}a_{21}) a_{21} \right. \right. \\ & \left. \left. + (1 + w_{12}w_{21})(a_{11} + w_{12}a_{21})^2 w_{21} \right) \right. \\ & \left. + \delta_{12} E\{I_2^2(n)\} \left((1 + w_{12}w_{21})^2 (a_{12} + w_{12}a_{22}) a_{22} \right. \right. \\ & \left. \left. + (1 + w_{12}w_{21})(a_{12} + w_{12}a_{22})^2 w_{21} \right) + E\{n_1^2(n)\} (1 + w_{12}w_{21}) w_{21} \right. \\ & \left. + E\{n_2^2(n)\} \left((1 + w_{12}w_{21})^2 w_{12} + (1 + w_{12}w_{21}) w_{12}^2 w_{21} \right) \right]. \quad (13) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial Q(w_{12}, w_{21})}{\partial w_{21}} = & 2 \left[\delta_{21} E\{I_1^2(n)\} \left((1 + w_{12}w_{21})^2 (a_{21} + w_{21}a_{11}) a_{11} \right. \right. \\ & \left. \left. + (1 + w_{12}w_{21})(a_{21} + w_{21}a_{11})^2 w_{12} \right) \right. \\ & \left. + \delta_{22} E\{I_2^2(n)\} \left((1 + w_{12}w_{21})^2 (a_{22} + w_{21}a_{12}) a_{12} \right. \right. \\ & \left. \left. + (1 + w_{12}w_{21})(a_{22} + w_{21}a_{12})^2 w_{12} \right) + E\{n_1^2(n)\} (1 + w_{12}w_{21}) w_{12} \right. \\ & \left. + E\{n_2^2(n)\} \left((1 + w_{12}w_{21})^2 w_{21} + (1 + w_{12}w_{21}) w_{21}^2 w_{12} \right) \right]. \quad (14) \end{aligned}$$

III. GRADIENT DESCENT

An alternative way to get to the optimum weights is to use gradient descent method. With this technique, the weight vector at time $i+1$ is computed by using the gradients of (13) and (14) according to the following recursive relation,

$$w_{12}^{i+1} = w_{12}^i - \mu_1 \frac{\partial}{\partial w_{12}^i} E\{y_{1d}^i(n)^2\} \quad (15)$$

$$w_{21}^{i+1} = w_{21}^i - \mu_2 \frac{\partial}{\partial w_{21}^i} E\{y_{2d}^i(n)^2\}. \quad (16)$$

3.1 Gradient Estimation Using Orthogonal Perturbation Sequences

The use of random search and weight perturbation techniques for gradient estimation in adaptive systems have been reported by many authors ([2], [3]). In this report, we will follow Cantoni's gradient estimation definition [3] in which the estimate of the gradient is obtained by perturbing the weights (different sequence for each weight) simultaneously with different orthogonal perturbation sequences and correlating the outputs with the same sequences.

That is, the estimates of the gradients of the outputs are obtained from

$$g_l(i) = \frac{1}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} y_{ld}^2(m) p_l(m) \quad l = 1, 2 \quad (17)$$

where $y_{1d}^2(m)$ and $y_{2d}^2(m)$ are the instantaneous powers of the outputs after discriminator; $g_1(i)$ and $g_2(i)$ are taken as the estimates of the true gradients given in (13) and (14), respectively.

The perturbation sequences $p_l(m)$ are each taken to be periodic with period N and size Λ which assumed a positive real constant and are much smaller than unity. Different sequences are mutually orthogonal and zero mean over N cycle. For normalization, we divide by the sequence period N to obtain unit average power. Therefore, such sequences will satisfy the following,

$$\frac{1}{N} \sum_{n=1}^N p_k(n) p_l(n) = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \quad (18)$$

$$\frac{1}{N} \sum_{n=1}^N p_l(n) = 0 \quad l = 1, 2, \dots \quad (19)$$

and for higher moments

$$\frac{1}{N} \sum_{n=1}^N p_k^x(n) p_l^y(n) = \begin{cases} 1 & k = l \text{ and if both } \nu \text{ and } x \text{ are even} \\ 0 & k \neq l \text{ or if at least one of them is odd.} \end{cases} \quad (20)$$

It is possible to find such a family of sequences that yield an unbiased gradient estimate (for example, the rows of Hadamard matrix [3]).

For updating the weight vectors we use the estimated gradients $g_1(i)$ and $g_2(i)$ as follows

$$w_{12}(i+1) = w_{12}(i) - \mu g_1(i) \quad (21)$$

$$w_{21}(i+1) = w_{21}(i) - \mu g_2(i) \quad i = 1, 2, \dots \quad (22)$$

and after each recursion step the weights are operated on by

$$w_{12}(i+1) = \text{clip}(w_{12}(i+1)) \quad (23)$$

$$w_{21}(i+1) = \text{clip}(w_{21}(i+1)) \quad (24)$$

where, $\text{clip}(\cdot)$ is a clipping function such that

$$\text{clip}(w) = \begin{cases} w & |w| \leq \alpha \\ \frac{w\alpha}{|w|} & |w| > \alpha \end{cases} \quad (25)$$

and μ is the constant which determines the stability of convergence. The clipping operation ensures that each weight is bounded by a constant α to be less than one so that the desired equilibrium point, is reached.

3.2. Gradient Estimates, $g_1(i)$ and $g_2(i)$ for the Power-Power Separator

From (17), we write the expected value of gradient estimates conditioned on the weights.

$$E\{g_l(i)|\mathbf{w}(i)\} = \frac{1}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} E\{y_{ld}^2(m)|\mathbf{w}(i)\} p_l(m) \quad l = 1, 2. \quad (26)$$

Substituting (9) in (26), and taking the expectation of both sides conditioned on the weight vector $\mathbf{w}(i)$, we get

$$E\{g_1(i)|\mathbf{w}(i)\} = \frac{1}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} \left[\delta_{11} E\{I_1^2(m)\} (a_{11} + w_{12}(m)a_{21})^2 \right]$$

$$\begin{aligned}
& +\delta_{12}E\{I_2^2(m)\}(a_{12} + w_{12}(m)a_{22})^2 + E\{n_1(m)^2\} \\
& + E\{n_2^2(m)\}w_{12}^2(m)\left[1 + w_{12}(m)w_{21}(m)\right]^2 p_1(m) \quad (27)
\end{aligned}$$

and

$$\begin{aligned}
E\{g_2(i)|\mathbf{w}(i)\} & = \frac{1}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} \left[\delta_{21}E\{I_1^2(m)\}(a_{21} + w_{21}(m)a_{11})^2 \right. \\
& + \delta_{22}E\{I_2^2(m)\}(a_{22} + w_{21}(m)a_{12})^2 + E\{n_2(m)^2\} \\
& \left. + E\{n_1^2(m)\}w_{21}^2(m)\right] \left[1 + w_{12}(m)w_{21}(m)\right]^2 p_2(m) \quad (28)
\end{aligned}$$

where $w_{12}(m)$ and $w_{21}(m)$ are the perturbed weights and related to the nominal values $w_{12}(i)$ and $w_{21}(i)$ by equation (6) and (7). Using these relations and the orthogonal properties of the perturbation sequence stated in (18) to (20), we find in the of this report,

$$\begin{aligned}
E\{g_1(i)|\mathbf{w}(i)\} & = 2 \left[\delta_{11}E\{I_1^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 [a_{11} + w_{12}(i)a_{21}]a_{21} \right. \right. \\
& \quad \left. \left. + [1 + w_{12}(i)w_{21}(i)][a_{11} + w_{12}(i)a_{21}(i)]^2 w_{21}(i) \right) \right. \\
& \quad \left. + \delta_{12}E\{I_2^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 [a_{12} + w_{12}(i)a_{22}]a_{22} \right. \right. \\
& \quad \left. \left. + [1 + w_{12}(i)w_{21}(i)][a_{12} + w_{12}(i)a_{22}]^2 w_{21}(i) \right) \right. \\
& \quad \left. + E\{n_1^2(i)\} [1 + w_{12}(i)w_{21}(i)]w_{21}(i) \right. \\
& \quad \left. + E\{n_2^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 w_{12}(i) \right. \right. \\
& \quad \left. \left. + [1 + w_{12}(i)w_{21}(i)]w_{12}^2(i)w_{21}(i) \right) \right]. \quad (29)
\end{aligned}$$

From a similar evaluation as well as exploiting symmetry properties between $y_{1d}(m)$ and $y_{2d}(m)$, we get for the expected value of gradient estimator $g_2(i)$ conditioned on $\mathbf{w}(i)$:

$$\begin{aligned}
E\{g_2(i)|\mathbf{w}(i)\} & = 2 \left[\delta_{21}E\{I_1^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 [a_{21} + w_{21}(i)a_{11}]a_{11} \right. \right. \\
& \quad \left. \left. + [1 + w_{12}(i)w_{21}(i)][a_{21} + w_{21}(i)a_{11}]^2 w_{12}(i) \right) \right. \\
& \quad \left. + \delta_{22}E\{I_2^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 [a_{22} + w_{21}(i)a_{12}]a_{12} \right. \right. \\
& \quad \left. \left. + [1 + w_{12}(i)w_{21}(i)][a_{22} + w_{21}(i)a_{12}]^2 w_{12}(i) \right) \right]
\end{aligned}$$

$$\begin{aligned}
& + E\{n_1^2(i)\}[1 + w_{12}(i)w_{21}(i)]w_{12}(i) \\
& + E\{n_2^2(i)\}\left([1 + w_{12}(i)w_{21}(i)]^2 w_{21}(i)\right. \\
& \left. + [1 + w_{12}(i)w_{21}(i)]w_{21}^2(i)w_{12}(i)\right) \Big] \quad (30)
\end{aligned}$$

IV. CONVERGENCE IN THE MEAN

In this section, we will investigate the convergence in the mean of the recursive relation in (21) (22) to the optimum weights. We will also find an upper bound for the step size μ with which such convergence in the mean will occur.

4.1 The Error's Mean

Convergence in the mean of weight vector; $\mathbf{w}(i) = [w_{12}(i), w_{21}(i)]^T$ to optimum $\mathbf{w}_{\text{opt}} = [w_{12\text{opt}}, w_{21\text{opt}}]^T$, means

$$\lim_{i \rightarrow \infty} E\{\mathbf{w}(i)\} = \mathbf{w}_{\text{opt}}. \quad (31)$$

We first define a weight error vector at time i as $\mathbf{e}(i) = [e_1(i), e_2(i)]^T$, with

$$e_1(i) = w_{12}(i) - w_{12\text{opt}} \quad (32)$$

$$e_2(i) = w_{21}(i) - w_{21\text{opt}}. \quad (33)$$

We know that at the optimum weights, the gradients of the mean output powers P and Q are equal to zero, i.e

$$\frac{\partial P(w_{12}, w_{21})}{\partial w_{12}} \Big|_{\mathbf{w}_{\text{opt}}} = 0 \quad (34)$$

$$\frac{\partial Q(w_{12}, w_{21})}{\partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} = 0. \quad (35)$$

Subtracting the optimum weights $w_{12\text{opt}}$ and $w_{21\text{opt}}$ from both sides of (21) and (22), respectively and the gradients in (34) and (35) from the estimate of the gradients $g_1(i)$ and $g_2(i)$ in these equation respectively, we can write,

$$w_{12}(i+1) - w_{12\text{opt}} = w_{12}(i) - w_{12\text{opt}} - \mu \left[g_1(i) - \frac{\partial P(w_{12}, w_{21})}{\partial w_{12}} \Big|_{\mathbf{w}_{\text{opt}}} \right] \quad (36)$$

$$w_{21}(i+1) - w_{21\text{opt}} = w_{21}(i) - w_{21\text{opt}} - \mu \left[g_2(i) - \frac{\partial Q(w_{12}, w_{21})}{\partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} \right]. \quad (37)$$

By using the definition in (32) and (33) in (36) and (37), respectively, we write.

$$e_1(i+1) = e_1(i) - \mu \left[g_1(i) - \frac{\partial P(w_{12}, w_{21})}{\partial w_{12}} \Big|_{\mathbf{w}_{\text{opt}}} \right] \quad (38)$$

$$e_2(i+1) = e_2(i) - \mu \left[g_2(i) - \frac{\partial Q(w_{12}, w_{21})}{\partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} \right]. \quad (39)$$

In order to investigate the convergence in the mean, we take the expectation of (38) and (39) conditioned on the weight vector, $\mathbf{w}(i)$. We get,

$$\begin{aligned} E\{e_1(i+1)|\mathbf{w}(i)\} &= E\{e_1(i)|\mathbf{w}(i)\} - \mu \left[E\{g_1(i)|\mathbf{w}(i)\} \right. \\ &\quad \left. - E\left\{ \frac{\partial P(w_{12}, w_{21})}{\partial w_{12}} \Big|_{\mathbf{w}_{\text{opt}}} \Big| \mathbf{w}(i) \right\} \right] \end{aligned} \quad (40)$$

$$\begin{aligned} E\{e_2(i+1)|\mathbf{w}(i)\} &= E\{e_2(i)|\mathbf{w}(i)\} - \mu \left[E\{g_2(i)|\mathbf{w}(i)\} \right. \\ &\quad \left. - E\left\{ \frac{\partial Q(w_{12}, w_{21})}{\partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} \Big| \mathbf{w}(i) \right\} \right]. \end{aligned} \quad (41)$$

4.2 Approximate Terms for the True and Estimate Gradients

In the case when the cross coupling constants $\left| \frac{a_{12}}{a_{22}} \right|$ and $\left| \frac{a_{21}}{a_{11}} \right|$ are -10 to -15 dB then $w_{12}w_{21} \ll 1$, and we can approximate the true gradients from (13) and (14), by:

$$\begin{aligned} \frac{\partial E\{y_{1d}(n)^2\}}{\partial w_{12}} \Big|_{\mathbf{w}_{\text{opt}}} &\approx 2 \left[\delta_{11} E\{I_1^2(n)\} [(a_{11} + w_{12\text{opt}} a_{21}) a_{21} + a_{11}^2 w_{21\text{opt}}] \right. \\ &\quad + \delta_{12} E\{I_2^2(n)\} [(a_{12} + a_{22} w_{12\text{opt}}) a_{22} + a_{12}^2 w_{21\text{opt}}] \\ &\quad \left. + E\{n_1^2(n)\} w_{21\text{opt}} + E\{n_1^2(n)\} w_{12\text{opt}} \right] \end{aligned} \quad (42)$$

and

$$\begin{aligned} \frac{\partial E\{y_{2d}(n)^2\}}{\partial w_{21}} \Big|_{\mathbf{w}_{\text{opt}}} &\approx 2 \left[\delta_{21} E\{I_1^2(n)\} [(a_{21} + a_{11} w_{21\text{opt}}) a_{11} + a_{21}^2 w_{12\text{opt}}] \right. \\ &\quad + \delta_{22} E\{I_2^2(n)\} [(a_{22} + a_{12} w_{21\text{opt}}) a_{12} + a_{22}^2 w_{12\text{opt}}] \\ &\quad \left. + E\{n_1^2(n)\} w_{21\text{opt}} + E\{n_1^2(n)\} w_{12\text{opt}} \right]. \end{aligned} \quad (43)$$

Similarly, under the assumption that $E\{I_k^2(n)\} = E\{I_k^2(i)\}$ and $E\{n_k^2(n)\} = E\{n_k^2(i)\}$ $k = 1, 2$, then, the estimated gradients can be approximated from (29), (30) to get

$$\begin{aligned}
E\{g_1(i)|\mathbf{w}(i)\} \approx & 2\left[\delta_{11}E\{I_1^2(n)\}[(a_{11} + a_{21}w_{12}(i))a_{21} + a_{11}^2w_{21}(i)]\right. \\
& + \delta_{12}E\{I_2^2(n)\}[(a_{12} + a_{22}w_{12}(i))a_{22} + a_{12}^2w_{21}(i)] \\
& \left. + E\{n_2^2(n)\}w_{21}(i) + E\{n_1^2(n)\}w_{12}(i)\right] \quad (44)
\end{aligned}$$

and

$$\begin{aligned}
E\{g_2(i)|\mathbf{w}(i)\} \approx & 2\left[\delta_{21}E\{I_1^2(n)\}[(a_{21} + a_{11}w_{21}(i))a_{11} + a_{21}^2w_{12}] + \delta_{22}E\{I_2^2(n)\}\right. \\
& \left. [(a_{22} + a_{12}w_{21}(i))a_{12} + a_{22}^2w_{12}(i)]\right. \\
& \left. + E\{n_2^2(n)\}w_{21}(i) + E\{n_1^2(n)\}w_{12}(i)\right]. \quad (45)
\end{aligned}$$

Subtracting the true gradients in (42) and (43) from the estimate gradients in (44) and (45), we get respectively;

$$\begin{aligned}
E\{g_1(i)|\mathbf{w}(i)\} - \frac{\partial E\{y_{1d}(n)^2\}}{\partial w_{12}}|_{\mathbf{w}_{\text{opt}}} = & 2\left[\delta_{11}E\{I_1^2(n)\}a_{11}^2[w_{21}(i) - w_{21\text{opt}}]\right. \\
& \delta_{11}E\{I_1^2(n)\}a_{21}^2[w_{12}(i) - w_{12\text{opt}}] + \delta_{12}E\{I_2^2(n)\}a_{22}^2[w_{12}(i) - w_{12\text{opt}}] \\
& + \delta_{12}E\{I_2^2(n)\}a_{12}^2[w_{21}(i) - w_{21\text{opt}}] + E\{n_1^2(n)\}[w_{21}(i) - w_{21\text{opt}}] \\
& \left. + E\{n_2^2(n)\}[w_{12}(i) - w_{12\text{opt}}]\right], \quad (46)
\end{aligned}$$

$$\begin{aligned}
E\{g_2(i)|\mathbf{w}(i)\} - \frac{\partial E\{y_{2d}(n)^2\}}{\partial w_{21}}|_{\mathbf{w}_{\text{opt}}} = & 2\left[\delta_{21}E\{I_1^2(n)\}a_{11}^2[w_{21}(i) - w_{21\text{opt}}]\right. \\
& \delta_{21}E\{I_1^2(n)\}a_{21}^2[w_{12}(i) - w_{12\text{opt}}] + \delta_{22}E\{I_2^2(n)\}a_{22}^2[w_{12}(i) - w_{12\text{opt}}] \\
& + \delta_{22}E\{I_2^2(n)\}a_{12}^2[w_{21}(i) - w_{21\text{opt}}] + E\{n_1^2(n)\}[w_{21}(i) - w_{21\text{opt}}] \\
& \left. + E\{n_2^2(n)\}[w_{12}(i) - w_{12\text{opt}}]\right] \quad (47)
\end{aligned}$$

Using (46), (47) in (40) and (41) respectively, we have

$$\begin{aligned}
E\{e_1(i+1)|\mathbf{w}(i)\} = & (1 - \mu a)E\{e_1(i)|\mathbf{w}(i)\} \\
& - \mu b E\{e_2(i)|\mathbf{w}(i)\} \quad (48)
\end{aligned}$$

$$E\{e_2(i+1)|\mathbf{w}(i)\} = -\mu c E\{e_1(i)|\mathbf{w}(i)\} + (1 - \mu d) E\{e_2(i)|\mathbf{w}(i)\} \quad (49)$$

where

$$a = 2[\delta_{12} E\{I_2^2(n)\} a_{22}^2 + \delta_{11} E\{I_1^2(n)\} a_{21}^2 + E\{n_2^2(n)\}] \quad (50)$$

$$b = 2[\delta_{11} E\{I_1^2(n)\} a_{11}^2 + \delta_{12} E\{I_2^2(n)\} a_{12}^2 + E\{n_1^2(n)\}] \quad (51)$$

$$c = 2[\delta_{22} E\{I_2^2(n)\} a_{22}^2 + \delta_{21} E\{I_1^2(n)\} a_{21}^2 + E\{n_2^2(n)\}] \quad (52)$$

$$d = 2[\delta_{21} E\{I_1^2(n)\} a_{11}^2 + \delta_{22} E\{I_2^2(n)\} a_{12}^2 + E\{n_1^2(n)\}]. \quad (53)$$

In matrix notation, we can write

$$E\{\mathbf{e}(i+1)|\mathbf{w}(i)\} = (\mathbf{I} - \mu\mathbf{A})E\{\mathbf{e}(i)|\mathbf{w}(i)\} \quad (54)$$

where \mathbf{I} and \mathbf{A} are the identity and the weight error matrices, respectively.

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (55)$$

Taking the expected value of (53) over the weights $\mathbf{w}(i)$ we can write .

$$E\{\mathbf{e}(i+1)\} = (\mathbf{I} - \mu\mathbf{A})E\{\mathbf{e}(i)\}. \quad (56)$$

If $\|\mathbf{I} - \mu\mathbf{A}\| < 1$ then $\lim_{i \rightarrow \infty} E\{\mathbf{e}(i)\} \rightarrow 0$. Therefore, we can establish an upper bound for the convergence constant μ

$$0 < \mu < \frac{1}{\lambda_{\max}} \quad (57)$$

where λ_{\max} is the maximum eigenvalue of the weight error matrix \mathbf{A} .

V. RESULTS

Using a computer, we simulated nondispersive fading channel and employed a power-power canceler to eliminate the effect of cross-pol interference. Perturbation sequences

used in the computer simulation are chosen from the rows of Hadamard matrix (i.e $p_1 = [1, -1, 1, -1, \dots]$ and $p_2 = [1, 1, -1, -1, 1, 1, \dots]$).

The block diagram of the power-power canceler using perturbation sequences in the control algorithm is given in Fig. 1. We applied two independent uniformly distributed bipolar data to the nondispersive channel. Then, corrupted data is applied to the canceler. In Fig. 2, the interference power residue versus the data sample is given for -14 dB cross-pol interference with the perturbation length $N=8$ and different perturbation magnitudes Λ . The same experiment is done for different perturbation sequence sizes and depicted in Fig. 3. The results depicted in these figures aforementioned are the average of four random experiments.

VI. CONCLUSION

In this report, we studied the dynamic analysis of the power-power canceler by using orthogonal perturbation sequences in the control algorithm. The results of the computer analysis shows that as the perturbation magnitude is reduced, the interference power residue decreases. Also, as the perturbation sequence length is increased a smooth estimate of the gradient is obtained, but the convergence time takes longer, as expected.

We conclude perturbation sequences can be used effectively in cross-pol interference cancellation.

VII. REFERENCES

- [1] Y. Barness, A. Dinç and H. Messer-Yaron, *The Bootstrapped Algorithm: Fast Algorithm for Blind Signal Separation*, Final Technical Report. RL-TR-93-24. April, 1993.
- [2] B. Widrow and J. McCool, "A Comparison of Adaptive Algorithms used on the Method of Steepest Decent and Random Search," *IEEE Trans. Acoustic, Speech and Signal Processing*, Vol. ASP-224, No. 5, pp 615-637. Sep. 1976.
- [3] A. Cantoni, "Application of Orthogonal Perturbation Sequences to Adaptive Beamforming," *IEEE Trans. Antenna Propagation*, Vol. AP-28, No. 2, pp 191-202. March, 1980.

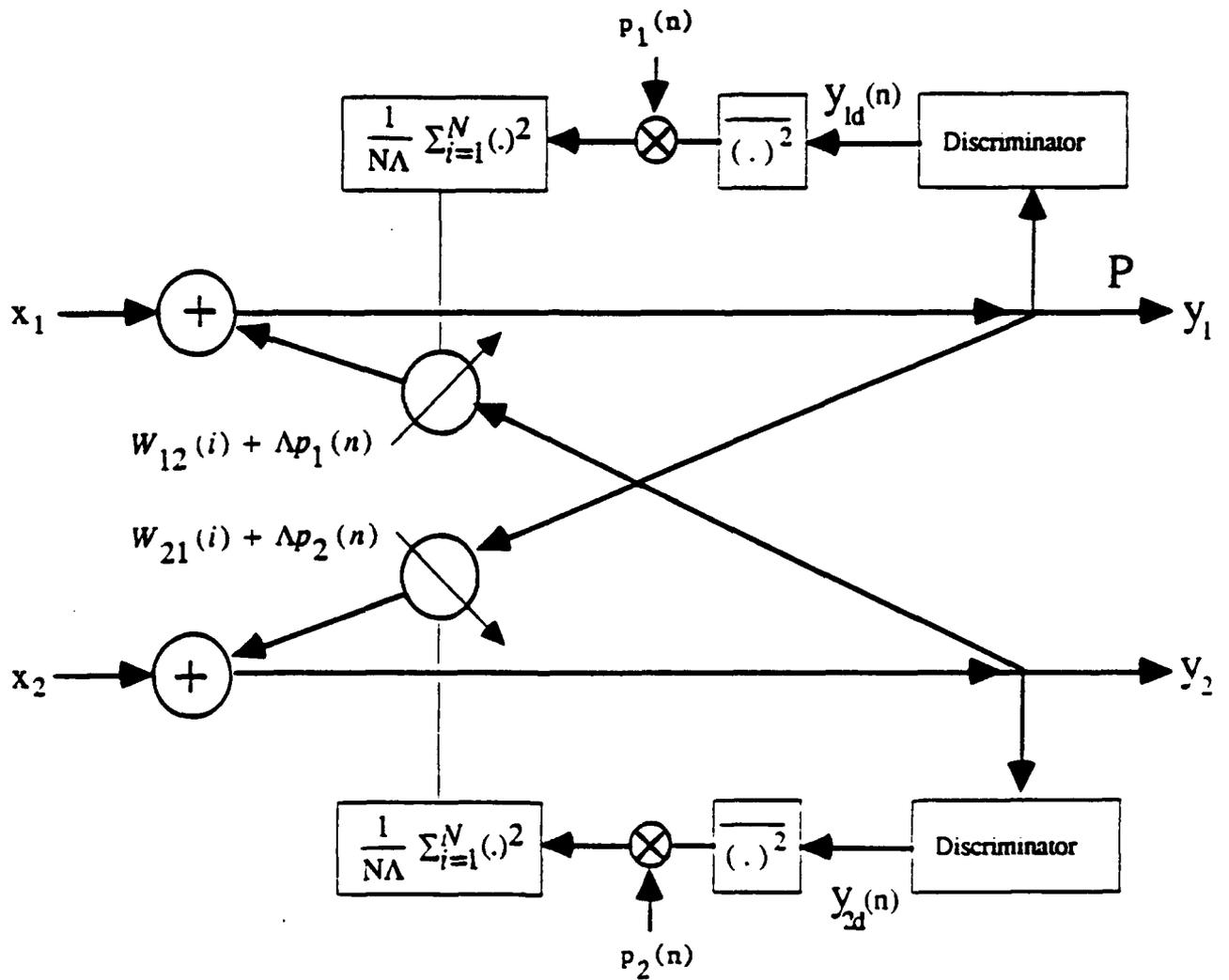


Figure 1 Power-Power Cross-Pol Interference Canceler controlled by orthogonal perturbation sequences.

POWER-POWER SCHEME
USING PERTURBATION SEQUENCES

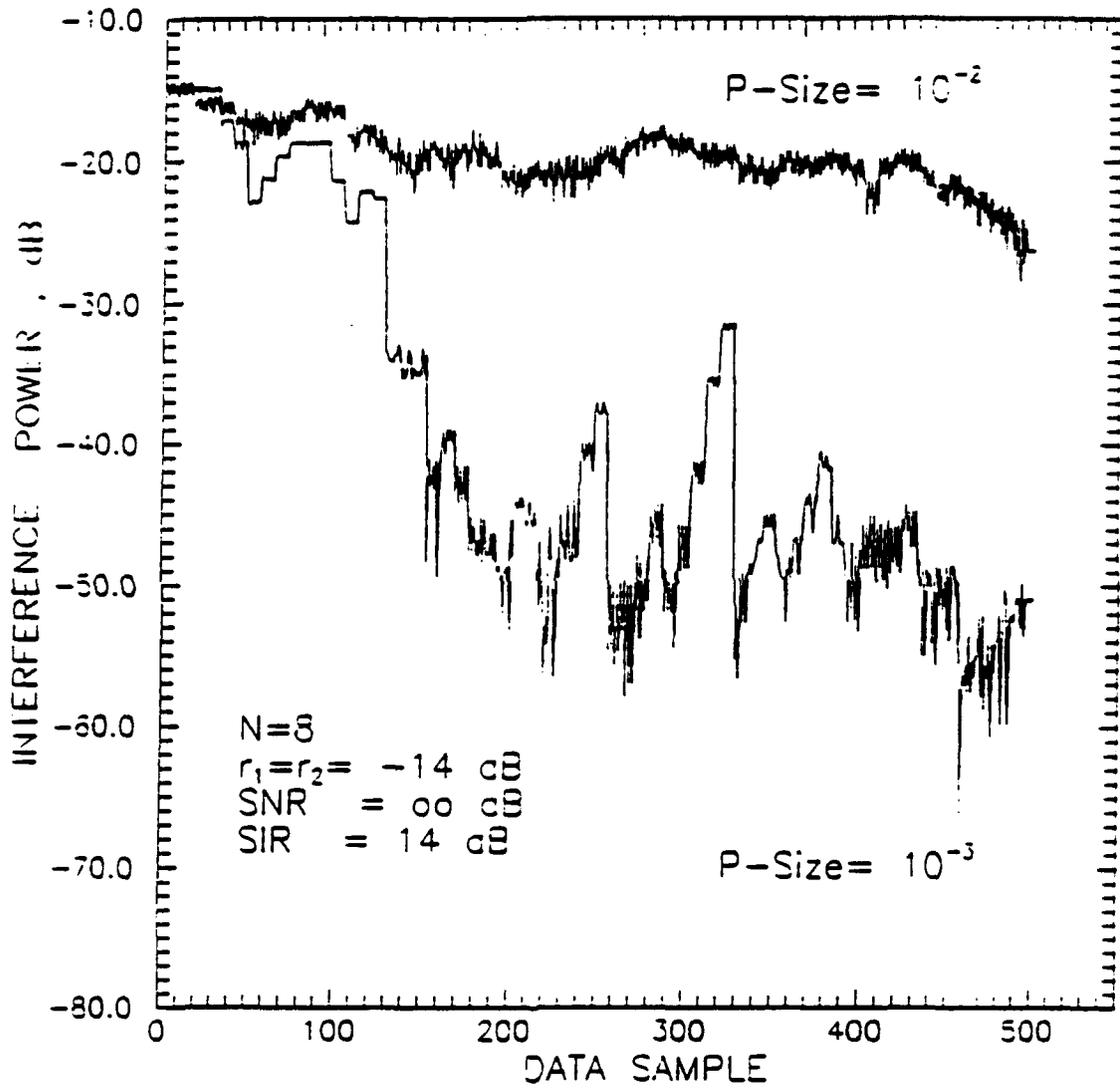


Figure 2 Effect of different perturbation magnitudes on convergence.

POWER-POWER SCHEME
USING PERTURBATION SEQUENCES

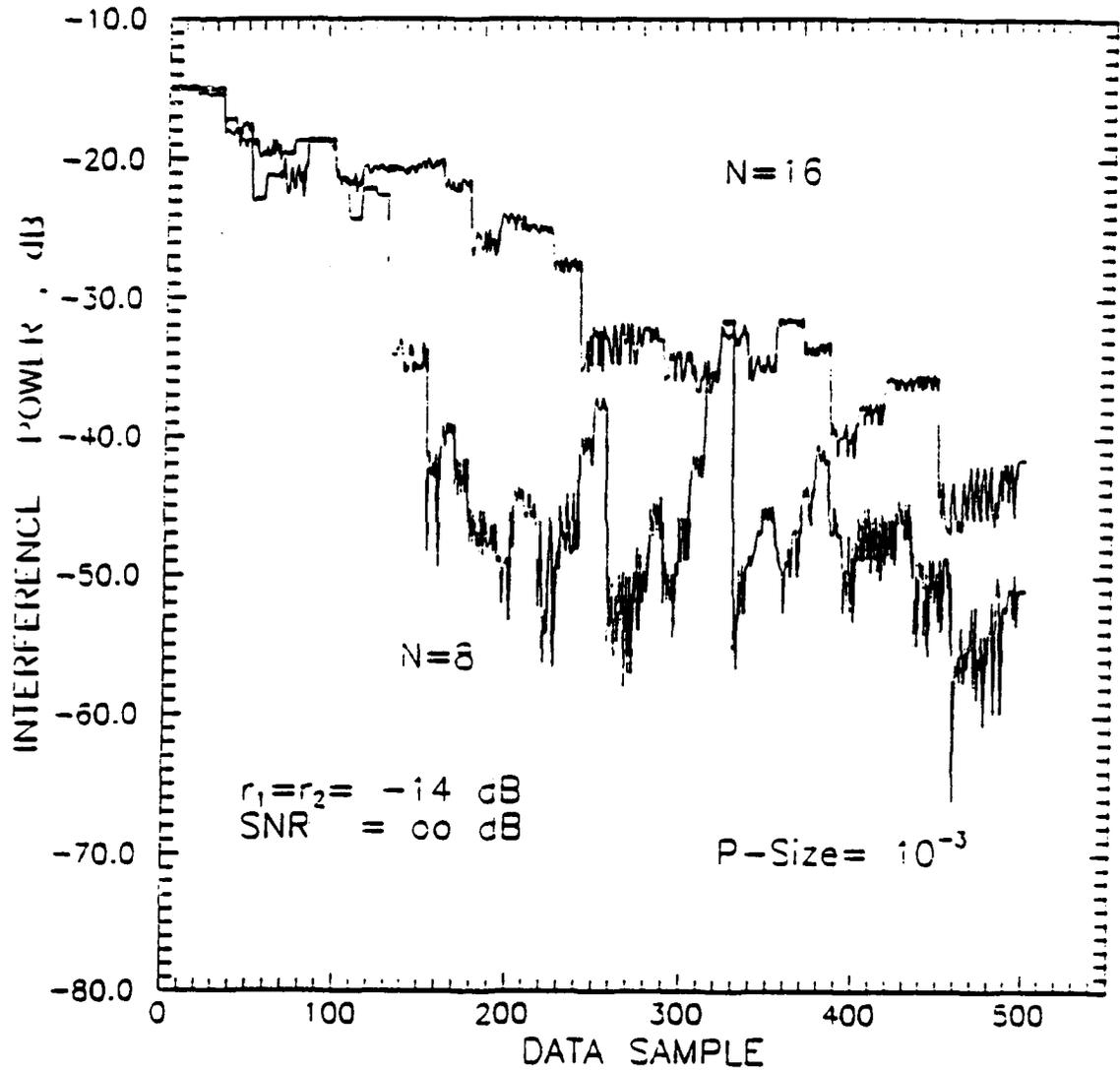


Figure 3 Comparison of different perturbation lengths on convergence.

VIII. APPENDIX

Let

$$E\{g_1(i)|\mathbf{w}(i)\} \triangleq E\{g_{11}(i)|\mathbf{w}(i)\} + E\{g_{12}(i)|\mathbf{w}(i)\} \\ + E\{g_{13}(i)|\mathbf{w}(i)\} + E\{g_{14}(i)|\mathbf{w}(i)\} \quad (\text{P-1})$$

where g_{1k} $k = 1, \dots, 4$ are the different four terms in (27). For the sake of simplicity, we evaluate (27) term by term.

That is:

$$E\{g_{11}(i)|\mathbf{w}(i)\} \triangleq \frac{\delta_{11}E\{I_1^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} (a_{11} + w_{12}(m)a_{21})^2 \\ \cdot [1 + w_{12}(m)w_{21}(m)]^2 p_1(m) \quad (\text{P-2})$$

$$E\{g_{12}(i)|\mathbf{w}(i)\} \triangleq \frac{\delta_{12}E\{I_2^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} (a_{12} + w_{12}(m)a_{22})^2 \\ \cdot [1 + w_{12}(m)w_{21}(m)]^2 p_1(m) \quad (\text{P-3})$$

$$E\{g_{13}(i)|\mathbf{w}(i)\} \triangleq \frac{E\{n_1^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} [1 + w_{12}(m)w_{21}(m)]^2 p_1(m) \quad (\text{P-4})$$

$$E\{g_{14}(i)|\mathbf{w}(i)\} \triangleq \frac{E\{n_2^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} w_{12}^2(m) [1 + w_{12}(m)w_{21}(m)]^2 p_1(m) \quad (\text{P-5})$$

where we also used the fact that the random sequences $I_l(i)$ and $n_l(i)$ $l = 1, 2$ are stationary.

Using (6), (7) in (P-2), we can get;

$$E\{g_{11}(i)|\mathbf{w}(i)\} = \frac{\delta_{11}E\{I_1^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} (a_{11} + [w_{12}(i) + \Lambda p_1(m)]a_{21})^2 \\ \cdot \left(1 + [w_{12}(i) + \Lambda p_1(m)][w_{21}(i) + \Lambda p_2(m)]\right)^2 p_1(m) \\ = \frac{\delta_{11}E\{I_1^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} X_{11}(m)X_{12}(m)p_1(m) \quad (\text{P-6})$$

where,

$$X_{11}(m) \triangleq (a_{11} + [w_{12}(i) + \Lambda p_1(m)]a_{21})^2 \quad (\text{P-7})$$

$$X_{12}(m) \triangleq \left(1 + [w_{12}(i) + \Lambda p_1(m)][w_{21}(i) + \Lambda p_2(m)]\right)^2 \quad (\text{P-8})$$

As a function of the perturbation $p_1(m)$ and $p_2(m)$, we write (P-7) and (P-8) as

$$X_{11}(m) = A + B\Lambda p_1(m) + C\Lambda^2 p_1(m)p_1(m) \quad (\text{P-9})$$

$$\begin{aligned} X_{12}(m) = & D + E\Lambda p_1(m) + F\Lambda p_2(m) + G\Lambda^2 p_1(m)p_2(m) + H\Lambda^2 p_1(m)p_1(m) \\ & + I\Lambda^2 p_2(m)p_2(m) + J\Lambda^3 p_1(m)p_2(m)p_2(m) + K\Lambda^3 p_2(m)p_1(m)p_1(m) \\ & + \Lambda^4 p_1(m)p_1(m)p_2(m)p_2(m) \end{aligned} \quad (\text{P-10})$$

where,

$$A = [a_{11} + a_{21}w_{12}(i)]^2 \quad (\text{P-11})$$

$$B = 2a_{21}[a_{11} + a_{21}w_{12}(i)] \quad (\text{P-12})$$

$$C = a_{21}^2 \quad (\text{P-13})$$

$$D = [1 + w_{12}(i)w_{21}(i)]^2 \quad (\text{P-14})$$

$$E = 2[1 + w_{12}(i)w_{21}(i)]w_{21}(i) \quad (\text{P-15})$$

$$F = 2[1 + w_{12}(i)w_{21}(i)]w_{12}(i) \quad (\text{P-16})$$

$$G = 2[1 + 2w_{12}(i)w_{21}(i)] \quad (\text{P-17})$$

$$H = w_{21}^2(i) \quad (\text{P-18})$$

$$I = w_{12}^2(i) \quad (\text{P-19})$$

$$J = 2w_{12}(i) \quad (\text{P-20})$$

$$K = 2w_{21}(i) \quad (\text{P-21})$$

Therefore, from (P-6), and by using the orthogonal properties of the perturbation sequences $p_1(m)$ and $p_2(m)$, we get

$$E\{g_{11}(i)|\mathbf{w}(i)\} = \frac{\delta_{11}E\{I_1(i)^2\}}{N\Lambda} [(AE + BD)\Lambda N$$

$$+(AJ + BH + BI + CE)\Lambda^3 N + (B + CJ)\Lambda^5 N] \quad (\text{P-22})$$

For $\Lambda \ll 1$ the second and the third terms in the parenthesis is small in comparison to the first and we have,

$$E\{g_{11}(i)|\mathbf{w}(i)\} \approx \delta_{11}E\{I_1^2(i)\}(AE + BD) \quad (\text{P-23})$$

with A, B, D and E defined in (P-11), (P-12), (P-14) and (P-15) respectively. Therefore:

$$E\{g_{11}(i)|\mathbf{w}(i)\} \approx 2\delta_{11}E\{I_1^2(i)\} \left[[a_{11} + a_{21}w_{12}(i)]^2 [1 + w_{12}(i)w_{21}(i)]w_{21}(i) + [a_{11} + a_{21}w_{12}(i)]a_{21}[1 + w_{12}(i)w_{21}(i)]^2 \right] \quad (\text{P-24})$$

Similarly, using (6) and (7) in (P-3), we get,

$$E\{g_{12}(i)|\mathbf{w}(i)\} = \frac{\delta_{12}E\{I_2^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} X_{13}(m)X_{12}(m)p_1(m) \quad (\text{P-25})$$

where,

$$X_{13}(m) \triangleq (a_{12} + [w_{12}(i) + \Lambda p_1(m)]a_{22})^2 \quad (\text{P-26})$$

and $X_{12}(m)$ as in (P-10).

As a function of the perturbation $p_1(m)$, we write (P-26) as

$$X_{13}(m) = A_1 + B_1\Lambda p_1(m) + C_1\Lambda^2 p_1(m)p_1(m) \quad (\text{P-27})$$

where,

$$A_1 = [a_{12} + a_{22}w_{12}(i)]^2 \quad (\text{P-28})$$

$$B_1 = 2[a_{12} + a_{22}w_{12}(i)]a_{22} \quad (\text{P-29})$$

$$C_1 = a_{22}^2. \quad (\text{P-30})$$

Therefore, from (P-25) and by using the orthogonal properties of $p_1(m)$ and $p_2(m)$, we get,

$$E\{g_{12}(i)|\mathbf{w}(i)\} = \frac{\delta_{12}E\{I_2^2(i)\}}{N\Lambda} [(A_1E + B_1D)\Lambda N + (A_1J + B_1H + B_1I + C_1E)\Lambda^3 N + (B_1 + C_1J)\Lambda^5 N]. \quad (\text{P-31})$$

For $\Lambda \ll 1$, the second and the third terms in the parenthesis is small in comparison to the first and we have,

$$E\{g_{12}(i)|\mathbf{w}(i)\} \approx \delta_{12}E\{I_2^2(i)\}(A_1E + B_1D) \quad (\text{P} - 32)$$

with A_1 , B_1 , D and E defined in (P-28), (P-29), (P-14), (P-15), respectively. Therefore,

$$E\{g_{12}(i)|\mathbf{w}(i)\} \approx 2\delta_{12}E\{I_2^2(i)\} \left[[a_{12} + a_{22}w_{12}(i)]^2 [1 + w_{12}(i)w_{21}(i)]w_{21}(i) + [a_{12} + a_{22}w_{12}(i)]a_{22}[1 + w_{12}(i)w_{21}(i)]^2 \right]. \quad (\text{P} - 33)$$

Again, using (6) and (7) in (P-4), we get,

$$E\{g_{13}(i)|\mathbf{w}(i)\} = \frac{E\{n_1^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} X_{12}(m)p_1(m) \quad (\text{P} - 34)$$

with $X_{12}(m)$ as in (P-10).

Applying the orthogonal properties of the perturbation, we get

$$E\{g_{13}(i)|\mathbf{w}(i)\} \approx \frac{E\{n_1^2(i)\}}{N\Lambda} (E\Lambda + J\Lambda^3)N. \quad (\text{P} - 35)$$

For Λ small the second term can be neglected and using (P-15) and (P-20) in (P-35), we get;

$$E\{g_{13}(i)|\mathbf{w}(i)\} \approx 2E\{n_1^2(i)\}[1 + w_{12}(i)w_{21}(i)]w_{21}(i). \quad (\text{P} - 36)$$

Finally, using (6) and (7) in (P-5), we get,

$$E\{g_{14}(i)|\mathbf{w}(i)\} = \frac{E\{n_2^2(i)\}}{N\Lambda} \sum_{m=iN+1}^{(i+1)N} X_{14}(m)X_{12}(m)p_1(m) \quad (\text{P} - 37)$$

where;

$$\begin{aligned} X_{14}(m) &\triangleq [w_{12}(i) + \Lambda p_1(m)]^2 \\ &= w_{12}^2(i) + 2w_{12}(i)\Lambda p_1(m) + \Lambda^2 p_1(m)p_1(m) \end{aligned} \quad (\text{P} - 38)$$

and $X_{12}(m)$ as in (P-10).

Again applying the orthogonal properties of the perturbation, we get

$$\begin{aligned}
E\{g_{14}(i)|\mathbf{w}(i)\} &= \frac{E\{n_2^2(i)\}}{N\Lambda} \left[[Ew_{12}^2(i) + 2Dw_{12}(i)]\Lambda.V \right. \\
&\quad + [E + 2w_{12}(i)H + 2Iw_{12}(i) + Jw_{12}^2(i)]\Lambda^3.V \\
&\quad \left. + (J + 2w_{12}(i))\Lambda^5.N. \right] \tag{P-39}
\end{aligned}$$

With the same approximation in (P-33), we can write

$$E\{g_{14}(i)|\mathbf{w}(i)\} \approx E\{n_2^2(i)\}[2D + Ew_{12}(i)]w_{12}(i). \tag{P-40}$$

With D and E defined by (P-14) and (P-15), respectively. Therefore,

$$\begin{aligned}
E\{g_{14}(i)|\mathbf{w}(i)\} &\approx 2E\{n_2^2(i)\} \left[[1 + w_{12}(i)w_{21}(i)]^2 \right. \\
&\quad \left. + [1 + w_{12}(i)w_{21}(i)]w_{21}(i)w_{12}(i) \right] w_{12}(i). \tag{P-41}
\end{aligned}$$

Combining the four terms from (P-24), (P-32), (P-36) and (P-41) in (P-1), we write,

$$\begin{aligned}
E\{g_1(i)|\mathbf{w}(i)\} &= 2 \left[\delta_{11} E\{I_1^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 [a_{11} + w_{12}(i)a_{21}] a_{21} \right. \right. \\
&\quad \left. \left. + [1 + w_{12}(i)w_{21}(i)] [a_{11} + w_{12}(i)a_{21}(i)]^2 w_{21}(i) \right) \right. \\
&\quad + \delta_{12} E\{I_2^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 [a_{12} + w_{12}(i)a_{22}] a_{22} \right. \\
&\quad \left. \left. + [1 + w_{12}(i)w_{21}(i)] [a_{12} + w_{12}(i)a_{22}]^2 w_{21}(i) \right) \right. \\
&\quad + E\{n_1^2(i)\} [1 + w_{12}(i)w_{21}(i)] w_{21}(i) \\
&\quad + E\{n_2^2(i)\} \left([1 + w_{12}(i)w_{21}(i)]^2 w_{12}(i) \right. \\
&\quad \left. \left. + [1 + w_{12}(i)w_{21}(i)] w_{12}^2(i) w_{21}(i) \right) \right]. \tag{P-42}
\end{aligned}$$

Similar lengthy manipulation can be applied to obtain the *estimate* of the gradient of Q with respect to w_{21} . However, due to the symmetry, we can obtain $Q(w_{12}, w_{21})$ from $P(w_{12}, w_{21})$ by taking $I_2(i)$, $n_2(i)$, a_{22} , a_{21} , w_{21} respectively instead of $I_1(i)$, $n_1(i)$, a_{11} , a_{12} , w_{12} and vice versa. Also, replacing δ_{11} and δ_{12} by δ_{22} , δ_{21} .

Applying these changes and replacements to (P-42) we get,

$$\begin{aligned}
 E\{g_2(i)|\mathbf{w}(i)\} = & 2\left[\delta_{21}E\{I_1^2(i)\}\left([1 + w_{12}(i)w_{21}(i)]^2[a_{21} + w_{21}(i)a_{11}]a_{11}\right.\right. \\
 & \left. + [1 + w_{12}(i)w_{21}(i)][a_{21} + w_{21}(i)a_{11}]^2w_{12}(i)\right) \\
 & + \delta_{22}E\{I_2^2(i)\}\left([1 + w_{12}(i)w_{21}(i)]^2[a_{22} + w_{21}(i)a_{12}]a_{12}\right. \\
 & \left. + [1 + w_{12}(i)w_{21}(i)][a_{22} + w_{21}(i)a_{12}]^2w_{12}(i)\right) \\
 & + E\{n_1^2(i)\}[1 + w_{12}(i)w_{21}(i)]w_{12}(i) \\
 & + E\{n_2^2(i)\}\left([1 + w_{12}(i)w_{21}(i)]^2w_{21}(i)\right. \\
 & \left. + [1 + w_{12}(i)w_{21}(i)]w_{21}^2(i)w_{12}(i)\right)].
 \end{aligned}
 \tag{P-43}$$

APPENDIX D
BOOTSTRAPPED ADAPTIVE SEPARATION OF TWO
SUPERIMPOSED SIGNALS -
DYNAMIC STUDY OF POWER-CORRELATOR SCHEME
USING WEIGHTS DITHERING WITH PN SEQUENCES ¹

by

Ruth Onn and Yeheskel Bar-Ness

ABSTRACT

This report considers the dynamic behavior of a blind signal separator, suggested and study by on of the author (Bar-ness). In this implementation, the error functions, which are derivatives of easily calculated functions, are estimated by adding a dither series to the weights. The resulting differential equations are analyzed. Constraints, imposed on the various parameters by the need for stability in the control loops, are found, as well as the sensitivity of the output to the interfering input.

¹This work was presented at CISS held at Johns Hopkins University, March 1993.
This work was partially supported by a grant from Rome Air Force Lab, (AFSC), Griffis Air Force Base, N.Y., under contract F30602-88-D-0025, Task C-2-2404.

I. INTRODUCTION

The topic of blind signal separation has recently been attracting vast interest among researchers [1,2]. Three structures of two signal blind signal separator were suggested in [3]. They can be adaptively controlled by either output power minimization and/or output de-correlation. It was shown that, in the absence of noise, these criteria lead to total signal separation. The research group at the Center for Communication and Signal Processing at NJIT has been examining this separator, analyzing both its theoretical aspects and its practical applications. This work has resulted in a few journal articles dealing with the bandwidth-complexity tradeoff [4], the question of noise effect [5], spatial separation of wide-band sources [6], etc. Application to satellite and microwave digital communication links were also examined [7,8]. Recently some results were obtained in applying the blind separator to decision feedback adaptive channel equalization.

This paper considers the dynamic behavior of these separators. In order to obtain a gradient of the output power needed to control the weights, a pseudo-random sequence is used to dither the weights. Minimizing the magnitude of the correlation between the two outputs is effected directly by forcing the correlation to zero.

As was shown in [5], one of the structures, called Forward/Forward, must be controlled by correlation zero-forcing. Another structure, termed Backward/Backward, can be controlled by either minimum power or correlation zero-forcing.

In this paper we will concentrate on a two input two output configuration. with the Forward/Backward structure, see figure . One weight will be controlled by minimum output power, while the other by zero-forcing of the correlation between the two outputs. Dynamic behavior of the other two structures can be similarly analyses.

Following the notation in [5] the two inputs to the interference cancelled are given by:

$$\begin{aligned}x_1(t) &= s_1(t) + a_1 s_2(t) \\x_2(t) &= a_2 s_1(t) + s_2(t)\end{aligned}$$

where a_1 and a_2 are unknown complex numbers, $s_1(t)$ and $s_2(t)$ uncorrelated signals. The separation is attempted by applying complex weights w_1 and w_2 to the inputs to obtain (possibly scaled) estimates of the signals.

$$\begin{aligned}\hat{s}_1(t) = y_1(t) &= x_1(t) + w_1 x_2(t) \\ \hat{s}_2(t) = y_2(t) &= w_2 x_1(t) + x_2(t)\end{aligned}$$

Based on considerations of complexity and bandwidth, error functions are chosen to control the separating weights. Estimates of the error functions are integrated to form the weights. In this implementation a dither series is added to the weights, and the error functions which are derivatives of easily calculated functions are estimated in this way. Thus, the weights are composed from the controller output w_n^C and the dither sequence w_n^D .

$$w_n = w_n^C + \mu w_n^D \quad n = 1, 2 \quad (1)$$

The dither sequence is scaled by a small constant μ , chosen so that

$$\mu \ll |w_n^C| \quad n = 1, 2 \quad (2)$$

The complex dither sequences, w_n^D $n = 1, 2$, are produced from two uncorrelated pseudo-random sequences by staggering the real and imaginary part.

$$w_n^D(t) = w_{nR}^D(t) + j w_{nI}^D(t) \quad n = 1, 2$$

where j is the square root of minus one. Thus

$$E [w_{1K}^D(t) w_{2L}^D(t)] = 0 \quad \text{for } K, L = R, I$$

where $E[\cdot]$ is the expectation operator. This implies

$$E [w_1(t) w_2(t)] = 0$$

II. THE POWER CORRELATOR DITHER CONTROLLED SIGNAL SEPARATOR

This scheme of separating the signal involves a trade-off between complexity and bandwidth [4]. It is composed of two loops, one is controlled by the power in one of the outputs, the other by the cross-correlation of the two outputs. The power controlled weight, which we deal with first necessitates estimating the derivative of the power function with respect to the controlled weight. The second weight is controlled directly by the sample cross-correlation, and therefore no dither has to be added to it.

2.1 The Power Controlled Loop

The output, the power of which controls the first weight is

$$y_1(t) = (1 + w_1(w_2 + a_2))s_1 + (a_1 + w_1(1 + a_1w_2))s_2 \quad (3)$$

Substituting (1) into (3) and dropping the dependence on t we have

$$\begin{aligned} y_1 &= (1 + (w_1^C + \mu w_1^D)(w_2 + a_2))s_1 + \\ &\quad (a_1 + w_1^C + w_1^D)(1 + a_1w_2)s_2 \\ &= (1 + w_1^C(w_2 + a_2))s_1 + (a_1 + w_1^C(1 + a_1w_2))s_2 + \\ &\quad + \mu \left(((w_2 + a_2)s_1 + (1 + a_1w_2)s_2)w_1^D \right) \end{aligned}$$

Where we used (2) to disregard entries with μ^2 . Taking the power of y_1 gives

$$\begin{aligned} |y_1|^2 &= \left| (1 + w_1^C(w_2 + a_2))s_1 + (a_1 + w_1^C(1 + a_1w_2))s_2 \right|^2 + \\ &\quad + 2\mu \Re \left[\left((1 + w_1^C(w_2 + a_2))s_1 + (a_1 + w_1^C(1 + a_1w_2))s_2 \right) \right. \\ &\quad \left. \left((w_2^* + a_2^*)s_1^* + ((1 + a_1^*w_2^*)s_2^*)w_1^{D*} \right) \right] \end{aligned}$$

where the superscript $*$ denotes the complex conjugate and $\Re(\cdot)$ denotes the real part. When we average the power over a proper time period, we can assume that $\overline{s_1 s_2^*} = \overline{s_1^* s_2} = 0$, thus

$$\begin{aligned} \overline{|y_1|^2} &= |1 + w_1^C(w_2 + a_2)|^2 \overline{|s_1|^2} + |a_1 + w_1^C(1 + a_1w_2)|^2 \overline{|s_2|^2} \\ &\quad + 2\mu \Re \left[\left((1 + w_1^C(w_2 + a_2))(w_2^* + a_2^*) \overline{|s_1|^2} + \right. \right. \\ &\quad \left. \left. (a_1 + w_1^C(1 + a_1w_2))(1 + a_1^*w_2^*) \overline{|s_2|^2} \right) w_1^{D*} \right] \end{aligned}$$

This signal is passed through a capacitor which we will assume ideally suppresses all entries that do not depend on w_1^D or w_1^{D*}

$$\overline{|y_1|^2}|_{AC} = 2\mu\Re \left[\left((1 + w_1^C(w_2 + a_2))(w_2^* + a_2^*)\overline{|s_1|^2} + (a_1 + w_1^C(1 + a_1w_2))(1 + a_1^*w_2^*)\overline{|s_2|^2} \right) w_1^{D*} \right] \quad (4)$$

Taking the product of the above with the dither w_1^D gives

$$e_1 = 2\mu \left((1 + w_1^C(w_2 + a_2))(w_2^{C*} + a_2^*)\overline{|s_1|^2} + \right. \quad (5)$$

$$\left. (a_1 + w_1^C(1 + a_1w_2))(1 + a_1^*w_2^{C*})\overline{|s_2|^2} \right) + \quad (6)$$

$$\mu \left((1 + w_1^{C*}(w_2^* + a_2^*))(w_2 + a_2)\overline{|s_1|^2} + \right. \quad (7)$$

$$\left. (a_1^* + w_1^{C*}(1 + a_1^*w_2^*))(1 + a_1w_2)\overline{|s_2|^2} \right) (w_1^D)^2 \quad (8)$$

We used $|w_1^D|^2 = 2$, which follows directly from the definition of the dither signal.

The control signal is next passed through a filter with the response

$$w_1^C = \frac{-g}{1 + \tau p} e_1$$

where p is the Heavyside differentiation operator. Note that $(w_1^D)^2 = -2jw_{1R}^D w_{1I}^D$. We will choose the time constant τ in the filter to be such that the time average $-2j\langle w_{1R}^D w_{1I}^D \rangle$ can be substituted for $(w_1^D)^2$.

$$(1 + \tau \frac{d}{dt} + 2g\mu \left[|w_2 + a_2|^2 \overline{|s_1|^2} + |1 + a_1w_2|^2 \overline{|s_2|^2} \right]) w_1^C = -2g\mu \left[(w_2 + a_2)^* \overline{|s_1|^2} + a_1(1 + a_1w_2)^* \overline{|s_2|^2} \right] + \quad (9)$$

$$-2jg\mu \left[|w_2 + a_2|^2 \overline{|s_1|^2} + |1 + a_1w_2|^2 \overline{|s_2|^2} \right] w_1^{C*} \langle w_{1R}^D w_{1I}^D \rangle + \quad (10)$$

$$-2jg\mu \left[(w_2 + a_2) \overline{|s_1|^2} + a_1^*(1 + a_1w_2) \overline{|s_2|^2} \right] \langle w_{1R}^D w_{1I}^D \rangle \quad (11)$$

The operator notation was replaced by the regular $\frac{d}{dt}$ notation. The terms in (9) through (11) have the following interpretation. The first term is the desired one, for in the steady state, when $\frac{d}{dt}w_1^C = 0$, it leads to

$$w_1^C = \frac{-2g\mu \left[(w_2 + a_2)^* \overline{|s_1|^2} + a_1(1 + a_1w_2)^* \overline{|s_2|^2} \right]}{1 + 2g\mu G}$$

with

$$G = |w_2 + a_2|^2 |s_1|^2 + |1 + a_1 w_2|^2 |s_2|^2 \quad (12)$$

This is the optimal solution for the power controlled weight, provided

$$1 \ll 2g\mu G \quad (13)$$

(see [3] equation (30)). The rest of the terms, as well as the dither added to w_1^C will cause degradation of the separation, relative to the optimal value predicted in [3].

These must be held to an acceptably small level. The terms in (10) represent the cross coupling of the real and imaginary parts of w_1^C , and thus affect its dynamic performance. In (10) the coupling is weighted by the correlation between the real and imaginary part of the dither sequence w_1^D .

Taking into account at first only the contributions of (9) and (10) we obtain the following differential equation

$$\begin{aligned} \tau \frac{d}{dt} w_1^C = & -(1 + 2g\mu G)w_1^C - 2jg\mu \langle w_{1R}^D w_{1I}^D \rangle w_1^C - \\ & - 2g\mu \left[(w_2 + a_2)^* |s_1|^2 + a_1 (1 + a_1 w_2)^* |s_2|^2 \right] \end{aligned} \quad (14)$$

The differential equation (14) can be written in vector form

$$\begin{aligned} \tau \frac{d}{dt} \begin{bmatrix} w_{1R}^C \\ w_{1I}^C \end{bmatrix} = & - \begin{bmatrix} 1 + 2g\mu G & 2g\mu G \langle w_{1R}^D w_{1I}^D \rangle \\ 2g\mu G \langle w_{1R}^D w_{1I}^D \rangle & 1 + 2g\mu G \end{bmatrix} \begin{bmatrix} w_{1R}^C \\ w_{1I}^C \end{bmatrix} \\ & - 2g\mu \begin{bmatrix} \left((w_2^C + a_2) |s_1|^2 + \right. \\ \left. (a_1 + |a_1|^2 w_2^C) |s_2|^2 \right)_R \\ - \left((w_2^C + a_2) |s_1|^2 + \right. \\ \left. (a_1 - |a_1|^2 w_2^C) |s_2|^2 \right)_I \end{bmatrix} \end{aligned} \quad (15)$$

The eigenvalues of (15) are given by

$$\lambda_1 = 1 + 2g\mu G (1 + \langle w_{1R}^D w_{1I}^D \rangle) \quad (16)$$

$$\lambda_2 = 1 + 2g\mu G (1 - \langle w_{1R}^D w_{1I}^D \rangle) \quad (17)$$

Thus the system is stable only if

$$1 + 2g\mu G > 2g\mu G \langle w_{1R}^D w_{1I}^D \rangle$$

and taking (13) that means

$$1 > \langle w_{1R}^D w_{1I}^D \rangle \quad (18)$$

The term in (11) is expected to affect mainly the steady state response of w_1^C . We are especially concerned in how it affects the sensitivity of \hat{s}_1 to s_2 . Thus we define y_{11} and y_{12} as follows

$$y_1(t) = y_{11}s_1(t) + y_{12}s_2(t) \quad (19)$$

$$y_{11}(t) = 1 + w_1^C(w_2 + a_2) + \mu w_1^D(w_2 + a_2) \quad (20)$$

$$y_{12}(t) = a_1 + w_1^C(1 + a_1w_2) + \mu w_1^D(1 + a_1w_2) \quad (21)$$

Rearranging (9) and (11) we have

$$w_1^C = \frac{-2g\mu [(w_2 + a_2)^* |s_1|^2 + a_1(1 + a_1w_2)^* |s_2|^2]}{1 + \tau p + 2g\mu G} + \frac{-2g\mu [(w_2 + a_2) |s_1|^2 + a_1^*(1 + a_1w_2) |s_2|^2] \langle w_{1R}^D w_{1I}^D \rangle}{1 + \tau p + 2g\mu G} \quad (22)$$

Substituting (22) into (21), and using the definition of G in (12), we get

$$y_{12}(t) = \frac{a_1 - 2g\mu(w_2^* + a_2^*)(1 - a_1a_2) |s_1|^2}{1 + \tau p + 2g\mu G} + \frac{-2g\mu [(w_2 + a_2) |s_1|^2 + a_1^*(1 + a_1w_2) |s_2|^2] (1 + a_1w_2) \langle w_{1R}^D w_{1I}^D \rangle}{1 + \tau p + 2g\mu G} + \mu w_1^D (1 + a_1w_2)$$

To get the sensitivity we are seeking we will consider $E[|y_{12}|^2]$.

$$E[|y_{12}|^2] = \left| \frac{a_1 - 2g\mu(w_2^* + a_2^*)(1 - a_1a_2) |s_1|^2}{1 + \tau p + 2g\mu G} \right|^2 + 4g^2\mu^2 |1 + a_1w_2|^2 |(w_2 + a_2) |s_1|^2|^2 + a_1^*(1 + a_1w_2) |s_2|^2|^2 E \left[\left| \frac{\langle w_{1R}^D w_{1I}^D \rangle}{1 + \tau p + 2g\mu G} \right|^2 \right] + 2\mu^2 |1 + a_1w_2|^2 \quad (23)$$

To evaluate (23) we will use the fact that the sequence chip rate $\frac{1}{T}$ is much greater than the closed loop bandwidth. We will approximate the sequence by Gaussian noise with the same power density value of $2\pi T \frac{\text{watts}}{\text{hertz}}$ over the closed loop bandwidth. For a low pass filter, the two sided noise bandwidth equals the reciprocal of twice its time constant. Therefore

$$\begin{aligned} E \left[\left| \frac{\langle w_{1R}^D w_{1I}^D \rangle}{1 + \tau p + 2g\mu G} \right|^2 \right] &= \frac{1}{(1 + 2\mu g G)^2} \frac{2\pi T(1 + 2\mu g G)}{2\tau} \\ &= \frac{\pi T}{\tau} \frac{1}{1 + 2\mu g G} \end{aligned}$$

Substituting in (23), the equation reads

$$E [|y_{12}|^2] = \left| \frac{a_1 - 2g\mu(w_2^* + a_2^*)(1 - a_1 a_2) \overline{|s_1|^2}}{1 + \tau p + 2g\mu G} \right|^2 \quad (24)$$

$$+ \frac{4\pi T g^2 \mu^2}{\tau} \times \quad (25)$$

$$\begin{aligned} &\frac{|1 + a_1 w_2|^2 (w_2 + a_2) \overline{|s_1|^2} + a_1^* (1 + a_1 w_2) \overline{|s_2|^2}}{1 + 2\mu g G} \\ &+ 2\mu^2 |1 + a_1 w_2|^2 \quad (26) \end{aligned}$$

The expression in (24) dominates the right hand side of the equation. Under the assumption $1 \ll 2g\mu G$ it will yield

$$E [|y_{12}|^2] = \frac{|(w_2^* + a_2^*)(1 - a_1 a_2) \overline{|s_1|^2}|^2}{G^2}$$

Which is exactly the one in [3]. (25) represents the effect of non-zero chip duration, and will be made small by making the chip duration as small as possible. It will increase with μg .

(26) imposes the limit on the size of μ . This will make obtaining a large μg hard without a large g . The size of g is in turn limited by considerations of stability, which will come up if the loop is not simply first order.

2.2 The correlation controlled loop

In this loop the weight is controlled by the average power in the sample correlation between the two outputs, $\overline{|y_1 y_2^*|^2}$. As is proven in [3], the relevant signal in the derivative of

the power is the cross-correlation itself, thus we do not need to add a dither component to w_2 . The averaging time constant is chosen such that the output depends only on the power in the signals s_1 and s_2 .

The inputs of the correlator are the two outputs of the circuit. Recall

$$\begin{aligned} y_1^* &= ((s_1 + a_1 s_2)(1 + w_1 w_2) + (a_2 s_1 + s_2)w_1)^* \\ &= s_1^*(w_1^*(w_2^* + a_2) + 1) + s_2^*(w_1^*(a_1^* w_2^* + 1) + a_1^*) \\ y_2 &= (s_1 + a_1 s_1)w_2 + (a_2 s_1 + s_2) \\ &= s_1(w_2 + a_2) + s_2(a_1 w_2 + 1) \end{aligned}$$

Averaging the product of the two inputs gives

$$\overline{i_1^* i_2} = \frac{\overline{|s_1|^2} (w_1^*(w_2^* + a_2) + 1)(w_2 + a_2) + \overline{|s_2|^2} (w_1^*(a_1^* w_2^* + 1) + a_1^*)(a_1 w_2 + 1)}{2}$$

Due to the averaging the terms that involved $s_1^* s_2$ or $s_1 s_2^*$ will null.

This is the error signal e_2 . It is passed through a filter with the response of $\frac{-2}{1+\tau p}$, and thus w_2 obeys the differential equation

$$\begin{aligned} (1 + \tau \frac{d}{dt})w_2 &= -g \left(\overline{|s_1|^2} (w_1^*(w_2^* + a_2) + 1)(w_2 + a_2) \right. \\ &\quad \left. + \overline{|s_2|^2} (w_1^*(a_1^* w_2^* + 1) + a_1^*)(a_1 w_2 + 1) \right) \end{aligned} \quad (27)$$

The steady state response will be

$$\begin{aligned} \overline{|s_1|^2} (w_1^*(w_2^* + a_2) + 1)(w_2 + a_2) + \overline{|s_2|^2} (w_1^*(a_1^* w_2^* + 1) + a_1^*)(a_1 w_2 + 1) \\ = -\frac{w_2}{g} \end{aligned} \quad (28)$$

Provided g is chosen large enough, the right hand side of (28) will be almost zero. This will give us the expected response for w_2 .

To study the dynamics of (27) around the solution, $w_2 = -a_2$ we will linearize the equation. Let $\epsilon = w_2 + a_2$. We will assume ϵ small enough, so we can disregard entries with ϵ^2 .

$$\begin{aligned} \tau \frac{d}{dt} \epsilon &= -(1 + g(\overline{|s_1|^2} + \overline{|s_2|^2} |a_1|^2))\epsilon - g \overline{|s_2|^2} (w_1^* a_1 (1 - a_1 a_2)^* \epsilon + w_1^* a^* (1 - a_1 a_2) \epsilon^*) \\ &\quad + a_2 - g \overline{|s_2|^2} (1 - a_1 a_2) (a_1^* + w_1^* (1 - a_1 a_2)^*) \end{aligned} \quad (29)$$

To obtain results on the stability of (29), it can be represented in vector form. The matrix in this representation has the following eigenvalues

$$\begin{aligned}\lambda_1 &= \frac{-1}{\tau}(1 + g\overline{|s_1|^2} + g\overline{|s_2|^2}|a_1|^2) \\ \lambda_2 &= \frac{-1}{\tau}(1 + g\overline{|s_1|^2} + g\overline{|s_2|^2}|a_1|^2 + g\overline{|s_2|^2}\Re(w_1^*a_1^*(1 - a_1a_2)))\end{aligned}$$

λ_1 is obviously always negative. λ_2 is also negative provided

$$\frac{1 + g\overline{|s_1|^2} + g\overline{|s_2|^2}|a_1|^2}{g\overline{|s_2|^2}} > |\Re(w_1a_1^*(1 - a_1a_2))|$$

Thus, the equation is stable if $\frac{\overline{|s_1|^2}}{\overline{|s_2|^2}}$ is not too small.

The steady state response of ϵ is

$$\begin{aligned}\epsilon_{ss} &= \frac{1}{\lambda_1\lambda_2}(1 + g\overline{|s_1|^2} + g\overline{|s_2|^2}|a_1|^2) \times (a_2 - g\overline{|s_2|^2}(1 - a_1a_2)(w_1(1 - a_1a_2) + a_1)^*) \\ &\quad + \frac{j}{\lambda_1\lambda_2}\overline{|s_2|^2}a_1(1 - a_1a_2)^* \times \Im(w_1(a_2 - g\overline{|s_2|^2}(1 - a_1a_2)(w_1(1 - a_1a_2) + a_1)^*))\end{aligned}\quad (30)$$

Using (30) and $g \gg 1$ we can get

$$\begin{aligned}\epsilon_{ss} &= \frac{1}{\lambda_1\lambda_2} \left[(\overline{|s_1|^2} + \overline{|s_2|^2}|a_1|^2) \times (\overline{|s_2|^2}(1 - a_1a_2)(w_1(1 - a_1a_2) + a_1)^*) \right. \\ &\quad \left. + j\overline{|s_2|^2}a_1(1 - a_1a_2)^* \times \Im(w_1(\overline{|s_2|^2}(1 - a_1a_2)(w_1(1 - a_1a_2) + a_1)^*)) \right]\end{aligned}\quad (31)$$

If we now assume that $w_1 \approx \frac{-a_1}{1 - a_1a_2}$, we will get that the steady state ϵ is very small. Thus if $\delta = w_1 + \frac{a_1}{1 - a_1a_2} \ll 1$, (29) will become

$$\begin{aligned}\overline{|s_1|^2}\epsilon - \overline{|s_2|^2}\frac{a_1^*(1 - a_1a_2)}{(1 - a_1a_2)^*}\epsilon^* + \overline{|s_2|^2}\delta^*\Re((a_1(1 - a_1a_2)^*\epsilon)) &= -\overline{|s_2|^2}|1 - a_1a_2|^2\delta^* \\ \overline{|s_1|^2}\epsilon &\approx -\overline{|s_2|^2}|1 - a_1a_2|^2\delta^*\end{aligned}\quad (32)$$

From (32) it is clear that $|\epsilon|$ will get smaller with $|\delta|$, and will do so in a reasonable pace. provided $\frac{\overline{|s_2|^2}}{\overline{|s_1|^2}}$ is not too large.

III. CONCLUSIONS

Using the proposed model for the weight controller, we obtained a differential equation for the constant part of the first weight, w_1^C from (14)

$$\begin{aligned} \tau \frac{d}{dt} w_1^C = & - (1 + 2g\mu G) w_1^C - 2jg\mu \langle w_{1R}^D w_{1I}^D \rangle w_1^{C*} \\ & - 2g\mu \left[(w_2 + a_2)^* \overline{|s_1|^2} + a_1 (1 + a_1 w_2)^* \overline{|s_2|^2} \right] \end{aligned}$$

where, $G = |w_2 + a_2|^2 \overline{|s_1|^2} + |1 + a_1 w_2|^2 \overline{|s_2|^2}$, w_1^D is the dither sequence added to the weight, and $\langle \cdot, \cdot \rangle$ denotes the time average of the product.

For the weight w_2 we examined its behavior in the vicinity of its steady state. $w_2 = -a_2$. By letting $\epsilon = w_2 + a_2$, and assuming $\epsilon \ll 1$, we showed that $\epsilon(t)$ is controlled by the following differential equation (29)

$$\begin{aligned} \tau \frac{d}{dt} \epsilon = & - (1 + g(\overline{|s_1|^2} + \overline{|s_2|^2} |a_1|^2)) \epsilon \\ & - 2g \overline{|s_2|^2} w_1^* \Re(a_1 (1 - a_1 a_2)^* \epsilon) \\ & + a_2 - g \overline{|s_2|^2} (1 - a_1 a_2) (a_1^* + w_1^* (1 - a_1 a_2)^*) \end{aligned}$$

Studying the differential equations in (14) and (29) we obtained the conditions for convergence in the mean of the weights and their steady state values. We also studied the sensitivity of the quality of separation to the power of the dither and other system parameters.

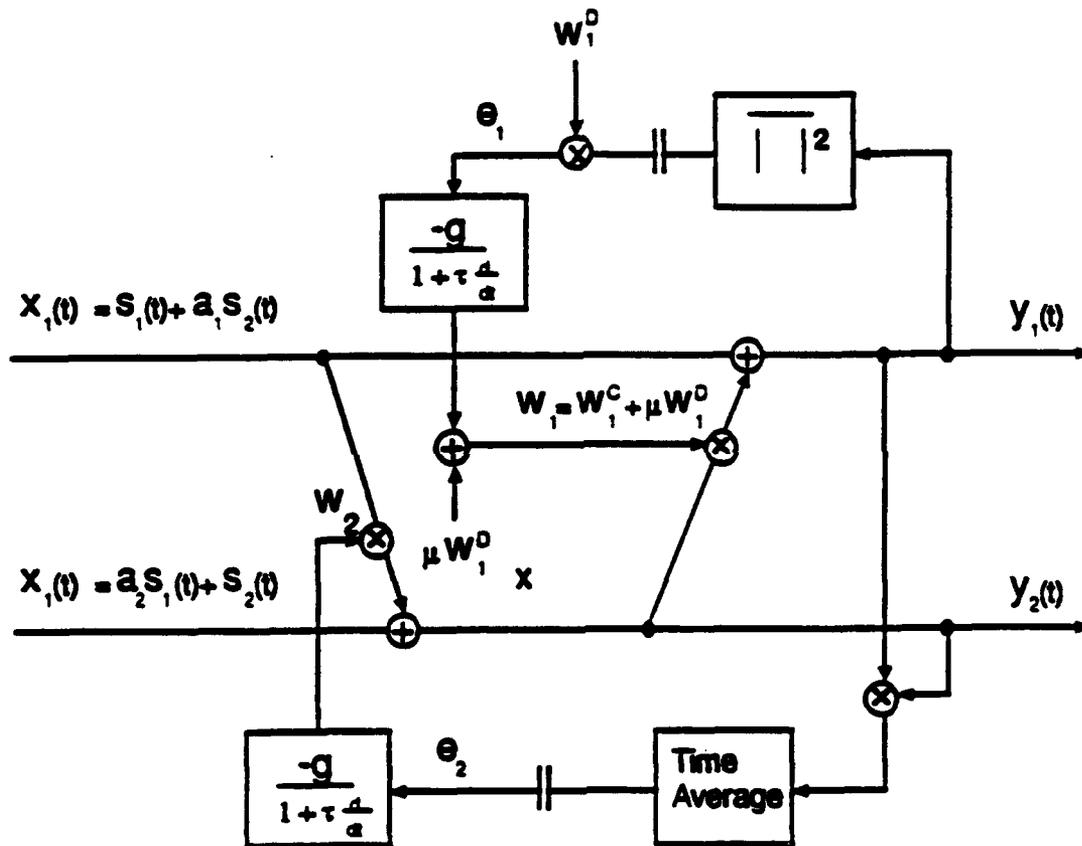


Figure 1 A diagram of the Forward/Backward bootstrap blind signal separator with dither control.

IV. REFERENCES

- [1] C. Jutten and J. Herault, "Blind Separation of Source Part I," *Signal Processing*, 245(1):1-10, July, 1991.
- [2] E. Weinstein, M. Fedor, and A. V. Oppenheim, "Multi-Channel Signal Separation Based on Decorrelation," Submitted to *IEEE Trans. on SP.*, 1992.
- [3] Y. Bar-ness, "Bootstrapped Cross-Pol Interference Canceling Techniques - Steady State Analysis," Submitted to *Signal Processing, Elsevier Science Publishers*, 1992.
- [4] Y. Bar-Ness, "Bootstrapped Adaptive Blind Signal Separation Techniques - Bandwidth Complexity Trade-off." Submitted to *Signal Processing, Elsevier Science Publishers*, 1992.
- [5] Y. Bar-Ness, A. Dinç and H. Messer, " Bootstrapped Adaptive Separation of Superimposed Signals - Analysis of the Effect Thermal Noise on Performance." Submitted to *IEE Proceedings-F, Radar and Signal Processing*, 1992.
- [6] Y. Bar-ness and H. Messer, "Application of the Bootstrapped Signal Separation Algorithms to Spatial Separation of Wideband Sources," Submitted to *IFAC Journal Automatica*, 1992.
- [7] A. Dinç and Y. Bar-Ness, "Convergence and Performance Comparison of Three Different Structures of Bootstrap Blind Algorithm for Multisignal Cochannel Separation." in *Proceedings of MILCOM '92*, pp 913-918, San Diego. CA, Oct.. 1992.
- [8] A. Dinç and Y. Bar-ness, " Error Probabilities of Bootstrapped Blind Adaptive Cross-Pol Canceler for M-ary QAM Signals over Non-Dispersive Fading Channel." In *Proceeding of IEEE ICC '92*, Chicago, Il., June 1992.

APPENDIX E
SIMULTANEOUS SPATIAL SEPARATION AND
DIRECTION-OF-ARRIVAL ESTIMATION OF
WIDEBAND SOURCES USING
BOOTSTRAPPED ALGORITHMS ¹

by

Hagit Messer and Yeheskel Bar-Ness

ABSTRACT

This paper presents two versions of a novel adaptive algorithm for simultaneous spatial separation and direction finding of two wideband, uncorrelated sources received by two separated sensors. The algorithm provides estimates of the sources (direction of arrival - DOA) and outputs two signals which are filtered versions of the source signals. We describe the different configurations of the separation system and study sensitivity of the two signal estimates to the accuracy in estimating the DOA of the sources. We show that perfect separation is achieved even if the DOAs are not perfectly estimated, at the cost of degradation in the output signal-to-noise-ratio. Then we propose implementation of an adaptive control algorithm and we discuss the steady state performance in the presence of an additive noise. We compare the complexity, performance and noise immunity of the two versions of the proposed algorithm.

¹This work was partially supported by a grant from Rome Air Force Lab, (AFSC), Griffis Air Force Base, N.Y., under contract F30602-88-D-0025, project E-21-T49, Task C-2-2404.

I. INTRODUCTION

Fig. 1 describes the scenario where two sources are received by an array of two sensors. The outputs of the two sensors are:

$$\begin{aligned} z_1(t) &= s_1(t - D_1) + s_2(t - D_2) + e_1(t) \\ z_2(t) &= s_1(t + D_1) + s_2(t + D_2) + e_2(t) \end{aligned} \quad (1)$$

where $s_1(t)$ and $s_2(t)$ are the signals radiated from the 1st and 2nd sources respectively, $e_1(t)$ and $e_2(t)$ are the additive noise processes in each of the sensors and D_i is given by,

$$D_i = \frac{d}{2c} \sin \theta_i \quad i = 1, 2 \quad (2)$$

d is the distance between the sensors, c is the propagation velocity and θ_i , $i = 1, 2$ is the bearing of the i th source. We assume that the random signals $s_1(t)$, $s_2(t)$, $e_1(t)$ and $e_2(t)$ are mutually statistically uncorrelated.

In this paper we present adaptive systems which receive $z_1(t)$ and $z_2(t)$ at the input, deliver as output signals $y_1(t) = \hat{s}_1(t)$ and $y_2(t) = \hat{s}_2(t)$, and simultaneously provide the estimates of D_1 and D_2 from which the bearings of the two signals θ_1 and θ_2 can be derived. However, in order to enable separation, some information (statistical or physical) about the signals $s_1(t)$ and $s_2(t)$ is required. This information will be utilized in the design of the control loop. It enables initiation of the separation procedure. If further information is available (such as knowledge, for example, of one of the two bearings) the same algorithm, which then bootstraps itself, will result in performance improvement.

In fact, the model described in this paper is applicable to many communication and signal processing problems which can be put in three groups: bearing estimation (source localization), source separation, and interference cancellation.

- For bearing estimation (source localization) one is interested in \hat{D}_1 and \hat{D}_2 and not in the signal estimates $\hat{s}_1(t)$ and $\hat{s}_2(t)$. In this application, particularly in active radar or sonar, much is known about $s_1(t)$ and $s_2(t)$.

- For source separation, the objective is to get the cleanest possible version of $s_1(t)$ and $s_2(t)$, while D_1 and D_2 are nuisance parameters. Indeed, in some cases there is some prior knowledge about θ_1 and θ_2 which can be used to restrict the possible values of D_1 and D_2 .
- For interference cancellation, the objective is to get a clean replica of one of the signals, say $s_1(t)$ while the other signal, as well as D_1 and D_2 are nuisance parameters which, in some applications, are partially known.

If the signals $s_1(t)$ and $s_2(t)$ are narrowband signals, centered around a known ω_0 , then a complex envelope notation can be used and Eq. (1) can be written as

$$\begin{aligned} z_1 &= e^{-j\omega_0 D_1} s_1 + e^{-j\omega_0 D_2} s_2 + e_1 \\ z_2 &= e^{j\omega_0 D_1} s_1 + e^{j\omega_0 D_2} s_2 + e_2 \end{aligned} \quad (3)$$

so the unknown model parameters are represented by the complex constant $a_{11} = \exp(-j\omega_0 D_1)$, $a_{12} = \exp(-j\omega_0 D_2)$, $a_{21} = a_{11}^* = \exp(j\omega_0 D_1)$ and $a_{22} = a_{12}^* = \exp(j\omega_0 D_2)$. Eq. (3) can then be written in a matrix form: $\underline{z} = A\underline{s} + \underline{e}$, where A is a matrix which is a function of two unknown complex parameters. The linear separation problem resulting from (3) was studied and discussed in many papers[1-5] and bootstrapped algorithms, similar to those proposed in this paper, were shown to be useful in its solution. However, such problem formalization is not applicable for wideband signal environment, where Eq. (3) is no longer a proper representation of Eq. (1). In such a case, the received signals are not a linear function of the unknown model parameters, and the model is inherently non-linear. We will show that the bootstrapped principle used in narrowband signal separation can be adapted to solve the wideband separation problem.

The wideband *DOA* estimation problem, that is of estimating D_1 and D_2 in Eq. (1), has attracted many research efforts in the last decade or so. Different algorithms for *DOA*

estimation were proposed, very few of them are adaptive. However, the estimation of the source signals is rarely considered. Our proposed algorithm can be regarded as an adaptive wideband *DOA* estimation algorithm as well as a signal separator and, in particular, a wideband interference canceller

The paper is organized as follows: We first study the different bootstrapped configurations, we discuss the topology of possible two-inputs two-outputs linear systems which contain variable delay lines and summers only, and we present two configurations for which separation of the signals $s_1(t)$ and $s_2(t)$ is achieved if and only if the variable delays equal D_1 and D_2 of Eq. (1). In Section 3 we study the sensitivity of these two structures to the deviation of the variable delays from the desired solution. We present a small-errors analysis of the second order statistics of the outputs, which shows the separation performance of the adaptive algorithm in a tracking mode of operation. We prove that with both configurations source separation is guaranteed in this mode of operation. Section 4 discuss the adaptive algorithm from the theoretical and practical points of view. We show that if the two signals to be separated do not have exactly the same spectrum, then the algorithm converges to a state of total signal separation with only ambiguity in assigning an output to a signal. We also discuss the effect of additive noise on the steady-state performance of the proposed two versions of the algorithm. Section 5 presents a comparison of the two configurations of the bootstrapped algorithms, draws guidelines for the question of when to use each of them, and emphasizes differences between the proposed algorithm and the one developed for narrow-band sources. Important issues such as stability of the algorithm, rate of convergence, etc. are studied in an ongoing research.

II. THE DIFFERENT CONFIGURATIONS OF THE DELAY CONTROL SEPARATION ALGORITHMS

We suggest that the delay-control separation system consists of pure delay lines and summation devices only. In the steady state the delays in this system are the estimates of the delays in the mixing models (Eq. (1)); that is $\tau_1 = \hat{D}_1$; $\tau_2 = \hat{D}_2$. The outputs of the system, $y_1(t)$ and $y_2(t)$, are used to control the delays τ_1 and τ_2 , following certain optimization criteria.

The frequency-domain mixing model of Eq. (1) can be described by the matrix

$$\mathbf{M}(\omega) = \begin{bmatrix} e^{-j\omega D_1} & e^{-j\omega D_2} \\ e^{j\omega D_1} & e^{j\omega D_2} \end{bmatrix} = \text{func}(D_1, D_2) \quad (4)$$

In open loop, the separation system is a linear system having the transfer function matrix

$$\mathbf{H}(\omega) = \begin{bmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{bmatrix} = \text{func}(\tau_1, \tau_2) \quad (5)$$

Now, since

$$\underline{Y}(\omega) = \mathbf{H}(\omega)\underline{Z}(\omega) = \mathbf{H}(\omega)[\mathbf{M}(\omega)\underline{S}(\omega) + \underline{E}(\omega)] \quad (6)$$

then perfect separation of the signals $s_1(t)$ and $s_2(t)$ is achieved if the matrix $\mathbf{T}(\omega) \triangleq \mathbf{H}(\omega)\mathbf{M}(\omega)$ is a diagonal matrix.

If narrowband signals are assumed (such that the model of Eq. (3) is valid), three different bootstrapped separation structures can be used [2]. These structures are the backward-backward (*BB*) configuration, the forward-forward (*FF*) configuration, and the backward-forward (*BF*) configuration [3]. When adapting the bootstrapped principles to the wideband case we found that the *BF* scheme must be excluded, if the filters $H_1(\omega)$ and $H_2(\omega)$ (see Fig. A1) are required to be pure delay lines (see appendix A).

The *BB structure* discussed in this paper is a special case of a feedback topology structure introduced in [4] for separation on N source signals using an array of $M > N$ sensors. It is shown there, that this structure is a feedback implementation of the least-squares estimates

of the source signals, if the direction of arrivals of the sources are perfectly known. This paper considers the special case where $M = N = 2$. However, its results for the BB configuration based on the topology presented in [4] and the control algorithm of section 5 can be generalized to any $M > N$. The generalization of the FF configuration is still to be studied.

For the two separation structures, the backward/backward and the forward/forward, we first find the transfer function matrix $\mathbf{H}(\omega)$ and use it to exhibit the output power spectrum matrix, and its properties. From [6],

$$\mathbf{S}_{yy}(\omega) = \mathbf{H}(\omega)\mathbf{S}_{zz}(\omega)\mathbf{H}^*(\omega) = \mathbf{T}(\omega)\mathbf{S}_{ss}(\omega)\mathbf{T}^*(\omega) + \mathbf{H}(\omega)\mathbf{S}_{ee}(\omega)\mathbf{H}^*(\omega) \quad (7)$$

where $\mathbf{T}(\omega) = \mathbf{H}(\omega)\mathbf{M}(\omega)$ is the transfer function matrix between the signals to be separated and their estimates.

2.1 The Backward-Backward Configuration

Fig. 2.a depicts the basic structure of the BB configuration. The transfer function matrix of this system is given by:

$$\mathbf{H}_{BB}(\omega) = \frac{j}{2 \sin \omega(\tau_1 - \tau_2)} \begin{bmatrix} e^{j\omega\tau_2} & -e^{-j\omega\tau_2} \\ -e^{j\omega\tau_1} & e^{-j\omega\tau_1} \end{bmatrix} \quad (8)$$

Using Eq. (4) we have

$$\begin{aligned} \mathbf{T}_{BB}(\omega) &= \mathbf{H}_{BB}(\omega)\mathbf{M}(\omega) \\ &= \frac{1}{\sin \omega(\tau_1 - \tau_2)} \begin{bmatrix} \sin \omega(D_1 - \tau_2) & \sin \omega(D_2 - \tau_2) \\ \sin \omega(\tau_1 - D_1) & \sin \omega(\tau_1 - D_2) \end{bmatrix} \end{aligned} \quad (9)$$

From Eq. (9) (and Eq. (6)) it is easy to see that if $\tau_1 = D_1$ and $\tau_2 = D_2$, then $\mathbf{T}_{BB}(\omega) = \mathbf{I}$ and a perfect separation of the signal is achieved. The output signals are then $y_1(t) = s_1(t) + \tilde{n}_1(t)$ and $y_2(t) = s_2(t) + \tilde{n}_2(t)$; where $\tilde{n}_1(t)$ and $\tilde{n}_2(t)$ contain no signal components.

The system of Fig. 2.b is exactly equivalent to that of Fig. 2.a. However, if one can tolerate the outputs to be a delayed versions of the signals to be separated, $\tilde{y}_1(t)$ and $\tilde{y}_2(t)$

can be regarded as the outputs of separation system, which is somewhat less complex than that of Fig. 2.b. For the system of Fig. 2.a, the transfer function matrix between \underline{z} and \underline{y} is given by,

$$\begin{aligned}\tilde{\mathbf{H}}_{BB}(\omega) &= \frac{e^{-j\omega(\tau_1 - \tau_2)}}{-2j \sin \omega(\tau_1 - \tau_2)} \begin{bmatrix} 1 & -e^{-j2\omega\tau_2} \\ -e^{j2\omega\tau_1} & 1 \end{bmatrix} \\ &= \frac{j}{2 \sin \omega(\tau_1 - \tau_2)} \begin{bmatrix} e^{-j\omega(\tau_1 - \tau_2)} & -e^{-j\omega(\tau_1 + \tau_2)} \\ -e^{j\omega(\tau_1 + \tau_2)} & e^{-j\omega(\tau_1 - \tau_2)} \end{bmatrix}\end{aligned}\quad (10)$$

and

$$\tilde{\mathbf{T}}_{BB}(\omega) = \tilde{\mathbf{H}}_{BB}(\omega)\mathbf{M}(\omega) = \frac{1}{\sin \omega(\tau_1 - \tau_2)} \begin{bmatrix} e^{-j\omega\tau_1} \sin \omega(D_1 - \tau_2) & e^{-j\omega\tau_1} \sin \omega(D_2 - \tau_2) \\ e^{j\omega\tau_2} \sin \omega(\tau_1 - D_1) & e^{j\omega\tau_2} \sin \omega(\tau_1 - D_2) \end{bmatrix}\quad (11)$$

$\tilde{\mathbf{T}}_{BB}(\omega)$ is a diagonal matrix for $\tau_1 = D_1$ and $\tau_2 = D_2$, but it is no longer the identity matrix. I .

Assuming that the signals to be separated are uncorrelated, wide sense stationary (WSS) processes with power spectral density (PSD) $S_1(\omega)$ and $S_2(\omega)$, respectively, the PSD matrix of the outputs of the BB configuration is given by:

$$\begin{aligned}\mathbf{S}_{BB}(\omega) &= \begin{bmatrix} S_1(\omega) \left(\frac{\sin \omega(D_1 - \tau_2)}{\sin \omega(\tau_1 - \tau_2)} \right)^2 + S_2(\omega) \left(\frac{\sin \omega(D_2 - \tau_2)}{\sin \omega(\tau_1 - \tau_2)} \right)^2 & S_1(\omega) \frac{\sin \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2)}{\sin \omega(\tau_1 - \tau_2) \sin \omega(\tau_1 - \tau_2)} \\ & + S_2(\omega) \frac{\sin \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2)}{\sin \omega(\tau_1 - \tau_2) \sin \omega(\tau_1 - \tau_2)} \\ S_1(\omega) \frac{\sin \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2)}{\sin \omega(\tau_1 - \tau_2) \sin \omega(\tau_1 - \tau_2)} & \\ + S_2(\omega) \frac{\sin \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2)}{\sin \omega(\tau_1 - \tau_2) \sin \omega(\tau_1 - \tau_2)} & S_1(\omega) \left(\frac{\sin \omega(\tau_1 - D_1)}{\sin \omega(\tau_1 - \tau_2)} \right)^2 + S_2(\omega) \left(\frac{\sin \omega(\tau_1 - D_2)}{\sin \omega(\tau_1 - \tau_2)} \right)^2 \end{bmatrix} \\ &+ \frac{N(\omega)}{2 \sin^2 \omega(\tau_1 - \tau_2)} \begin{bmatrix} 1 & -\cos \omega(\tau_1 - \tau_2) \\ -\cos \omega(\tau_1 - \tau_2) & 1 \end{bmatrix}\end{aligned}\quad (12)$$

Eq. (12) is derived from Eq. (7) using $\tilde{\mathbf{H}}_{BB}(\omega)$ and $\tilde{\mathbf{T}}_{BB}(\omega)$ of Eqs. (8) and (9), where we assumed that the noise processes are uncorrelated and have equal spectra $S_{e_1}(\omega) = S_{e_2}(\omega) = N(\omega)$. It can be shown that $\tilde{\mathbf{S}}_{BB}(\omega)$, which is related to $\tilde{\mathbf{H}}_{BB}(\omega)$ and $\tilde{\mathbf{T}}_{BB}(\omega)$ of Eqs. (10) and (11) is very similar. The only difference is that the off-diagonal terms are multiplied by $e^{-j\omega(\tau_1 + \tau_2)}$.

From Eq. (12), the output *PSD* at one of the outputs, say y_1 , is given by,

$$S_{y_1}(\omega) = S_1(\omega) \left(\frac{\sin \omega(D_1 - \tau_2)}{\sin \omega(\tau_1 - \tau_2)} \right)^2 + S_2(\omega) \left(\frac{\sin \omega(D_2 - \tau_2)}{\sin \omega(\tau_1 - \tau_2)} \right)^2 + \frac{N(\omega)}{2 \sin^2 \omega(\tau_1 - \tau_2)} \quad (13)$$

and a similar term for the power spectral density of the other output. The cross spectrum between the outputs consist of three terms each related, respectively, to the signals to be separated, $s_1(t)$, $s_2(t)$, and to the additive noise. Note from (13) that at the output, $y_1(t)$ the term related to $s_1(t)$, for example, can be viewed as having resulted from passing $s_1(t)$ via a filter with transfer function $\frac{\sin \omega(D_1 - \tau_2)}{\sin \omega(\tau_1 - \tau_2)}$. This transfer function is highly non-linear in the unknown parameter(s) and their estimate(s). Therefore, although a pure delay line is a special case of a linear filter, the problem discussed here is very different from the linear model in [5-7] and in [8], where the unknown linear filters are assumed to be rational (*FIR* or *IIR*) filters of unknown coefficients.

Also note that the possibility of $\tau_1 = \tau_2$ must be avoided, and that at $\omega = 0$ the output noise term diverges. Therefore, this *BB* configuration can only be used with band-pass signals that impinge on the array from directions not very close to one another.

2.2 The Forward-Forward Configuration

Fig. 2.c depicts the basic structure of the *FF* configuration. The transfer function of this system is given by,

$$\mathbf{H}_{FF}(\omega) = \begin{bmatrix} e^{j\omega\tau_2} & -e^{-j\omega\tau_2} \\ -e^{j\omega\tau_1} & e^{-j\omega\tau_1} \end{bmatrix} = -2j \sin \omega(\tau_1 - \tau_2) \mathbf{H}_{BB}(\omega) \quad (14)$$

Using Eq. (4) we have,

$$\mathbf{T}_{FF}(\omega) = \mathbf{H}_{FF}(\omega) \mathbf{M}(\omega) = -2j \begin{bmatrix} \sin \omega(D_1 - \tau_2) & \sin \omega(D_2 - \tau_2) \\ \sin \omega(\tau_1 - D_1) & \sin \omega(\tau_1 - D_2) \end{bmatrix} = -2j \sin \omega(\tau_1 - \tau_2) \mathbf{T}_{BB}(\omega) \quad (15)$$

Note again that if $\tau_1 = D_1$, $\tau_2 = D_2$ the two signals are indeed separated. However, each of the two outputs rather than being a delayed version of the corresponding input, consists

of superposition of delayed versions of the separated inputs. In fact, $y_1(t) = s_1(t - \Delta) - s_1(t + \Delta) + \tilde{n}_1(t)$ and $y_2(t) = s_2(t - \Delta) - s_2(t + \Delta) + \tilde{n}_2(t)$, where $\Delta = D_1 - D_2$ and $\tilde{n}_1(t)$ and $\tilde{n}_2(t)$ contain no signal components. Thus the separated outputs are distorted versions of the input signals, and for applications where such a distortion cannot be accepted, the *BB* configuration should be used. On the other hand, if the signals' waveforms are nuisance parameters and one is mostly interested in the unknown delays D_1 and D_2 , then the *FF* configuration of Fig. 2.c, or its equivalent simpler version Fig. 2.d can be used. In the system of Fig. 2.d, the outputs $\tilde{y}_1(t)$ and $\tilde{y}_2(t)$ decouple the inputs if $D_1 = \tau_1$ and $D_2 = \tau_2$, so $\tilde{y}_1(t)$ contains no part of $s_2(t)$ and $\tilde{y}_2(t)$ contains no part of $s_1(t)$. Both outputs, however, contain distorted versions of the decoupled inputs.

The power spectral density matrix for the *FF* configuration is given by (see Eq. (7)). From Fig. 2.c,

$$\mathbf{S}_{FF}(\omega) = 4 \begin{bmatrix} S_1(\omega) \sin^2 \omega(D_1 - \tau_2) & S_1(\omega) \sin \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2) \\ + S_2(\omega) \sin^2 \omega(D_2 - \tau_2) & + S_2(\omega) \sin \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2) \\ S_1(\omega) \sin \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2) & S_1(\omega) \sin^2 \omega(D_1 - \tau_1) \\ + S_2(\omega) \sin \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2) & + S_2(\omega) \sin^2 \omega(\tau_1 - D_2) \end{bmatrix} + 2N(\omega) \begin{bmatrix} 1 & -\cos \omega(\tau_1 - \tau_2) \\ -\cos \omega(\tau_1 - \tau_2) & 1 \end{bmatrix} = 4 \sin^2 \omega(\tau_1 - \tau_2) \mathbf{S}_{BB}(\omega) \quad (16)$$

By comparing (16) with (12) we notice that the power spectrum and cross spectrum of the outputs of the *BB* and *FF* configurations differ only by $4 \sin^2 \omega(\tau_1 - \tau_2)$. The *FF* configuration can be regarded as containing zeros-only, or FIR system, while the *BB* has poles and zeros, or IIR system. As such, the *FF* configuration can handle low-pass signals too, and is not sensitive to situations where $\tau_1 = \tau_2$. Also notice that the *BB* configuration can be looked upon as an *FF* configuration, followed by a 2×2 system with a transfer function matrix $\frac{1}{4 \sin^2 \omega(\tau_1 - \tau_2)} I$, where I is the 2-dimensional identity matrix.

III. SMALL-ERRORS ANALYSIS OF THE OUTPUT SPECTRUM

In both *FF* and *BB* structures, $\tau_1 = D_1$ and $\tau_2 = D_2$ result in perfect separation of the source signal. If D_1 and D_2 are unknown, they should be estimated. To study the effect of the estimation errors on the separation performance of the two configurations we next find the cross-spectrum of the outputs under small error assumption; (i.e., $|\psi_i| = |\tau_i - D_i| \ll |D_i|$, $i = 1, 2$).

3.1 The Output Power Spectral Density Matrix

We look at the linear approximation of the output power spectral density matrix. i.e.:

$$\mathbf{S}_y(\omega) = \mathbf{S}_y(\omega) |_{\psi_1=\psi_2=0} + \psi_1 \left[\frac{\partial \mathbf{S}_y(\omega)}{\partial \psi_1} \Big|_{\psi_1=\psi_2=0} \right] + \psi_2 \left[\frac{\partial \mathbf{S}_y(\omega)}{\partial \psi_2} \Big|_{\psi_1=\psi_2=0} \right] \quad (17)$$

where $\psi_1 = \tau_1 - D_1$, and $\psi_2 = \tau_2 - D_2$ are sufficiently small.

From Appendix B we know that the linear approximation of Eq. (12) is

$$\begin{aligned} \mathbf{S}_{BB}(\omega) \approx & \begin{bmatrix} S_1(\omega)(1 - \psi_1 \frac{2\omega}{tg(\omega\Delta)}) & (\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) \frac{\omega}{\sin(\omega\Delta)} \\ (\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) \frac{\omega}{\sin(\omega\Delta)} & S_2(\omega)(1 + \psi_2 \frac{2\omega}{tg(\omega\Delta)}) \end{bmatrix} \\ & + \frac{N(\omega)}{2 \sin^2(\omega\Delta)} \begin{bmatrix} 1 + \frac{2\omega}{tg(\omega\Delta)}(-\psi_1 + \psi_2) & -\cos(\omega\Delta) \\ -\cos(\omega\Delta) & + \frac{\omega}{\sin(\omega\Delta)}(1 + \cos^2(\omega\Delta))(-\psi_1 + \psi_2) \\ + \frac{\omega}{\sin(\omega\Delta)}(1 + \cos^2(\omega\Delta))(-\psi_1 + \psi_2) & 1 + \frac{2\omega}{tg\omega\Delta}(-\psi_1 + \psi_2) \end{bmatrix} \end{aligned} \quad (18)$$

where $\Delta = D_1 - D_2$. Similarly, for the *FF* configuration, applying Eq. (17) on Eq. (16) yields (see Appendix B):

$$\begin{aligned} \mathbf{S}_{FF}(\omega) \approx & 4 \begin{bmatrix} S_1(\omega)[\sin^2(\omega\Delta) - \psi_2 \omega \sin 2(\omega\Delta)] & \omega \sin(\omega\Delta)(\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) \\ \omega \sin(\omega\Delta)(\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) & S_2(\omega)[\sin^2(\omega\Delta) + \psi_1 \omega \sin 2(\omega\Delta)] \end{bmatrix} \\ & + 2N(\omega) \begin{bmatrix} 1 & (\psi_1 - \psi_2)\omega \sin(\omega\Delta) - \cos(\omega\Delta) \\ (\psi_1 - \psi_2)\omega \sin(\omega\Delta) - \cos(\omega\Delta) & 1 \end{bmatrix} \end{aligned} \quad (19)$$

Note that for both configurations, small errors in the estimates of D_1 and of D_2 , such that $|\frac{\psi_1}{D_1}| \approx 0, |\frac{\psi_2}{D_2}| \approx 0$ but not identically zero, still guarantee separation of the two signals. The power spectral density of each of the outputs (diagonal elements in (18) or (19)) contains a filtered version of only one of the signals (plus noise).

From Eqs. (18) and (19) we notice that:

1. For $\psi_1 = \psi_2 = 0$ the signals $s_1(t)$ and $s_2(t)$ are separated so that at $y_1(t)$ we have no component of $s_2(t)$ and at $y_2(t)$ we have no component of $s_1(t)$. This holds in both configurations and is still true if $\psi_1 \approx 0$ and $\psi_2 \approx 0$. However while for $\psi_1 = \psi_2 = 0$ the output of the *BB* configuration consists of an exact replica of the input signal, for $\psi_1 \approx 0$ and $\psi_2 \approx 0$, this separated signal is a filtered, distorted replica of the input signal. For the *FF* configuration the output signal is a filtered version of the input, when ψ_1 and ψ_2 are almost zero (small errors) or exactly zero.
2. If $N(\omega) = 0$ (no additive noise), then $\psi_1 = \psi_2 = 0$ guarantee uncorrelated outputs ($S_{y_1 y_2}(\omega) = 0$ in both configurations). If neither ψ_1 nor ψ_2 is zero the outputs are, in general, correlated.
3. In the presence of noise, the outputs are correlated even if $\psi_1 = \psi_2 = 0$.

3.2 The Output SNR

With perfect delay estimation; $\psi_1 = \psi_2 = 0$, the output signal-to noise ratios for the two structures are given, respectively, by

$$\text{SNR}_{o_i}^{BB} = \frac{\int_W S_i(\omega) d\omega}{\int_W \frac{N(\omega)}{2 \sin^2(\omega \Delta)} d\omega} \quad (20)$$

$$\text{SNR}_{o_i}^{FF} = \frac{2 \int_W S_i(\omega) \sin^2(\omega \Delta) d\omega}{\int_W N(\omega) d\omega} \quad (21)$$

where W is the processing bandwidth. The inputs SNR are given by:

$$SNR_{in}^i = \frac{\int_W S_i(\omega) d\omega}{\int_W N(\omega) d\omega} \quad ; \quad i = 1, 2 \quad (22)$$

We notice that for both configurations, the cancellation of the other signal (interference) causes degradation in the SNR —the output SNR is worse than the input SNR (but, hopefully better than the input $SINR$ - the signal to interference plus noise ratio!). From Eqs. (20) and (21) we see that as $\Delta = D_1 - D_2$ is smaller (the spatial separation between the sources is smaller), the output SNR is smaller. Notice, however, that the output SNR is different for the two configurations.

For the case of flat spectral density, $S_i(\omega) = P_i$ and $N(\omega) = P_N$ over $\omega_0 - W/2 \leq |\omega| \leq \omega_0 + W/2$, we have for the inputs SNR , $SNR_i = \frac{P_i}{P_N} = \rho_i$, $i = 1, 2$ and for the outputs SNR is (see Appendix C),

$$SNR_{o_i}^{BB} = \frac{2\rho_i W}{\int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \frac{1}{\sin^2(\omega\Delta)} d\omega} = \frac{\rho_i(\cos \Delta W - \cos 2\omega_0)}{\frac{\sin W\Delta}{W\Delta}} \quad ; \quad i = 1, 2 \quad (23)$$

$$SNR_{o_i}^{FF} = \frac{\rho_i}{W} \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \sin^2(\omega\Delta) d\omega = \rho_i(1 - \cos 2\omega_0\Delta) \frac{\sin W\Delta}{W\Delta} \quad ; \quad i = 1, 2 \quad (24)$$

Notice that if $W \ll \omega_0$ (a narrowband assumption), Eqs. (20) and (21) yield.

$$\begin{aligned} SNR_{o_i}^{BB} &\approx \frac{2S_i(\omega_0)}{\frac{N(\omega_0)}{\sin^2 \omega_0 \Delta}} \quad ; \quad SNR_{o_i}^{FF} \approx \frac{2S_i(\omega_0) \sin^2 \omega_0 \Delta}{N(\omega_0)} \\ \Rightarrow SNR_{o_i}^{FF} &= SNR_{o_i}^{BB} = 2 \sin^2 \omega_0 \Delta SNR_{in}^i \quad ; \quad i = 1, 2 \end{aligned} \quad (25)$$

That is, under perfect separation and narrowband assumption, the outputs SNR is the same for the two configurations, while, as can be seen from Eqs. (23) and (24), in the wideband case it can be very much different. Also notice that in the narrowband case, the outputs SNR satisfies $SNR_{o_i} = \gamma SNR_{in}^i$ where $0 \leq \gamma \leq 0.5$, so the SNR degradation is, at least 3 dB. In the wideband case, depending on the relative bandwidth, the maximum degradation in SNR can be smaller. Clearly, due to the perfect cancellation of the other signal at each of

the outputs, the outputs SNR are the same as the output signal-to-interference-plus-noise-ratio ($SINR$). Eqs. (23)-(25) show that the outputs $SINR$ is a function of the input SNR (not the input $SINR$, which is smaller) and of the weighted spatial separation between the sources, Δ .

As suggested by Eqs. (18) and (19), the output SNR in a "small errors" scenario, i.e., $\psi_1 \approx 0$ and $\psi_2 \approx 0$, are different from those obtained with no errors ($\psi_1 = \psi_2 = 0$), even though in both cases the signal-to-interference ratio (SIR) in the two outputs are infinite (perfect separation).

The important conclusion from this subsection is that when signal separation is achieved, the "cost" in degrading the output SNR can be severe. Unless the other signal (the interference) input SNR is high, it is not guaranteed that the output SNR is indeed larger than the input $SINR$ (signal to interference + noise ratio), even in the ideal case where $\psi_1 = \psi_2 = 0$.

IV. OPTIMIZATION CRITERIA AND SEPARATION ADAPTIVE ALGORITHMS

A signal separation problem similar to the one addressed in this paper, and which is highly related to a multi-channel identification problem, is widely considered lately under different names and/or scenarios (e.g. [8]). Most optimization criteria used for controlling the algorithms are related to the assumptions that the signals to be separated are statistically independent and that at least one of them is a non-Gaussian process[5-7]. Based on these assumptions a family of higher-order spectra (HOS) optimization criteria are proposed. The spatial separation problem addressed in the present paper is often related to cases where the signals to be separated are both Gaussian (e.g., passive localization in sonar). The signals can be assumed uncorrelated (or statistically independent) but no higher-order cumulants (or spectrum) can be used as optimization criteria. Decorrelation of the output signals, which is indeed a necessary condition for separation when no additive noise is present, can be used as

an optimization criterion. In the sequel, we study how outputs' decorrelation can be made a sufficient condition for source separation in our problem. This is a special case of [8], in which the outputs' decorrelation criterion was used for any linear-filtered combination of the two signals. However, the case we study, where the combination filters are pure delay lines, is of special interest because of the related applications. Also, the non-linear dependencies of the filter(s) on the unknown parameter(s) τ make the use of existing results in our problem, far from being trivial.

Assuming that the inputs are zero-mean processes, that no-noise is present ($V(\omega) = 0$), and that $\tau_1 \neq \tau_2$, then from Eq. (12) or (16) we conclude that for both *BB* and *FF* configurations the output signals are decorrelated if and only if

$$S_1(\omega) \sin \omega \psi_1 \sin \omega(\Delta - \psi_2) = S_2(\omega) \sin \omega \psi_2 \sin \omega(\Delta + \psi_1) : \forall \omega \quad (26)$$

It is simple to note that (26) will be satisfied for every ω and hence the outputs will be decorrelated if $\psi_1 = \psi_2 = 0$. On the other hand if $\psi_1 \neq 0$ and $\psi_2 \neq 0$ we want to find conditions under which this equation is satisfied, for any ω .

With the small errors approximation of Eq. (18) or (19), we get, instead of Eq. (26) :

$$\psi_1 S_1(\omega) = \psi_2 S_2(\omega) \quad (27)$$

Since ψ_1 and ψ_2 are to be constants (in the steady state of any algorithm), Eq. (27) is satisfied for $\psi_1 \neq 0, \psi_2 \neq 0$ if and only if $S_1(\omega) = \alpha S_2(\omega)$. That is :

If the signals have the same spectral shape, the outputs can be decorrelated with ψ_1 and ψ_2 being non-zero, but small, only if $\psi_1/\psi_2 = 1/\rho_{12}$, ρ_{12} being the power ratio between the signal $s_1(t)$ and the signal $s_2(t)$.

If $S_1(\omega) \neq \alpha S_2(\omega)$, the only solution to Eq. (27) for every ω is $\psi_1 = \psi_2 = 0$. Thus, we conclude that, at a small error scenario, then if the input spectra are of different shape (and

that $N(\omega) = 0$). a necessary and sufficient condition for $\psi_1 = 0$ and $\psi_2 = 0$ is decorrelation of the output signal.

Is this result valid also if no small-errors assumption is used? Given that $S_1(\omega) \neq \alpha S_2(\omega)$, then beside $\psi_1 = \psi_2 = 0$ (26) is satisfied only if

1. $\psi_1 = 0, \psi_2 \neq 0$ and $\Delta = -\psi_1 = 0$ which is equivalent to $D_1 = D_2$.
2. $\psi_1 = 0, \psi_2 \neq 0$ and $\Delta = \psi_1 = 0$ which is again equivalent to $D_1 = D_2$.
3. $\Delta + \psi_1 = 0, \Delta - \psi_2 \neq 0$ which is equivalent to $\Delta = \psi_2 = -\psi_1$. This condition leads to $\tau_1 = D_2$ and $\tau_2 = D_1$. That is, signal separation in the opposite direction ($y_1(t)$ is proportional to $s_2(t)$ and $y_2(t)$ is proportional to $s_1(t)$).

This ambiguity is inherent in the model identification problem[e.g., 8].

Thus, a proper criterion for separation of two uncorrelated Gaussian signals of different spectra is decorrelation of the separation outputs.

If the signals have the same spectral shape (so $S_1(\omega) = \alpha S_2(\omega)$), a discriminator, which exploit slight prior information about distinguishability of the signals to be separated, can be used to trigger the algorithm as suggested in [1] for narrowband signals.

4.1 Decorrelation Algorithm

Decorrelation can be performed both in frequency domain (imposing $S_{y_1 y_2}(\omega) = 0, \forall \omega$) or in the correlation-lag (time) domain (imposing $R_{y_1 y_2}(\alpha) = 0, \forall \alpha$)². Because of availability of hardware components, we are interested in time domain algorithms. The algorithm adaptively controls τ_1 and τ_2 to reach of outputs decorrelation. That is, we need to solve two delay control equations simultaneously, in order to obtain their unique solution $\psi_1 = \psi_2 = 0$.

²Notice that for the two configurations (*BB* and *FF*) the frequency domain requirements result in the same equation, while the time domain requirements are inherently different.

The time domain decorrelation criterion introduces an *infinite* number of equations in the form:

$$\mathbf{E}\{y_1(t)y_2(t - \alpha_i)\} = R_{y_1 y_2}(\alpha_i, \tau_1, \tau_2) = 0 \quad -\infty < \alpha_i < \infty \quad (28)$$

(\mathbf{E} stands for statistical expectation). For any choice of α_1 and α_2 we can solve simultaneously

$$R_{y_1 y_2}(\alpha_1; \tau_1, \tau_2) = 0 \quad (29)$$

$$R_{y_1 y_2}(\alpha_2; \tau_1, \tau_2) = 0 \quad (30)$$

for τ_1 and τ_2 . The choice of α_1, α_2 should be such that Eqs. (29) and (30) are linearly independent.

Under small-errors conditions and no-noise assumption, then using Eq. (18) for the *BB* configuration, Eqs. (29), (30) become:

$$\psi_1 \int_{-\infty}^{\infty} \frac{\omega S_1(\omega)}{\sin(\omega\Delta)} e^{j\omega\alpha_1} d\omega - \psi_2 \int_{-\infty}^{\infty} \frac{\omega S_2(\omega)}{\sin(\omega\Delta)} e^{j\omega\alpha_1} d\omega = 0 \quad (31)$$

$$\psi_1 \int_{-\infty}^{\infty} \frac{\omega S_1(\omega)}{\sin(\omega\Delta)} e^{j\omega\alpha_2} d\omega - \psi_2 \int_{-\infty}^{\infty} \frac{\omega S_2(\omega)}{\sin(\omega\Delta)} e^{j\omega\alpha_2} d\omega = 0 \quad (32)$$

If $\alpha_1 \neq \alpha_2$ then $\cos \omega\alpha_1 = \cos \omega\alpha_2$ only for discrete point of ω which has a measure zero on the real axis ω . Therefore, we have $\cos \omega\alpha_1 \neq \cos \omega\alpha_2$ on set of intervals with a positive measure, and hence the set of equations (31) and (30) are linearly independent. Finally we conclude that if the algorithm brings the controlled delays to the vicinity of $\psi_1 = \psi_2 = 0$, then the algorithm converges to these values. Similar argument apply to the *FF* configuration and equation (19).

The proposed control algorithm is depicted in Fig. 3. From which we note that it reaches a singular point only if $\hat{R}_{y_1 y_2}(\alpha_1) = 0$ and similarly for τ_2 . Assuming that ω_0 is the center frequency of the processing bandwidth we suggest, as a rule of thumb, to choose $\alpha_1 = 0$ and $\alpha_2 = \frac{\pi}{\omega_0}$.

4.2 Practical implementation considerations

1. Practically, α_1 and α_2 in Fig. 3 can be changed during the adaptation. The choice of these delays may have great effect on the rate of convergence, even if the steady-state values are the same. This subject is worth further investigation.
2. For the *BB* configuration and the narrowband case, decorrelation criterion can be replaced by minimization of the output powers[3]. In the more general wideband case, one can show (see Appendix D) that decorrelation of the outputs is a sufficient condition, but not a necessary one for minimum power, so that minimum power criteria can not replace decorrelation. In particular, for the delay control case since the output power is given by,

$$P_i = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{y_i y_i}(\omega) d\omega \quad i = 1, 2 \quad (33)$$

a necessary condition for minimum power is $\frac{\partial P_i}{\partial \tau_j} = 0$, $i \neq j$ $i, j = 1, 2$ for which it is sufficient to have:

$$\frac{\partial}{\partial \tau_2} S_{y_1 y_1}(\omega) = 0 \quad ; \quad \frac{\partial}{\partial \tau_1} S_{y_2 y_2}(\omega) = 0 \quad ; \quad \forall \omega \quad (34)$$

From Eq. (12), with $N(\omega) = 0$, we get,

$$\begin{aligned} \frac{\partial}{\partial \tau_2} S_{y_1 y_1}(\omega) &= -\frac{2\omega}{\sin \omega(\tau_1 - \tau_2)} \left[S_1(\omega) \frac{\sin \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2)}{\sin \omega(\tau_1 - \tau_2) \sin \omega(\tau_1 - \tau_2)} \right. \\ &\quad \left. + S_2(\omega) \frac{\sin \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2)}{\sin \omega(\tau_1 - \tau_2) \sin \omega(\tau_1 - \tau_2)} \right] \\ &= -\frac{2\omega}{\sin \omega(\tau_1 - \tau_2)} S_{y_2 y_1}(\omega) \end{aligned} \quad (35)$$

and similarly,

$$\frac{\partial}{\partial \tau_1} S_{y_2 y_2}(\omega) = \frac{2\omega}{\sin \omega(\tau_1 - \tau_2)} S_{y_1 y_2}(\omega) \quad (36)$$

We see that imposing decorrelation of the output signals ($S_{y_2 y_1}(\omega) = S_{y_1 y_2}(\omega) = 0$, for every ω) is equivalent to imposing Eq. (34). Using Eq. (35) or Eq. (36) in Eq. (33) we have,

$$\frac{\partial P_i}{\partial \tau_j} = \frac{(-1)^i}{\pi} \int_{-\infty}^{\infty} \frac{\omega S_{y_i y_j}(\omega)}{\sin \omega(\tau_1 - \tau_2)} d\omega \quad ; \quad i \neq j \quad i = 1, 2 \quad (37)$$

Thus, decorrelation of the outputs, i.e., $S_{y_1 y_2}(\omega) = 0$ for all ω , is sufficient, but not necessary, for minimum power. Therefore, the criterion of minimum power used in the narrowband case, does not necessarily leads to the desired separation in our case. [A special case is that when $S_{y_1 y_2}(\omega) = \text{constant}$, i.e. - the cross spectrum is flat over the frequency band of interest then decorrelation is a necessary and sufficient condition for minimum power].

3. In practice, decorrelation of the outputs \tilde{y}_1 and \tilde{y}_2 in Figs. 2.b and/or 2.d is sufficient for separation, and therefore these outputs can be used as inputs to the control algorithm of Fig. 3. If one can bare constant delay of the outputs then the part of the scheme which uses \tilde{y}_i as inputs to produce y_i can be removed, resulting in a simpler separation scheme.
4. The optimization criterion used assumes no noise ($N(\omega) = 0$). However, the noise can have dramatic effect on the performance of the algorithm as in the narrowband case [3]. We discuss this effect under the "small errors" assumption:

By comparing the off-diagonal entries of Eqs. (18) and (19) we see that at $\psi_1 \approx 0, \psi_2 \approx 0$ the cross-spectrum function of the separator output is approximately zero for the two configurations as long as $N(\omega) = 0$ and $\sin(\omega\Delta) \neq 0$. However, the effect of the additive noise is dramatically different for the two schemes. The average cross-power noise for the FF configuration is given by,

$$N_{12}^{FF} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} 2N(\omega)[\cos(\omega\Delta) + (\psi_2 - \psi_1)\omega \sin(\omega\Delta)] d\omega \quad (38)$$

while for the *BB* configuration it is given by

$$N_{12}^{BB} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{N(\omega)}{2 \sin^2(\omega\Delta)} [\cos(\omega\Delta) + (\psi_2 - \psi_1) \frac{\omega}{\sin(\omega\Delta)} (1 + \cos^2(\omega\Delta))] d\omega \quad (39)$$

We mainly notice that in the cases where Δ is small (which is related to a "high resolution" scenario) the effect of the noise on the performance of any output is more dramatic in the *BB* structure than for the *FF* one. As in the case of narrowband signals[3], the presence of an additive noise causes bias to the estimates of D_1 and D_2 . If we let $\tau_1 = \hat{D}_1$ and $\tau_2 = \hat{D}_2$ be these estimates:

$$\tau_1 = \hat{D}_1 = D_1 + \epsilon_1 \quad (40)$$

$$\tau_2 = \hat{D}_2 = D_2 + \epsilon_2 \quad (41)$$

then ϵ_1 and ϵ_2 are non-zero mean random processes. Their mean is proportional to the average cross-power noise of Eqs. (38) and (39), and is larger for the *BB* configuration than for the *FF*.

V. CONCLUSIONS AND DISCUSSION

In this paper we presented two bootstrapped-like algorithms for spatial separation of wideband sources. We show that by seeking decorrelation of the outputs of either a symmetric feedback or a forward structure we achieve both separation of the source signals and estimation of their relative delays. An adaptive bootstrapped algorithm for such output decorrelation is proposed and its performance for source separation is evaluated using a small error analysis. We show that if tracking is achieved (i.e., $\tau_1 \approx D_1$ and $\tau_2 \approx D_2$), perfect separation of the source signals is performed at the cost of increasing *SNR* and signal distortions. These distortions can be compensated for by standard equalization methods. The effect of delay estimation errors caused by the presence of additive noise at the inputs is studied and is shown to be potentially harmful.

Two different configurations of the algorithm, denoted by *BB* for backward-backward and *FF* for forward-forward, are discussed. While both are able to perform source separation and delay estimation, they exhibit several differences :

1. In the ideal situation, i.e., when $\psi_1 = \psi_2 = 0$, the *BB* configuration provides undistorted versions of the signal waveforms, while the *FF* configuration decouples the signals but provides, at the outputs, a filtered version of them. Therefore, if an exact replica of the source signal is needed, the *BB* configuration is to be preferred.
2. In the presence of additive noise, the *FF* configuration provides better estimates of the unknown source directions (the delays) than the *BB* configuration. Thus, the *FF* structure is better for direction of arrival (DOA) estimation.
3. The *BB* configuration exhibits difficulties in the presence of noise when applied at baseband (see (39)).

In comparison to narrowband separators [1-3], we conclude the following:

- (a) The mixed forward-backward configuration [2] can only be implemented in the narrowband case.
- (b) Unlike the narrowband in the broadband case decorrelation cannot be replaced by minimization of the output powers, when backward-backward separator is used.
- (c) If the signals are only known to be uncorrelated and have the same spectrum, then, as in the narrowband case, a discriminator is needed to achieve separation. If, however, their spectrum is known to be different, then such a discriminator is unnecessary.

Further study of the control algorithm under different scenarios is currently being pursued. Also under current examination is the question of generalization of the separation

algorithm to N inputs and M sensors, $M > N$. For this it was already shown in [4] that the BB configuration can be naturally generalized and it results in a least-square estimate of the signal waveforms. The generalization of the FF configuration and of the control algorithms is still ongoing.

VI. APPENDICES

Appendix A

The transfer function matrix for the FB configuration is given by (see Fig. A.1):

$$\mathbf{H}(\omega) = \begin{bmatrix} 1 + H_1(\omega)H_2(\omega) & -H_2(\omega) \\ -H_1(\omega) & 1 \end{bmatrix} \quad (\text{A-1})$$

Thus, $\mathbf{T}(\omega) = \mathbf{M}(\omega)\mathbf{H}(\omega)$; the transfer function matrix between the signals to be separated and their estimates are:

$$\begin{aligned} \mathbf{T}(\omega) &= \begin{bmatrix} e^{-j\omega D_1} & e^{-j\omega D_2} \\ e^{j\omega D_1} & e^{j\omega D_2} \end{bmatrix} \mathbf{H}(\omega) \\ &= \begin{bmatrix} e^{-j\omega D_1}(1 + H_1(\omega)H_2(\omega)) - H_1(\omega)e^{-j\omega D_2} & -H_2(\omega)e^{-j\omega D_1} + e^{-j\omega D_2} \\ e^{j\omega D_1}(1 + H_1(\omega)H_2(\omega)) - H_1(\omega)e^{j\omega D_2} & -H_2(\omega)e^{j\omega D_1} + e^{j\omega D_2} \end{bmatrix} \quad (\text{A-2}) \end{aligned}$$

For signal separation we need $\mathbf{T}(\omega)$ to be diagonal, so one needs:

$$H_2(\omega) = e^{j\omega(D_1 - D_2)} \quad (\text{A-3})$$

$$H_1(\omega) = e^{j\omega(D_1 - D_2)}(1 + H_1(\omega)e^{j\omega(D_1 - D_2)}) \quad (\text{A-4})$$

or

$$\begin{aligned} H_1(\omega) &= \frac{e^{j\omega(D_1 - D_2)}}{1 - e^{j2\omega(D_1 - D_2)}} = \frac{1}{e^{-j\omega(D_1 - D_2)} - e^{j\omega(D_1 - D_2)}} = \frac{1}{-2j \sin \omega(D_1 - D_2)} \\ &= \frac{j}{2 \sin \omega(D_1 - D_2)} \quad (\text{A-5}) \end{aligned}$$

Notice that $H_1(\omega)$ can not be implemented by a simple pure delay line, so the FB structure cannot be considered if this constraint is valid.

Appendix B

From Eq. (16),

$$\mathbf{S}_{FF}(\omega) = 4 \begin{bmatrix} S_1(\omega) \sin^2 \omega(D_1 - \tau_2) & S_1(\omega) \sin \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2) \\ + S_2(\omega) \sin^2 \omega(D_2 - \tau_2) & + S_2(\omega) \sin \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2) \end{bmatrix} \\ + 2N(\omega) \begin{bmatrix} 1 & -\cos \omega(\tau_1 - \tau_2) \\ -\cos \omega(\tau_1 - \tau_2) & 1 \end{bmatrix} \quad (\text{B-1})$$

Thus

$$\mathbf{S}_{FF}(\omega) |_{\psi_1=\psi_2=0} = \begin{bmatrix} 4S_1(\omega) \sin^2(\omega\Delta) + 2N(\omega) & -2N(\omega) \cos(\omega\Delta) \\ -2N(\omega) \cos(\omega\Delta) & 4S_2(\omega) \sin^2(\omega\Delta) + 2N(\omega) \end{bmatrix} \\ = 4 \sin^2(\omega\Delta) \begin{bmatrix} S_1(\omega) & 0 \\ 0 & S_2(\omega) \end{bmatrix} + 2N(\omega) \begin{bmatrix} 1 & -\cos(\omega\Delta) \\ -\cos(\omega\Delta) & 1 \end{bmatrix} \quad (\text{B-2})$$

From Eq. (B-1)

$$\frac{\partial \mathbf{S}_{FF}(\omega)}{\partial \tau_1} = 4 \begin{bmatrix} 0 & \omega \cos \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2) S_1(\omega) \\ \omega \cos \omega(\tau_1 - D_1) \sin \omega(D_1 - \tau_2) S_1(\omega) & + \omega \sin \omega(D_2 - \tau_2) \cos \omega(\tau_1 - D_2) S_2(\omega) \end{bmatrix} \\ + 2N(\omega) \begin{bmatrix} 0 & \omega \sin \omega(\tau_1 - \tau_2) \\ \omega \sin \omega(\tau_1 - \tau_2) & 0 \end{bmatrix} \quad (\text{B-3})$$

so

$$\frac{\partial \mathbf{S}_{FF}(\omega)}{\partial \tau_1} |_{\psi_1=\psi_2=0} = 4 \begin{bmatrix} 0 & \omega \sin(\omega\Delta) S_1(\omega) \\ \omega \sin(\omega\Delta) S_1(\omega) & 2\omega \sin(\omega\Delta) \cos(\omega\Delta) S_2(\omega) \end{bmatrix}$$

$$+ 2N(\omega) \begin{bmatrix} 0 & \omega \sin(\omega\Delta) \\ \omega \sin(\omega\Delta) & 0 \end{bmatrix} \quad (\text{B-4})$$

Also,

$$\frac{\partial S_{FF}(\omega)}{\partial \tau_2} = 4 \begin{bmatrix} -2\omega S_1(\omega) \sin \omega(D_1 - \tau_2) \cos \omega(D_1 - \tau_2) & -\omega S_1(\omega) \sin \omega(\tau_1 - D_1) \cos \omega(D_1 - \tau_2) \\ -2\omega S_2(\omega) \sin \omega(D_2 - \tau_2) \cos \omega(D_2 - \tau_2) & -\omega S_2(\omega) \cos \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2) \\ -\omega S_1(\omega) \sin \omega(\tau_1 - D_1) \cos \omega(D_1 - \tau_2) & 0 \\ -\omega S_2(\omega) \cos \omega(D_2 - \tau_2) \sin \omega(\tau_1 - D_2) & 0 \end{bmatrix} \\ + 2N(\omega) \begin{bmatrix} 0 & -\omega \sin(\omega(\tau_1 - \tau_2)) \\ -\omega \sin(\omega(\tau_1 - \tau_2)) & 0 \end{bmatrix} \quad (\text{B-5})$$

and

$$\frac{\partial S_{FF}(\omega)}{\partial \tau_2} \Big|_{\psi_1=\psi_2=0} = 4 \begin{bmatrix} -2\omega S_1 \sin(\omega\Delta) \cos(\omega\Delta) & -\omega S_2 \sin(\omega\Delta) \\ -\omega S_2 \sin \omega\Delta & 0 \end{bmatrix} - 2N(\omega) \begin{bmatrix} 0 & \omega \sin(\omega\Delta) \\ \omega \sin(\omega\Delta) & 0 \end{bmatrix} \quad (\text{B-6})$$

Therefore, from Eqs. (17), (B-2), (B-4) and (B-6)

$$S_{FF}(\omega) \approx S_{FF}(\omega) \Big|_{\psi_1=\psi_2=0} + \psi_1 \frac{\partial S_{FF}(\omega)}{\partial \tau_1} \Big|_{\psi_1=\psi_2=0} + \psi_2 \frac{\partial S_{FF}(\omega)}{\partial \tau_2} \Big|_{\psi_1=\psi_2=0} \\ = 4 \sin^2(\omega\Delta) \begin{bmatrix} S_1(\omega) & 0 \\ 0 & S_2(\omega) \end{bmatrix} + 2N(\omega) \begin{bmatrix} 1 & -\cos(\omega\Delta) \\ -\cos(\omega\Delta) & 1 \end{bmatrix} \\ + \psi_1 \omega \sin(\omega\Delta) \begin{bmatrix} 0 & 4S_1(\omega) + 2N(\omega) \\ 4S_1(\omega) + 2N(\omega) & 8S_2(\omega) \cos(\omega\Delta) \end{bmatrix} \\ - \psi_2 \omega \sin(\omega\Delta) \begin{bmatrix} 8S_1(\omega) \cos \omega\Delta & 4S_2(\omega) + 2N(\omega) \\ 4S_2(\omega) + 2N(\omega) & 0 \end{bmatrix} \quad (\text{B-7})$$

Alternatively,

$$\begin{aligned}
\mathbf{S}_{FF}(\omega) &= \begin{bmatrix} 4S_1(\omega) \sin^2(\omega\Delta) + 2N(\omega) & -2N(\omega) \cos(\omega\Delta) \\ -4\omega S_1(\omega) \sin 2(\omega\Delta) \psi_2 & +\omega \sin(\omega\Delta)(\psi_1(4S_1(\omega) + 2N(\omega)) \\ & -\psi_2(4S_2(\omega) + 2N(\omega))) \\ -2N(\omega) \cos(\omega\Delta) & \\ +\omega \sin(\omega\Delta)(\psi_1(4S_1(\omega) + 2N(\omega)) & 4S_2(\omega) \sin^2(\omega\Delta) + 2N(\omega) \\ -\psi_2(4S_2(\omega) + 2N(\omega))) & +\psi_1\omega 4S_2(\omega) \sin 2(\omega\Delta) \end{bmatrix} \\
&= 4 \begin{bmatrix} S_1(\omega)[\sin^2(\omega\Delta) - \omega \sin 2(\omega\Delta) \psi_2] & \omega \sin(\omega\Delta)(\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) \\ \omega \sin(\omega\Delta)(\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) & S_2(\omega)[\sin^2(\omega\Delta) + \omega \sin 2(\omega\Delta) \psi_1] \end{bmatrix} \\
&\quad + 2N(\omega) \begin{bmatrix} 1 & \omega \sin(\omega\Delta)(\psi_1 - \psi_2) - \cos(\omega\Delta) \\ \omega \sin(\omega\Delta)(\psi_1 - \psi_2) - \cos(\omega\Delta) & 1 \end{bmatrix}
\end{aligned} \tag{B-8}$$

Notice that $y_1(t)$ contains only $s_1(t)$ and noise and $y_2(t)$ contains only $s_2(t)$ and noise. so under linearization the signals are separated.

For the BB configuration, we use the fact that $\mathbf{S}_{BB}(\omega) = \frac{1}{4 \sin^2 \omega(\tau_1 - \tau_2)} \mathbf{S}_{FF}(\omega)$. Therefore, about $\psi_1 = 0, \psi_2 = 0$:

$$\begin{aligned}
\mathbf{S}_{BB}(\omega) &= \left[\frac{1}{4 \sin^2 \omega(\tau_1 - \tau_2)} \mathbf{S}_{FF}(\omega) \right]_{\psi_1=\psi_2=0} + \psi_1 \left[\frac{1}{4 \sin^2(\omega\Delta)} \frac{\partial \mathbf{S}_{FF}(\omega)}{\partial \tau_1} \right]_{\psi_1=\psi_2=0} \\
&\quad - \frac{2\omega \cos(\omega\Delta)}{4 \sin^3(\omega\Delta)} \mathbf{S}_{FF}(\omega) \Big|_{\psi_1=\psi_2=0} \\
&\quad + \psi_2 \left[\frac{1}{4 \sin^2(\omega\Delta)} \frac{\partial \mathbf{S}_{FF}(\omega)}{\partial \tau_2} \right]_{\psi_1=\psi_2=0} + \frac{2\omega \cos(\omega\Delta)}{4 \sin^3(\omega\Delta)} \mathbf{S}_{FF}(\omega) \Big|_{\psi_1=\psi_2=0}
\end{aligned}$$

$$\mathbf{S}_{BB}(\omega) = \frac{1}{\sin^2(\omega\Delta)} \mathbf{S}_{FF}(\omega) + \frac{2\omega \cos(\omega\Delta)}{\sin^3(\omega\Delta)} \mathbf{S}_{FF}(\omega) \Big|_{\psi_1=\psi_2=0} (-\psi_1 + \psi_2)$$

$$\begin{aligned}
&= \begin{bmatrix} S_1(\omega) & 0 \\ 0 & S_2(\omega) \end{bmatrix} + \frac{\omega}{\sin(\omega\Delta)} \begin{bmatrix} -2S_1(\omega) \cos(\omega\Delta) \psi_2 & \psi_1 S_1(\omega) - \psi_2 S_2(\omega) \\ \psi_1 S_1(\omega) - \psi_2 S_2(\omega) & 2S_2(\omega) \cos(\omega\Delta) \psi_1 \end{bmatrix} \\
&\quad + 2N(\omega) \begin{bmatrix} \frac{1}{\sin^2(\omega\Delta)} & \frac{\omega}{\sin(\omega\Delta)}(\psi_1 - \psi_2) - \frac{\cos(\omega\Delta)}{\sin^2(\omega\Delta)} \\ \frac{\omega}{\sin(\omega\Delta)}(\psi_1 - \psi_2) - \frac{\cos(\omega\Delta)}{\sin^2(\omega\Delta)} & \frac{1}{\sin^2(\omega\Delta)} \end{bmatrix} \\
&\quad + \frac{2\omega \cos(\omega\Delta)}{\sin(\omega\Delta)} \begin{bmatrix} S_1(\omega) & 0 \\ 0 & S_2(\omega) \end{bmatrix} (-\psi_1 + \psi_2) \\
&\quad + \frac{2N(\omega)2\omega \cos(\omega\Delta)}{\sin^3(\omega\Delta)} \begin{bmatrix} 1 & -\cos(\omega\Delta) \\ -\cos(\omega\Delta) & 1 \end{bmatrix} (-\psi_1 + \psi_2) \\
&= \begin{bmatrix} S_1(\omega) \left\{ 1 - \frac{2\omega}{\text{tg}(\omega\Delta)} \psi_1 \right\} & \frac{\omega}{\sin(\omega\Delta)} \{ \psi_1 S_1(\omega) - \psi_2 S_2(\omega) \} \\ \frac{\omega}{\sin(\omega\Delta)} \{ \psi_1 S_1(\omega) - \psi_2 S_2(\omega) \} & S_2(\omega) \left\{ 1 + \frac{2\omega}{\text{tg}(\omega\Delta)} \psi_2 \right\} \end{bmatrix} \\
&\quad + 2N(\omega) \begin{bmatrix} \frac{1}{\sin^2(\omega\Delta)} \left(1 - \frac{2\omega}{\text{tg}(\omega\Delta)} (\psi_1 - \psi_2) \right) & (\psi_1 - \psi_2) \left(\frac{\omega}{\sin(\omega\Delta)} + 2\omega \frac{\cos^2(\omega\Delta)}{\sin^3(\omega\Delta)} \right) \\ (\psi_1 - \psi_2) \left(\frac{\omega}{\sin(\omega\Delta)} + 2\omega \frac{\cos^2(\omega\Delta)}{\sin^3(\omega\Delta)} \right) & - \frac{\cos(\omega\Delta)}{\sin^2(\omega\Delta)} \\ - \frac{\cos(\omega\Delta)}{\sin^2(\omega\Delta)} & \frac{1}{\sin^2(\omega\Delta)} \left(1 - \frac{2\omega}{\text{tg}(\omega\Delta)} (\psi_1 - \psi_2) \right) \end{bmatrix} \tag{B-9}
\end{aligned}$$

Therefore,

$$\begin{aligned}
S_{BB}(\omega) &= \begin{bmatrix} S_1(\omega) \left(1 - \frac{2\omega}{\text{tg}(\omega\Delta)} \psi_1 \right) & \frac{\omega}{\sin(\omega\Delta)} (\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) \\ \frac{\omega}{\sin(\omega\Delta)} (\psi_1 S_1(\omega) - \psi_2 S_2(\omega)) & S_2(\omega) \left(1 + \frac{2\omega}{\text{tg}(\omega\Delta)} \psi_2 \right) \end{bmatrix} \\
&\quad + \frac{N(\omega)}{2 \sin^2(\omega\Delta)} \begin{bmatrix} 1 + \frac{2\omega}{\text{tg}(\omega\Delta)} (\psi_2 - \psi_1) & (\psi_1 - \psi_2) \frac{\omega}{\sin(\omega\Delta)} (1 + \cos^2(\omega\Delta)) \\ (\psi_1 - \psi_2) \frac{\omega}{\sin(\omega\Delta)} (1 + \cos^2(\omega\Delta)) & - \cos(\omega\Delta) \\ - \cos(\omega\Delta) & 1 + \frac{2\omega}{\text{tg}(\omega\Delta)} (\psi_2 - \psi_1) \end{bmatrix} \tag{B-10}
\end{aligned}$$

Here also, the small error ($\psi_1 \approx \psi_2 \approx 0$) guarantees separation each of the output contains one signal only (plus noise).

Appendix C

From [9]:

$$\int \sin^2 x dx = \frac{1}{2}(x - \sin x \cos x) \quad (\text{C-1})$$

$$\int \frac{1}{\sin^2 x} dx = -\cot x = -\frac{\cos x}{\sin x} \quad (\text{C-2})$$

Therefore,

$$\begin{aligned} \int_{\omega_0 - W/2}^{\omega_0 + W/2} \sin^2(\omega \Delta) d\omega &= \frac{1}{\Delta} \int_{\Delta(\omega_0 - W/2)}^{\Delta(\omega_0 + W/2)} \sin^2 x dx = \frac{1}{2\Delta} [\Delta W - \frac{1}{2} \sin \Delta(2\omega_0 + W) \\ &+ \frac{1}{2} \sin \Delta(2\omega_0 - W)] = \frac{W}{2} [1 - \cos 2\omega_0 \Delta \frac{\sin \Delta W}{\Delta W}] \quad (\text{C-3}) \end{aligned}$$

Also,

$$\begin{aligned} \int_{\omega_0 - W/2}^{\omega_0 + W/2} \frac{1}{\sin^2 \Delta \omega} d\omega &= \frac{1}{\Delta} \int_{\Delta(\omega_0 - W/2)}^{\Delta(\omega_0 + W/2)} \frac{1}{\sin^2 x} dx = -\frac{1}{\Delta} \left[\frac{\cos \Delta(\omega_0 + W/2)}{\sin \Delta(\omega_0 + W/2)} - \frac{\cos \Delta(\omega_0 - W/2)}{\sin \Delta(\omega_0 - W/2)} \right] \\ &= -\frac{1}{\Delta} \frac{\sin \Delta(\omega_0 - \frac{W}{2} - \omega_0 - \frac{W}{2})}{\sin \Delta(\omega_0 + \frac{W}{2}) \sin \Delta(\omega_0 - \frac{W}{2})} = \frac{W \frac{\sin W \Delta}{W \Delta}}{\sin \Delta(\omega_0 + \frac{W}{2}) \sin \Delta(\omega_0 - \frac{W}{2})} \\ &= 2W \frac{\sin W \Delta}{W \Delta} \frac{1}{\cos \Delta W - \cos 2\Delta \omega_0} \quad (\text{C-4}) \end{aligned}$$

Appendix D

In a general backward configuration, $y_i(t) = z_i(t) - h_j(t) * y_j(t)$ (see Fig. B.1).

The impulse response of the filter, $h(t)$, is a function of an unknown parameter (to be controlled) say θ_2 . Thus,

$$y_1(t) = z_1(t) - \int_{-\infty}^{\infty} h_2(\tau) y_2(t - \tau) d\tau = z_1(t) - \int_{-\infty}^{\infty} h(\tau, \theta_2) y_2(t - \tau) d\tau \quad (\text{D-1})$$

The output power is given by $P_1 = E\{y_1^2(t)\}$. A necessary condition for minimization of the output power is $\frac{\partial P_1}{\partial \theta_2} = 0$, That is:

$$\frac{\partial P_1}{\partial \theta_2} = \frac{\partial}{\partial \theta_2} E\{y_1^2(t)\} = E\{2y_1(t) \frac{\partial y_1(t)}{\partial \theta_2}\} = E\{2y_1(t) \frac{\partial}{\partial \theta_2} (z_1(t) - \int_{-\infty}^{\infty} h(\tau, \theta_2) y_2(t - \tau) d\tau)\}$$

$$\begin{aligned}
&= -2E\{y_1(t) \int_{-\infty}^{\infty} \frac{\partial h(\tau, \theta_2)}{\partial \theta_2} y_2(t - \tau) d\tau\} = -2 \int_{-\infty}^{\infty} \frac{\partial h(\tau, \theta_2)}{\partial \theta_2} E\{y_1(t) y_2(t - \tau)\} d\tau \\
&= -2 \int_{-\infty}^{\infty} \frac{\partial h(\tau, \theta_2)}{\partial \theta_2} R_{y_1 y_2}(\tau) d\tau \tag{D-2}
\end{aligned}$$

From Eq. (D-2) we see that $R_{y_1 y_2}(\tau) \approx 0$ is a sufficient, but not a necessary condition for $\frac{\partial P_1}{\partial \theta_2} = 0$: i.e., decorrelation guarantees minimum power, but not the other way around.

VI. REFERENCES

- [1] Y. Bar-Ness, J.W. Carlin, M.L. Steinberger, "Bootstrapping Adaptive Cross-Pol Canceller for Satellite Communication," *Proc. of ICC-82*, June 1982.
- [2] Y. Bar-Ness, "Bootstrapped Adaptive Cross-Pol Interference Canceling Techniques - Steady State Analysis," *Bell Labs. Report*, January 26, 1982. Also accepted for publication in *Signal Processing*.
- [3] Y. Bar-Ness, A. Dinc, H. Messer, "Bootstrapped Adaptive Separation of Superimposed Signals - Analysis of the Effect of Thermal Noise on Performance," Submitted.
- [4] H. Messer, Y. Bar-Ness, "Bootstrapped Spatial Separation of Wideband Superimposed Signals," *Proc. EUSIPCO '92*, August 1992.
- [5] C. Jutten, J. Herault, "Blind Separation of Sources, Part I," *Signal Processing*, July 1991.
- [6] L. Tong, V.C. Soon, Y.F. Huang, R. Liu, "Multiple Source Separation in Noise." 27th Annual Allerton Conference on Communication, Control and Computing, Urbana. IL.. September 1989.
- [7] P. Comon, "Independent Component Analysis," *International Signal Processing Workshop on Higher-Order Statistics*, July 10-12, 1991, Chamrousse, France.
- [8] E. Weinstein, M. Feder, A. V. Oppenheim, "Multi-Channel Signal Separation Based on Decorrelation," Submitted to the *IEEE Trans. on Signal Processing*.
- [9] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products*, 1980.

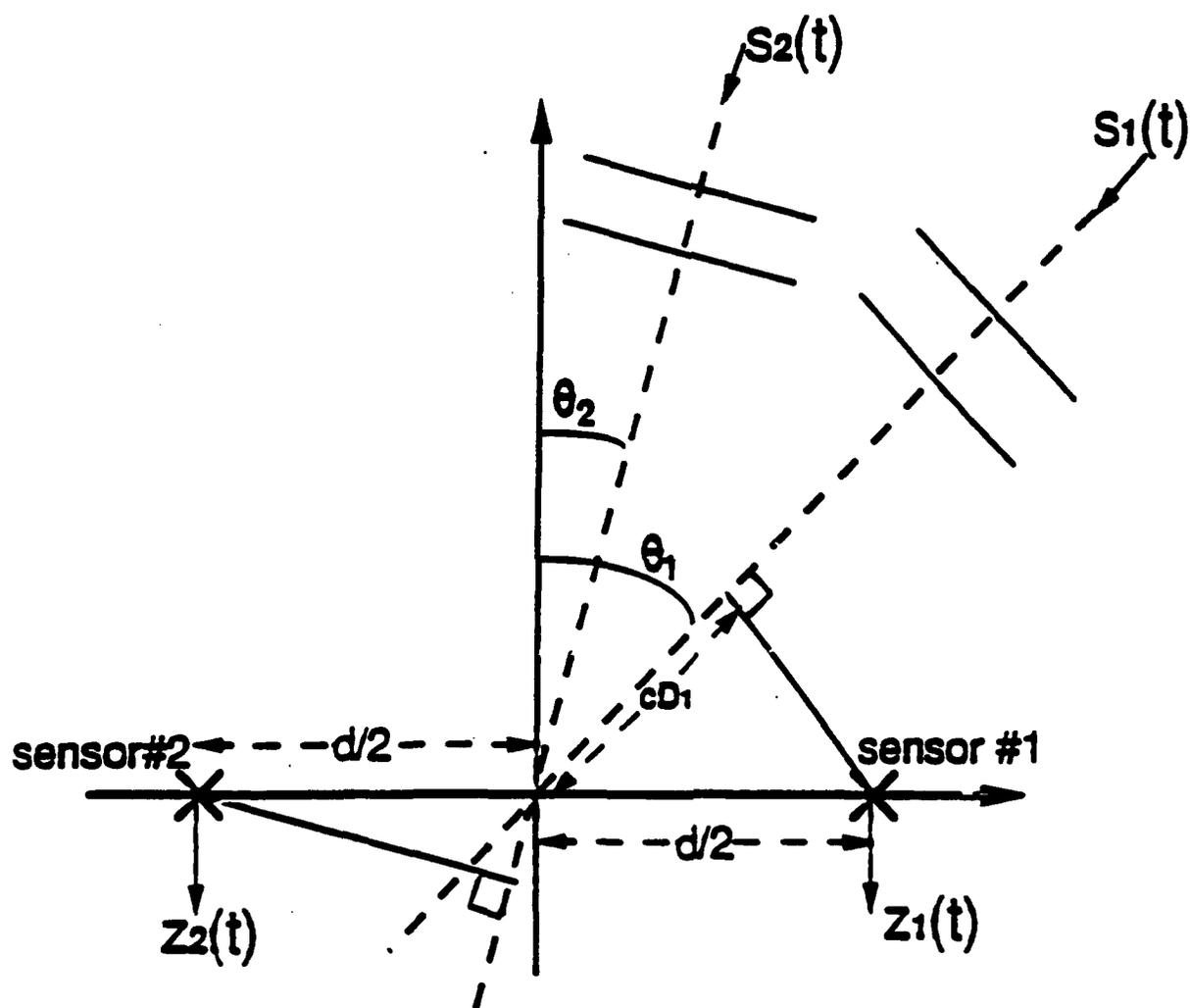
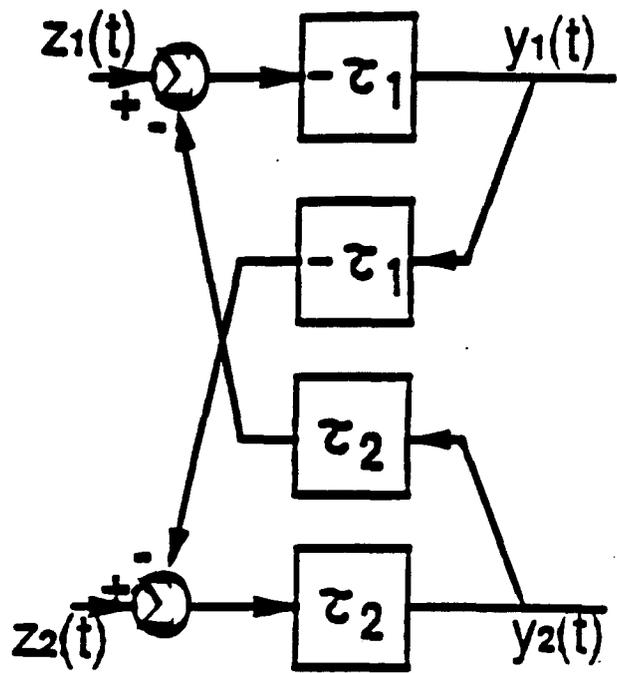
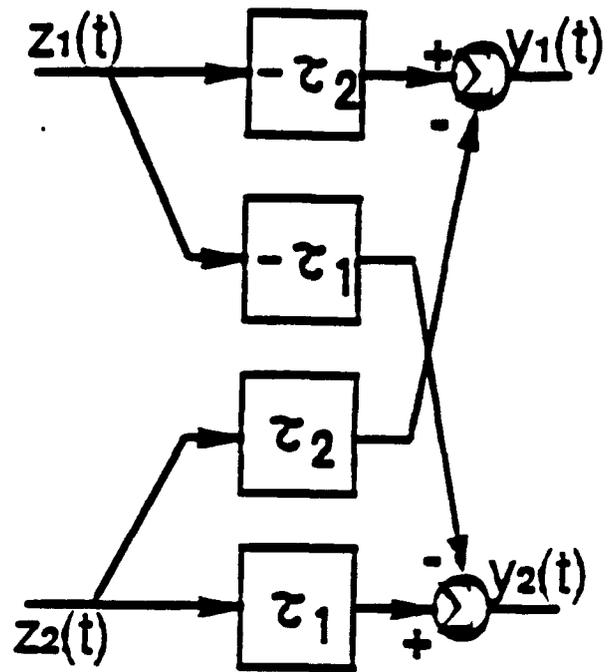


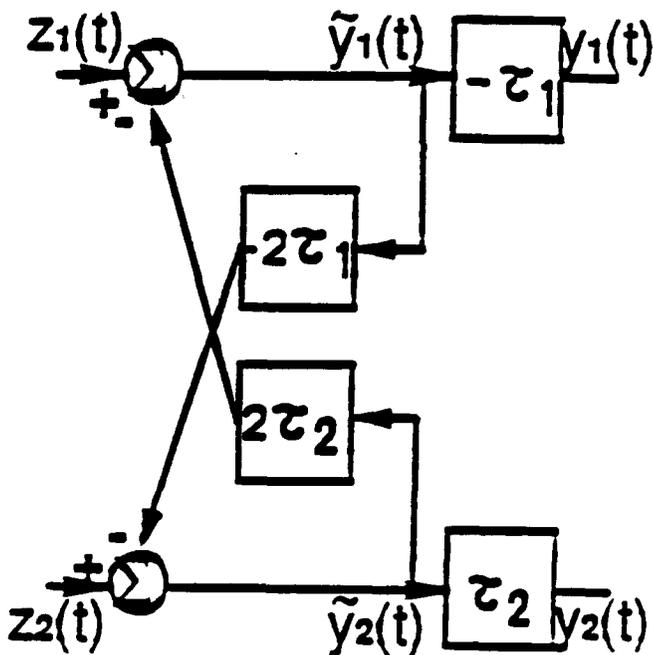
Figure 1 The model of the spatial signal mixture.



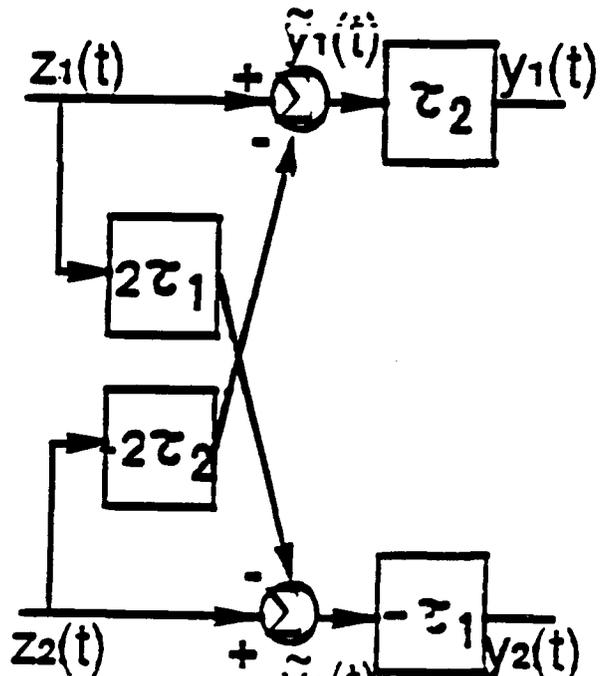
(a)



(c)



(b)



(d)

Figure 2 The different structures of the delay control separators - (a) the basic BB structure. (b) the modified BB structure. (c) the basic FF structure. (d) the modified FF structure.

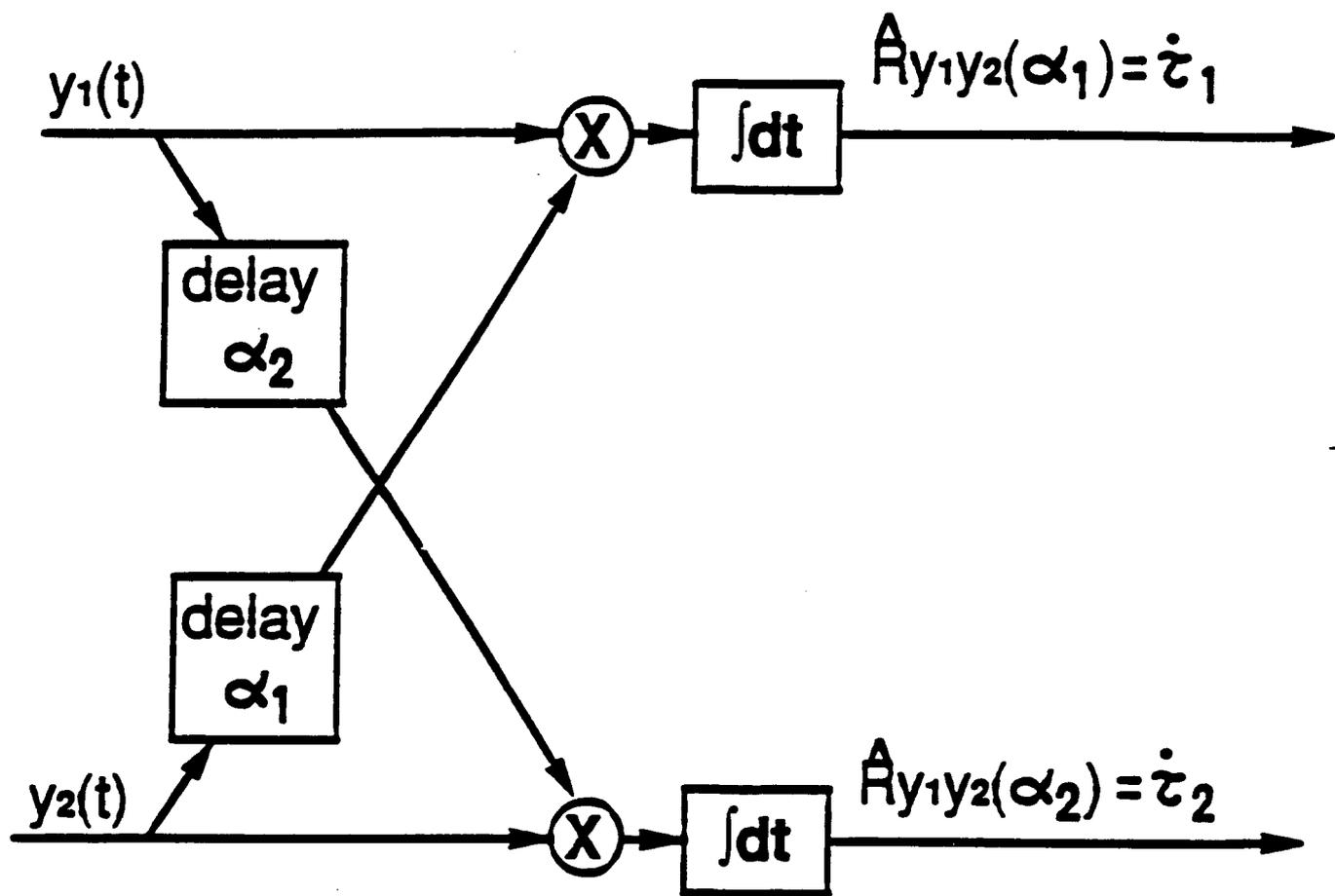


Figure 3 Possible implementation of the adaptive control algorithm.

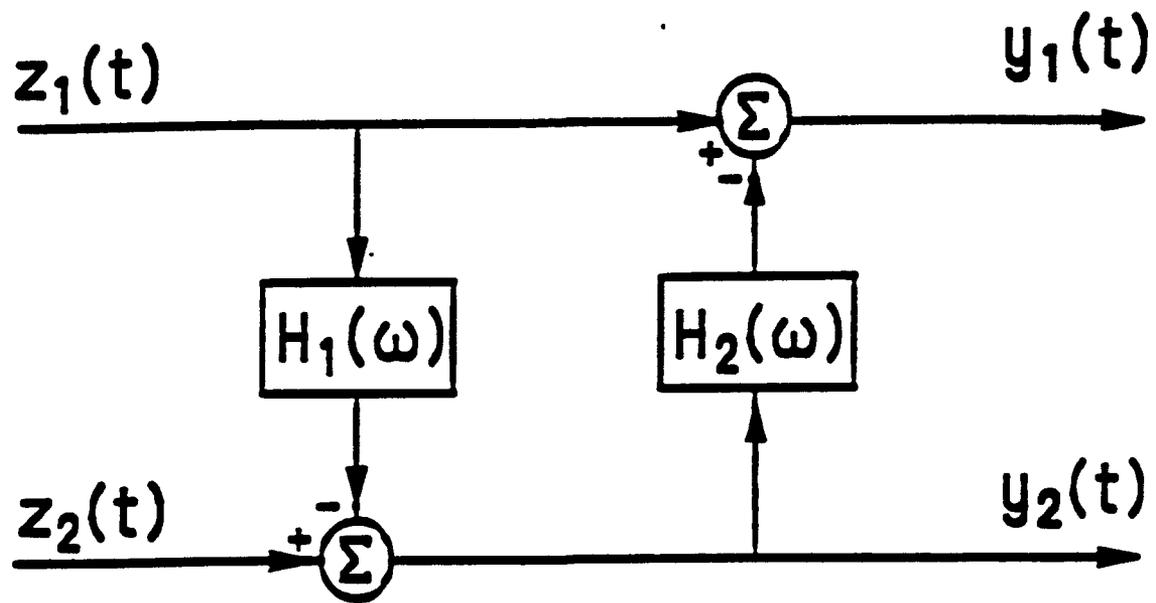


Figure A1 The topology of the FB structure.

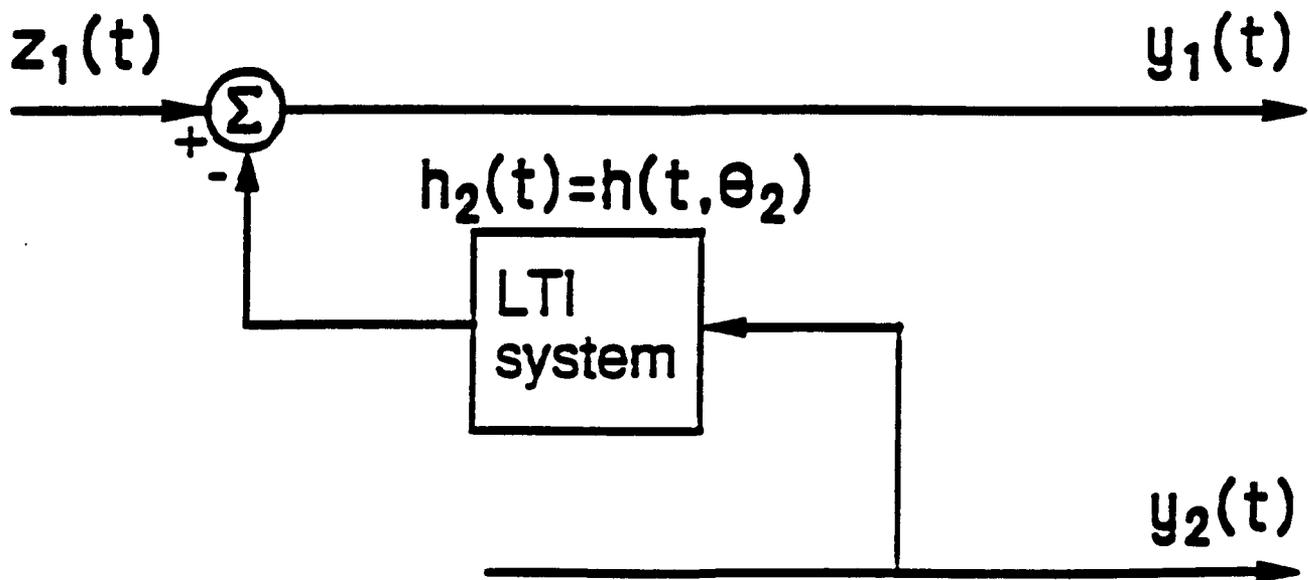


Figure B1 The general feedback structure.

APPENDIX F
BROADBAND INTERFERENCE CANCELATION USING
A BOOTSTRAPPED APPROACH ¹

by

Yehezkel Bar-Ness and Hagit Messer

ABSTRACT

In this paper we present a novel approach for rejecting a broadband interference from unknown direction when received by an array of two sensors. Two configurations of such an approach termed "bootstrapped-based algorithms" are presented. Both configurations perform perfect interference cancelation when the input signal-to-noise-ratio (SNR) is large enough, and do it much faster than the common LMS interference canceler. However, additive noise causes performance degradation to both. It is shown that no general claim can be made about the superiority of one of the configurations with respect to the other. The output signal-to interference-plus-noise-ratio (SINR) depends on the spatial separation between the interference and the desired signal, as well as on the interference-to-noise ratio (SNR), in a different manner for both configurations. The paper provides guidelines for the choice of one or the other configuration in different scenarios.

¹This work was presented at ICASSP '93.

This work was partially supported by grant from ROME (AFSC, Griffiss Air Force Base, NY) under contract F 30602-88-D-0025, Task C-2-2404.

I. INTRODUCTION AND BASIC THEORY

A common problem in many fields (e.g., communication, radar, sonar, EM etc.) is the need to cancel a spatial interference in the presence of a desired source. The conventional approach to solve this problem is based on the LMS algorithm [1]. For narrowband signals, the LMS algorithm is used to adapt complex weights so as to minimize mean square error between the array output and a reference signal. For canceling an interference from an unknown direction when no reference is available a, minimum output power criterion might be used. It was shown by Compton [2] that this criterion leads to *power inversion*. That is, if the input signal to interference ratio (SIR) is ρ , then the output SIR is $1/\rho$. Bar-Ness et. al developed a class of "bootstrapped algorithms" [3-5] that perform perfect interference cancellation independent of the input SIR. In this paper we use the bootstrapped approach for *broadband* interference cancellation, where the complex weights of the narrowband case are replaced by pure delay lines. Unlike the narrowband one, in this case the open-loop system is a non-linear function of the weights (the delays), resulting in inherent differences between the narrow band and the broadband. However, following the Bar-Ness *approach* we developed two configurations of a broadband interference canceler; both perform perfect interference cancellation when no additive, thermal noise is present.

The two bootstrapped configurations for broadband interference cancellation denoted by: backward-backward (BB) and forward-forward (FF), are depicted in Fig. 1. $z_1(t)$ and $z_2(t)$ are the outputs of the two sensors and are assumed to be:

$$\begin{aligned} z_1(t) &= s(t - D_s) + i(t - D_i) + e_1(t) \\ z_2(t) &= s(t + D_s) + i(t + D_i) + e_2(t) \end{aligned} \quad (1)$$

where $s(t)$ and $i(t)$ are the desired signal and the interference signal, respectively, radiated from bearings θ_s and θ_i , $e_1(t)$ and $e_2(t)$ are the additive noise processes in each of the sensors

and

$$D_s = \frac{d}{2c} \sin \theta_s \quad \text{and} \quad D_i = \frac{d}{2c} \sin \theta_i \quad (2)$$

where d is the separation between the sensors and c is the propagation velocity. We assume that the signal direction, θ_s , which is the look direction, is *known*, while the interference direction, θ_i is unknown. The same system can be used for *separation* of two sources [6]. It can be shown that the signal denoted by $y_2(t)$ in Fig. 1 is an estimate of the interference waveform. In our special case, we are only interested in the estimating the desired signal waveform. This is given by the output $y_1(t)$. We assume that the random signals $s(t)$, $i(t)$, $e_1(t)$ and $e_2(t)$ are mutually uncorrelated, wide sense stationary Gaussian processes, with power density spectrum (PDS) $S(\omega)$, $I(\omega)$, and $E_1(\omega) = E_2(\omega) = N(\omega)$, respectively. Based on this assumption, the PDS matrix of the signals $y_1(t)$ (the output) and $y_2(t)$ in the BB configuration of Fig. 1a is given by:

$$\begin{aligned} \mathbf{S}_{BB}(\omega) = & \\ & \begin{bmatrix} S(\omega) + I(\omega) \left(\frac{\sin \omega(D_i - \tau)}{\sin \omega(D_s - \tau)} \right)^2 & I(\omega) \frac{\sin \omega(D_i - \tau)}{\sin \omega(D_s - \tau)} \frac{\sin \omega(D_s - D_i)}{\sin \omega(D_s - \tau)} \\ I(\omega) \frac{\sin \omega(D_i - \tau)}{\sin \omega(D_s - \tau)} \frac{\sin \omega(D_s - D_i)}{\sin \omega(D_s - \tau)} & I(\omega) \left(\frac{\sin \omega(D_s - D_i)}{\sin \omega(D_s - \tau)} \right)^2 \end{bmatrix} \\ & + \frac{N(\omega)}{2 \sin^2 \omega(D_s - \tau)} \begin{bmatrix} 1 & -\cos \omega(D_s - \tau) \\ -\cos \omega(D_s - \tau) & 1 \end{bmatrix} \end{aligned} \quad (3)$$

From Eq. (3), the output PDS, from which the output signal to interference plus noise ratio (SINR) can be calculated for any τ :

$$S_{y_1}(\omega) = S(\omega) + I(\omega) \left(\frac{\sin \omega(D_i - \tau)}{\sin \omega(D_s - \tau)} \right)^2 + \frac{N(\omega)}{2 \sin^2 \omega(D_s - \tau)} \quad (4)$$

Similarly, it can be shown that the power density spectrum matrix for the FF configuration of Fig. 1.b is given by :

$$\mathbf{S}_{FF}(\omega) = 4 \sin^2 \omega(D_s - \tau) \mathbf{S}_{BB}(\omega) \quad (5)$$

From this equation we see that the power spectrum (and cross-spectrum of the outputs) of the BB and FF configurations differ only by $4 \sin^2 \omega(D_s - \tau)$. The transfer function of the BB configuration includes a pole at $\omega = 0$ and therefore cannot process baseband signals, while the transfer function of the FF configuration is a zeros-only one. As such, the FF configuration can handle low-pass signals too, and is also not sensitive to situations where $D_s = \tau$. Notice that the BB configuration can be regarded as an FF one, followed by a 2×2 system with a transfer function matrix $\frac{1}{4 \sin^2 \omega(D_s - \tau)} \mathbf{I}$, where \mathbf{I} is the 2-dimensional identity matrix.

Eq. (3) suggests that *decorrelation* of the outputs $y_1(t)$ and $y_2(t)$ is a reasonable optimization criterion for interference cancellation. If no noise is present ($N(\omega) = 0$) then by imposing decorrelation of $y_1(t)$ and $y_2(t)$, i.e., by controlling τ so that the off-diagonal entries of the first matrix in Eq. (3) are zero, a perfect interference cancellation is performed. Based on this idea, the control algorithm in Fig. 1 is chosen to be decorrelation of its two outputs. That is, the voltage controlled delay line which implements τ is controlled by the output of an integrator, whose input is the correlation of $y_1(t)$ and $y_2(t)$. Therefore, the steady-state value of τ is that for which the off diagonal entries of Eq. (3) are zero for all ω . In the next section, we present a comparative study of the two canceler structures.

II. THE CONTROL ALGORITHM

The algorithm described in the last section is extremely non-linear in τ . To make the theoretical analysis traceable, we assume that the error, $\phi = D_i - \tau$ is small and we take the linearization of equations (3) and (5) to be about $\phi = 0$. The control algorithm in the two configurations is determined by the cross correlation between the output $y_1(t)$ and $y_2(t)$, which is the inverse Fourier transform of their cross-spectrum, $S_{12}(\omega)$. In general, it can be

described by the differential equation:

$$\frac{d\tau}{dt} = kR_{12}(0; \tau) = kE\{y_1(t)y_2(t)\} \quad (6)$$

where k is a constant and $R_{12}(\alpha; \tau) = E\{y_1(t)y_2(t-\alpha)\}$ is the cross-correlation between the outputs. Under small errors assumption, this cross spectrum is given by:

$$S_{12}^{FF}(\omega) \approx \phi\omega \sin(\omega\Delta)[4I(\omega) + 2N(\omega)] - 2N(\omega) \cos(\omega\Delta) \quad (7)$$

where $\Delta = D_s - D_i$. Also

$$S_{12}^{BB}(\omega) \approx \frac{S_{12}^{FF}(\omega)}{4 \sin^2(\omega\Delta)} + \phi N(\omega) \frac{\omega \cos^2(\omega\Delta)}{\sin^3(\omega\Delta)} \quad (8)$$

In this case, the non-linear differential equation (6) is linear in τ and, for $k = 1$, is given by:

$$2\pi \frac{d\phi_{FF}}{dt} = - \phi_{FF} \int_{-\infty}^{\infty} \omega \sin(\omega\Delta)[4I(\omega) + 2N(\omega)]d\omega + 2 \int_{-\infty}^{\infty} N(\omega) \cos(\omega\Delta)d\omega \quad (9)$$

$$2\pi \frac{d\phi_{BB}}{dt} = - \phi_{BB} \int_{-\infty}^{\infty} \left(\frac{\omega \sin(\omega\Delta)[4I(\omega) + 2N(\omega)]}{4 \sin^2(\omega\Delta)} + N(\omega) \frac{\omega \cos^2(\omega\Delta)}{\sin^3(\omega\Delta)} \right) d\omega + 2 \int_{-\infty}^{\infty} \frac{N(\omega) \cos(\omega\Delta)}{4 \sin^2(\omega\Delta)} d\omega \quad (10)$$

For the special case where the spectrum of all signals is flat, that is, $S(\omega) = S$, $I(\omega) = I$ and $N(\omega) = N$ in the band, and zero elsewhere, the approximated control equations become:

$$\frac{\pi}{2N} \frac{d\phi_{FF}}{dt} = -\phi_{FF}(1 + 2\rho_I) \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \omega \sin(\omega\Delta)d\omega + \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \cos(\omega\Delta)d\omega \quad (11)$$

$$\frac{2\pi}{N} \frac{d\phi_{BB}}{dt} = - \phi_{BB} \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \frac{\omega}{\sin(\omega\Delta)} \left(2 \frac{\cos^2(\omega\Delta)}{\sin^2(\omega\Delta)} + 2\rho_I + 1 \right) d\omega + \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \frac{\cos(\omega\Delta)}{\sin^2(\omega\Delta)} d\omega \quad (12)$$

where $\omega_0 = 2\pi f_0$ is the center frequency and $W = 2\pi B$ is the bandwidth of the processing band. In Fig. 2 we present $\tau = D_i - \phi$ as a function of time as derived numerically from Eqs. (11) and (12). In both cases the separation between the sensors is $d = 30m$. $D_i = 25nSec$ and $\Delta = 18nSec$. In Fig. 2a $\rho_I = 60dB$ (high INR), $f_0 = 3.005MHz$ and $B = 5.99MHz$. In Fig. 2b $\rho_I = -30dB$ (low INR), $f_0 = 12.75MHz$ and $B = 24.5MHz$. In both cases, the BB configuration (the dashed line) converges faster than the FF configuration (the solid line). However, due to the presence of the additive noise, the algorithms don't converge to the true delay ($D_i = 25nSec$ in our example) so there is a bias in the estimate of the unknown delay D_i . This bias is larger for the BB configuration in both cases depicted in Fig. 2.

III. OUTPUT SINR, DISCUSSION AND CONCLUSIONS

From Eq. (4) we see that indeed, if the algorithm converges to $\phi = 0$, the interference is completely rejected. For any other $\tau \neq D_i$, the output signal-to-interference-plus-noise-ratio (SINR), for the case of flat-spectrum signals, is given by:

$$SINR_o^{BB} = 2\rho_s W \frac{1}{\int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \frac{1 + 2\rho_I \sin^2 \omega(D_i - \tau)}{2 \sin^2 \omega(D_s - \tau)} d\omega} \quad (13)$$

$$SINR_o^{FF} = \rho_s \frac{\int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \sin^2 \omega(D_s - \tau) d\omega}{W + \rho_I \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \sin^2 \omega(D_i - \tau) d\omega} \quad (14)$$

where $\rho_s = \frac{S}{N}$ is the input signal-to-noise-ratio. However, under a small-errors assumption, this output SINR become,

$$SINR_o^{BB} \approx \frac{2\rho_s W}{\int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} \frac{1 - \phi \frac{2\omega}{\tan(\omega\Delta)}}{2 \sin^2(\omega\Delta)} d\omega} \quad (15)$$

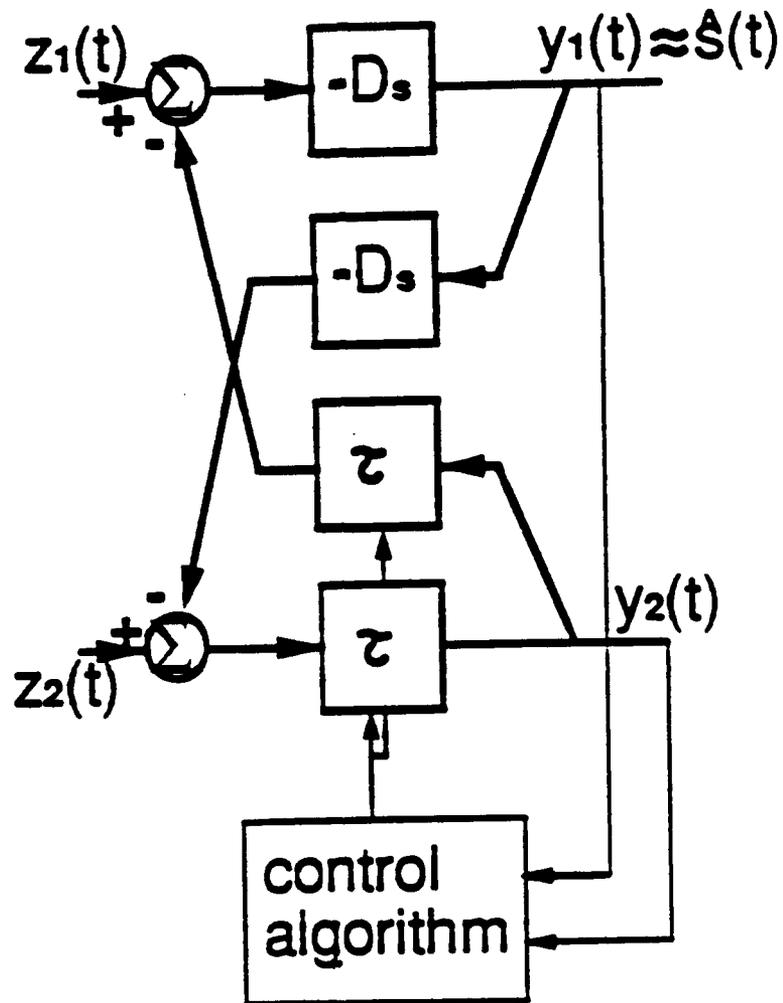
$$SINR_o^{FF} \approx \frac{\rho_s}{W} \int_{\omega_0 - \frac{W}{2}}^{\omega_0 + \frac{W}{2}} (\sin^2(\omega\Delta) - \phi \omega \sin(2\omega\Delta)) d\omega \quad (16)$$

From these last two equations we see that, in both configurations, the output SINR is not a function of ρ_I . That is, if τ is closed to D_i , it is sufficient to guarantee that no compo-

ment of the interference appears in the output $y_1(t)$. Notice, however, that the interference rejection causes reduction in the SNR. That is, the bias in the estimate of D , resulted from the presence of additive noise causes only degradation in output SNR (relative to the input SNR ρ_s) while leaving the output SIR (signal-to-interference-ratio) infinite, provided that it is kept small. As shown in the numerical examples, although both configurations of the proposed algorithms guarantee interference rejections, there are differences in their performances: rate of convergence, bias, region of operation (bandwidth etc.). These differences, as well as quantitative comparison of the proposed approach with the equivalent LMS canceler, are now under ongoing research.

IV. REFERENCES

- [1] B. Widrow, J. McColl, M. Ball: "The Complex LMS Algorithm", *IEEE Proc.* Vol. 63. No. 12, pp. 719-720, Dec. 1975.
- [2] R.T. Compton Jr.: *Adaptive Antennas: Concepts and Performance*. Prentice Hall, 1988.
- [3] Y. Bar-Ness, J. Rokach: "Cross-Coupled Bootstrapped Interference Canceler". *Proc. of the 1981 SP-S International Symposium*, pp. 292-295.
- [4] Y. Bar-Ness, J.W. Carlin, M.L. Steinberger: "Bootstrapping Adaptive Interference Cancelers: Some Practical Limitations", *Proc. of the Globecom*, Paper F3.7, 1982.
- [5] Y. Bar-Ness, A. Dinc: "Performance Comparison of LMS, Diagonalizer and Bootstrapped Adaptive Cross-Pol Canceler Over Non-Dispersive Channel", *MILCOM '1990*, paper 3.7
- [6] H. Messer, Y. Bar-Ness: "Bootstrapped Spatial Separation of Wideband Superimposed Signals", *Proc. of EUSIPCO '92*.



(a)

Figure 1a The backward-backward structure.

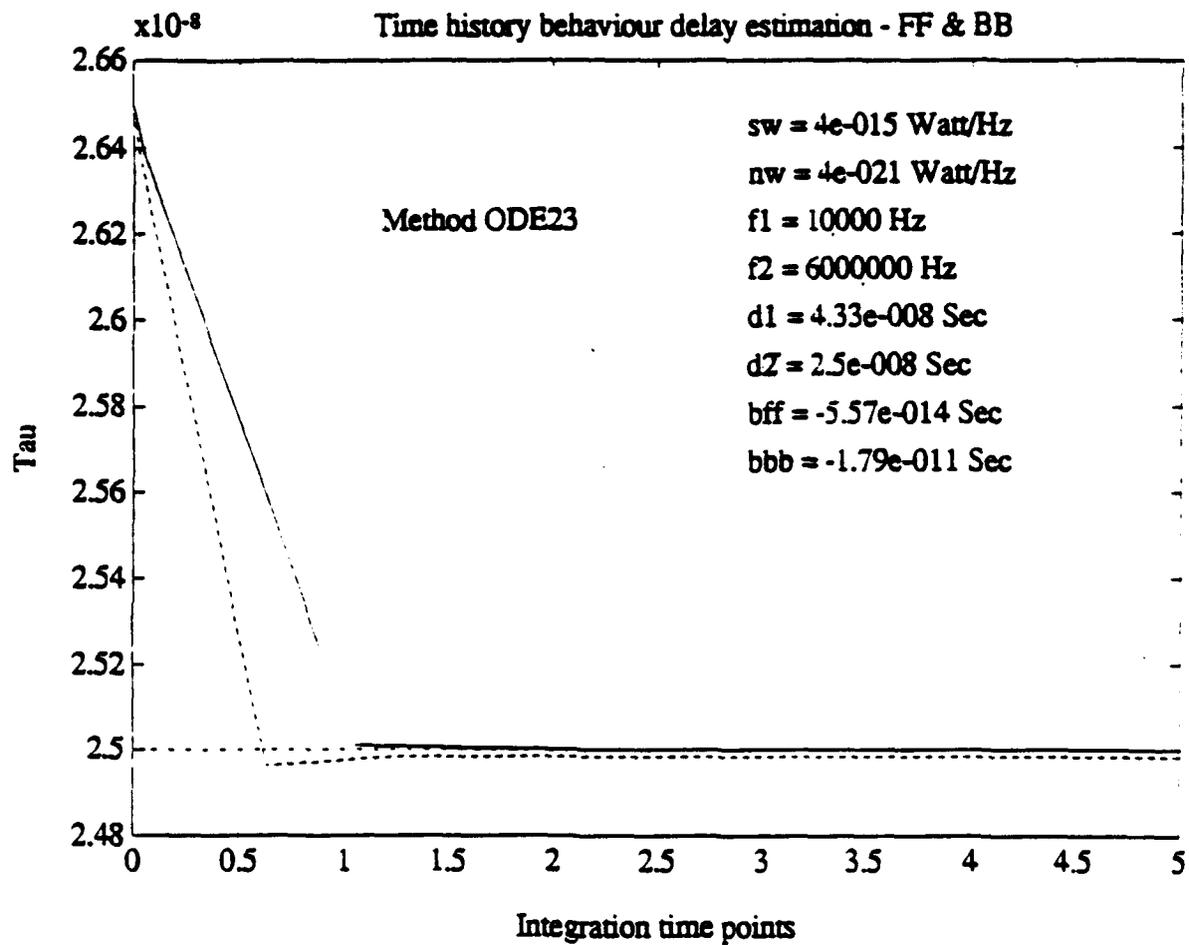


Figure 2a Delay estimation of D_i as a function of time. For FF configuration (solid lines) and for BB configuration (dotted lines), $D_s = 4.33 \cdot 10^{-8}$ sec (known) $D_i = 2.5 \cdot 10^{-8}$ sec (unknown). $I = 4 \cdot 10^{-15}$ Watts/Hz, $N = 4 \cdot 10^{-21}$ Watts/Hz $B = 5.99$ MHz ($f_1 = 10$ KHz, $f_2 = 6$ MHz) resulted bias for FF = $-5.57 \cdot 10^{-14}$ sec for BB = $-1.79 \cdot 10^{-11}$ sec.

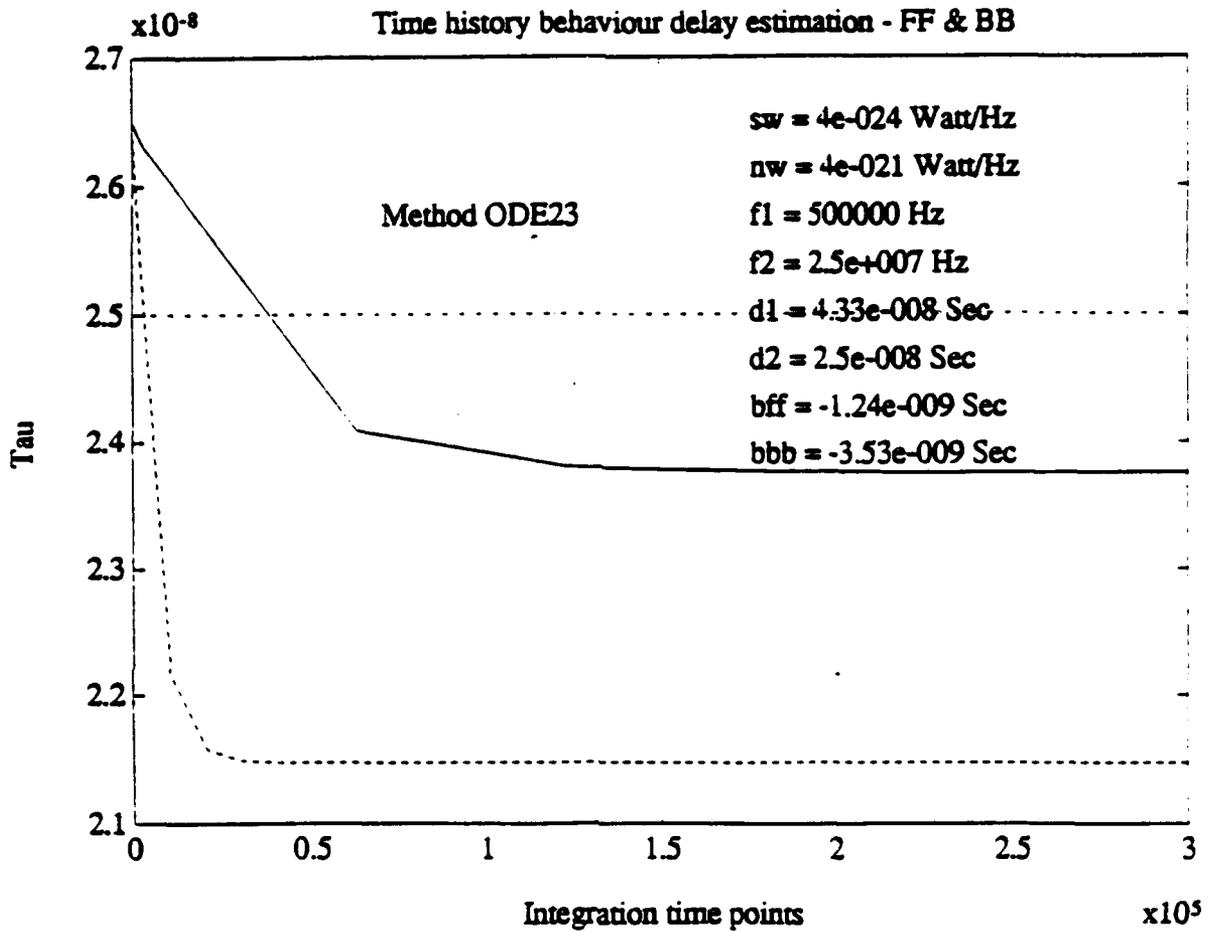


Figure 2b Delay estimation of D_i as a function of time. For FF configuration (solid lines) and for BB configuration (dotted lines), $D_s = 4.33 \cdot 10^{-8}$ sec (known) $D_i = 2.5 \cdot 10^{-8}$ sec (unknown). $I = 4 \cdot 10^{-24}$ Watts/Hz, $N = 4 \cdot 10^{-21}$ Watts/Hz $B = 24.5$ MHz ($f_1 = 500$ KHz. $f_2 = 25$ MHz) resulted bias for FF = $-1.24 \cdot 10^{-9}$ sec for BB = $-3.53 \cdot 10^{-9}$.

APPENDIX G
A FORWARD/BACKWARD BOOTSTRAPPED STRUCTURE
FOR BLIND SEPARATION OF
SIGNALS IN A MULTI-CHANNEL
DISPERSIVE ENVIRONMENT ¹

by

Abdulkadir Dinç and Yeheskel Bar-Ness

Abstract

This paper proposes a new multidimensional adaptive algorithm with *Forward/Backward* bootstrapped structure for dispersive channel environment. It is an alternative multi-signal separator where the loop-bandwidth of the signal separator structure and steady state performance are crucial. It separates superimposed convolutive multi uncorrelated signals. The Bootstrapped adaptive algorithm which does not require a training sequence employs the minimization of output signal correlations as optimization criteria. The control algorithm is set for the multidimensional case. The learning process of the 2 dimensional signal separator using computer simulation is investigated and compared to that of the least mean square (LMS) algorithm for different cross channel eigenvalue spreads.

¹This work was presented at ICASSP '93.

This work is supported by grant from ROME Air Development Center (AFSC, Griffiss Air Force Base, NY under contract F30602-88-D-0025, Task C-0-2456).

I. INTRODUCTION

Special structure multi-loop separators of multi-channel superimposed signals sometimes termed "bootstrapped separators" have been discussed in the literature. Applied to the two signal case and non-dispersive channel, three structures of this canceller were introduced [1]. These were referred to as Forward/Forward (FF), Backward/Backward (BB) and Forward/Backward (FB) (or, according to the performance criterion used, they are termed: correlator/correlator, power/power, and power/correlator, respectively). One or the other of these structures proved to be useful in practical application of cross-polarization cancellation. e.g satellite communication and microwave radio. Applied to digital communication with high M-ary QAM dually polarized signals, it was recently shown [2] that the bootstrapped separator outperforms other signal separators including the LMS (least mean square). The average symbol error probability on both signals was shown to be much lower, which clearly indicates that the bootstrapped separators result in deeper cancellation of the undesired signal in spite of the fact that *no supervisory signal* is needed. Furthermore, due to the feedback properties of the bootstrapped structure, it was shown [3] that the learning process is shorter than that of the LMS algorithm, particularly at low signal-to-interference ratios.

Bandwidth-complexity trade-offs of the three structure are discussed in [4]. It was found that, particularly when it is applied as a *blind separator* at very high RF (radio frequency) such as in dually polarized satellite or microwave terrestrial communications, the FB structure has the best system bandwidth complexity trade-off when compared to the other two structures.

Recently, the FF and BB structures of bootstrapped blind signal separators have been considered as signal separators in a dispersive environment [5,6]. But, while the first can tolerate a wide system bandwidth it requires twice as many channel equalizers as the FB structure. On the other hand, the BB structure can use simple power measuring devices to control the weight and requires no complex channel equalizer. However, it is very much

limited in handling wide bandwidth signals.

The main purpose of this paper is to present stochastic convergence and steady state performance of the Forward/Backward structure and compare it to that of the LMS algorithm.

Without loss of generality, and for simplicity, the analysis is done with a two input-output separator. After introducing the structure of the forward/backward scheme of the bootstrapped separator, the optimal control weights will be presented and shown to lead to a total separation of the two input signals in the noise-free case. Stochastic approximations are employed to show convergence in the mean of these weights to their predicted optimal values. The performance of the FB bootstrapped separator for a dispersive two inputs-two-outputs interference channel with a different signal-to-interference ratio is compared to that of the LMS algorithm.

II. CHANNEL MODEL AND PROBLEM STATEMENT

A discrete time model of an N-dimensional dispersive interference channel is given by

$$\mathbf{x}(n) = \mathbf{H}(n) * \mathbf{I}(n) + \mathbf{n}(n) \quad (1)$$

where * denotes convolution and \mathbf{H} is an (NxN) channel matrix. \mathbf{I} is an (Nx1) information vector assumed to be independent and identically distributed sequence and \mathbf{n} is an (Nx1) white Gaussian noise and \mathbf{x} is the received signal vectors, respectively.

$$\mathbf{H} = \begin{bmatrix} 1 & \cdots & h_{1N}(n) \\ \cdot & 1 & \cdot \\ h_{N1}(n) & \cdots & 1 \end{bmatrix}. \quad (2)$$

The channel responses are assumed to be slowly time varying finite impulse response (FIR). That is, the channel interference filters are assumed to vary slowly with respect to the signal $\mathbf{I}_i(n)$ rate and to be less than unity in magnitude. $h_{ij}(n)$ $i \neq j$ $i, j = 1, 2, \dots, N$ are transversal filters. $h_{ij}^0, h_{ij}^1, \dots, h_{ij}^{K-1}$ are cross channel filter tap coefficients. The diagonal coefficients h_{ii} $i = 1, 2, \dots, N$ are assumed to be unity.

Our objective is to find a multi dimensional bootstrapped adaptive algorithm structure that will diagonalize the channel matrix \mathbf{H} without using a training sequence and compare its convergence with that of an LMS algorithm .

III. LMS AND BOOTSTRAPPED ADAPTIVE SIGNAL SEPARATORS

3.1 Multidimensional LMS Adaptive Signal Separator

The traditional LMS algorithm which minimizes the error $E\{e_1^2 + e_2^2 + \dots e_N^2\}$ at the output of the separator Fig. 1² can be used as a multidimensional signal separator.

The error at each output is given by

$$e_i(n) = x_i(n) + \sum_{j=1}^N \mathbf{w}_{ij} * x_j(n) - I_i(n) \quad (3)$$

where

$$\mathbf{w}_{ij} * x_j(n) = \sum_{m=0}^{M_{ij}-1} w_{ij}^m x_j(n-m) \quad (4)$$

and $w_{ij}^0, w_{ij}^1, \dots w_{ij}^{M_{ij}-1}$ are the tap weight of the \mathbf{w}_{ij} transversal filter. From (3) we write

$$\frac{\partial e_i^2}{\partial w_{ij}^m} = 2E_i\{e_i(n)x_j(n-m)\} \quad m = 0, 1, \dots (M_{ij} - 1) \quad (5)$$

The recursive stochastic weight updating algorithm to search for the optimum weights is given by,

$$\begin{aligned} \mathbf{w}_{ij}^m(n+1) &= \mathbf{w}_{ij}^m(n) - \mu e_i(n) x_j(n-m), \quad i, j = 1, 2, \dots, N \\ e_i(n) &= y_i(n) - \hat{I}_i(n) \quad i \neq j \end{aligned} \quad (6)$$

where $\mathbf{I}_i(n)$ is the reference signal vector, obtained by the help of the training sequence or some other decision feedback means.

²In order to have the same number of weight as in the bootstrap structure w_{ii} were taken to equal unity

3.2 Multidimensional Forward/Backward Bootstrapped Signal Separator

From Fig. 2, it can be easily shown that the output is given by;

$$y_i(n) = x_i(n) - \sum_{k=1}^{i-1} w_{ik} * x_k(n) - \sum_{j=i+1}^N w_{ij} * y_j(n) \quad (7)$$

where $w_{ij} * y_j(n)$ as in (4). The z-transform of the tap weight vector is $w_{ij}(z) = \sum_{m=0}^{M_{ij}-1} w_{ij}^{(m)} z^{-m}$.

Denoting each tap weight vectors as the entries of a matrix, we define two (N x N) weight matrices

$$\mathbf{W}_{upper} = \begin{bmatrix} 1 & w_{12} & \cdots & w_{1N} \\ 0 & 1 & & w_{2N} \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdot & 1 \end{bmatrix} \quad (8)$$

$$\mathbf{W}_{lower} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -w_{21} & 1 & & 0 \\ \cdot & \cdot & \cdots & \cdot \\ -w_{N1} & -w_{N2} & \cdot & 1 \end{bmatrix} \quad (9)$$

Using (7) together with (8) and (9), we can write

$$\mathbf{W}_{upper} * \mathbf{y}(n) = \mathbf{W}_{lower} * \mathbf{x}(n) \quad (10)$$

Let the z-transform of the signal separator output vector $\mathbf{y}(n)$ be given by

$$Y(z) = \mathbf{W}_{upper}(z)^{-1} \mathbf{W}_{lower}(z) X(z) \quad (11)$$

It can be shown that in aa no noise condition, the suggested bootstrapped recursive algorithm will result in a matrix $\mathbf{N}(z) = \mathbf{W}_{upper}^{-1}(z) \mathbf{W}_{lower}(z) \mathbf{H}(z)$ whose off-diagonal entries are zero.

3.2.1 Optimal Weights

For simplicity, in the rest of the analysis we show the optimal weight vectors for two dimensional signal separation in the absence of noise. The optimum weights that minimizes

the output correlations are obtained by requiring

$$R_{y_i y_j}(k) = E\{y_i(n+k)y_j(n)\} = 0, ; i \neq j, i, j = 1, 2 \dots N \quad (12)$$

For $I_1(n)$ and $I_2(n)$ identically distributed zero mean independent sequences we have

$$R_{I_i}(k) = \sigma_i^2 \delta k, \quad i = 1, 2 \quad (13)$$

Hence by using (11) we get for the cross-power spectrum

$$\begin{aligned} S_{y_1 y_2}(z) = & \sigma_1^2 [1 + W_{12}(z)[W_{21}(z) - H_{21}(z)]] [H_{21}(z^{-1}) - W_{21}(z^{-1})] \\ & + \sigma_2^2 [H_{12}(z)[1 + W_{12}(z)W_{21}(z)] - W_{12}(z)[1 - H_{12}(z^{-1})W_{21}(z^{-1})] \end{aligned} \quad (14)$$

There are two optimum adaptive weight vector solutions that make the cross power spectrum in (16) to be zero. They are respectively,

$$\begin{aligned} \mathbf{W}_{opt1}(z) &= [W_{1opt12}(z), W_{2opt21}(z)] \\ &= \left[\frac{H_{12}(z)}{1 - H_{21}(z)H_{12}(z)}, H_{21}(z) \right] \end{aligned} \quad (15)$$

$$\mathbf{W}_{opt2}(z) = \left[\frac{-H_{12}(z)}{1 - H_{21}(z)H_{12}(z)}, \frac{1}{H_{12}(z)} \right] \quad (16)$$

Using the desired solution (14), the outputs of the two-dimensional signal separator are given by

$$\mathbf{Y}(z) = \mathbf{N}(z)\mathbf{I}(z) \quad (17)$$

where

$$\mathbf{N}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 - H_{12}(z)H_{21}(z) \end{bmatrix} \quad (18)$$

3.2.2 The Search for the Optimal Weights

The control algorithm simultaneously minimizes the output correlations $f(y_i(n))y_j(n - m)$ $i, j = 1..N$ $i \neq j$, $m = 0, \dots, (M_{ij} - 1)$, where M_{ij} is the number of filter taps in \mathbf{w}_{ij} . The optimum weights can be found by successive use of the following recursive equations

$$\begin{aligned} \mathbf{w}_{ij}^m(n+1) &= \mathbf{w}_{ij}^m(n) - \mu f(y_i(n))y_j(n-m) \\ i, j &= 1..N \quad i \neq j, \quad m = 0, 1, \dots, M_{ij} - 1 \end{aligned} \quad (19)$$

where $f(y) = y^3$ is an odd nonlinear function and μ is the stability convergence constant.

IV. SIMULATION

The inputs $I_i(n)$ $i = 1, 2..N$ to the channel are taken to be binary ± 1 with equal probability. The simulation is realized for a two dimensional channel, where the cross channel filters $\mathbf{h}_{12}(n)$ and $\mathbf{h}_{21}(n)$ are chosen to be 3 taps raised cosine, with $h(n) = \frac{1}{2}[1 + \cos(\frac{2\pi}{W}(n-2))]$ $n=1, 2, 3$, and $W=2.9$ [9]. To the outputs of the channel, a Gaussian white noise with signal-to-noise ratio (SNR), of 40 dB is added. The two outputs of the channel are used in one hand as inputs to a 4-tap LMS, and to the *forward/backward bootstrapped* adaptive signal separators in the other. The results of 500 Monte Carlo runs are given for different cross channel coupling ratios, between the i th and j th channel inputs defined as

$$SIR = \frac{E\{I_i^2(n)\}}{k^2 E\{I_j^2(n)\} \sum_{m=0}^{M_{ij}-1} (h_{ij}^m)^2}$$

where k is varied to get the required SIR. We took signal attenuations $\mathbf{h}_{ii} = 1$ and interference filter coefficients \mathbf{h}_{ij} to be of same value. By setting all the tap-weight vectors initially to zero, and providing a constraint $1 - H_{12}(z)H_{21}(z) \neq 0$ we search for the optimum \mathbf{W}_{opt} . The learning processes of LMS algorithm and the forward/backward bootstrapped algorithm are compared for different cross channel coupling levels.

In order to compare the two algorithms with the same number of adaptive filters, the weight vectors $\mathbf{w}_{11} = \mathbf{w}_{22}$ for LMS are taken to be unity, that is $\mathbf{w}_{ii} = [1, 0..0]^T$.

V. RESULTS

In this section, we present the results of the computer simulations for different SIR. In Fig. 3 and Fig. 4, we present the interference of the power residue. The learning curves with input SIR = 10.0, 7.6 and 5.6 dB for LMS and F/B bootstrap respectively. The convergence constants, μ used were 0.08 and 0.04 for F/B bootstrap and LMS, respectively. These constant values were chosen to be slightly less than their maximum allowable values. In Fig. 5 we compare the output power residues of the LMS to those of the bootstrap. Note that output Y_2 of the bootstrap has high residue due to co channel distortion as is reflected in (17). Channel equalization will reduce this distortion and make power residue comparable to that of output y_1 .

VI. CONCLUSION

From this study, we conclude that forward/backward structure of the bootstrapped algorithm might be proposed as a multi-signal separator. However, some, but not all, of the outputs of the separator require channel equalizers. For total signals separation, $h_{12}(n)$ and $h_{21}(n)$ are FIR filters, , then the adaptive filters, $w_{21}(n)$ and $w_{12}(n)$ must be FIR and IIR adaptive filters, respectively (see Fig 1). For the system to be stable we must have $1 - H_{12}(z)H_{21}(z) \neq 0$. That is, there should be no solution on the unit circle. It is clear from Fig. 5 that the Forward/Backward separator outperforms the LMS separator despite the fact that the reference signal was not used in the former.

VIII. REFERENCES

- [1] Bar-Ness, Y., "Bootstrapped Adaptive Cross-Pol Interference Cancelling Techniques- Steady State Analysis," *submitted for publication to the Signal Processing*, Elsevier Science Publishers, Nov. 27 1991. (accepted pending revision Aug. 1992)

- [2] Dinç, A. and Bar-Ness, Y., " Error Probabilities of Bootstrapped Blind Adaptive Cross-Pol Cancellers For M-ary QAM over Non-dispersive Fading Channel," *Proceedings of International Conf. on Communications* , ICC '92, paper 353.5, Chicago, Ill. June 1992.
- [3] Dinç, A. and Bar-Ness, Y., " Bootstrap: A Fast Blind Adaptive Signal Separator." *Proceedings of ICASSP '92*, paper 43.8, San Francisco, CA, March 1992.
- [4] Bar-Ness, Y., " Bootstrapped Adaptive Blind Signal Separator Techniques-Bandwidth-Complexity Tradeoff ", submitted to *Signal Processing*, Elsevier Science Publishers. July 27, 1992.
- [5] Jutten, C, et. al , " Blind Separation of Sources: An Algorithm for Separation of Convolutional Mixtures," *Inter. Signal Process. Workshop on HOS*, pp. 273-276. Chamrousse. France, July 1991.
- [6] Weinstein, E., Feder, M. and Oppenheim, A. V., " Multi-Channel Signal Separation Based on Decorrelation " *submitted to IEEE Trans. on ASSP* , 1992.
- [7] Haykin, S., *Adaptive Filter Theory*, Prentice Hall 1991.

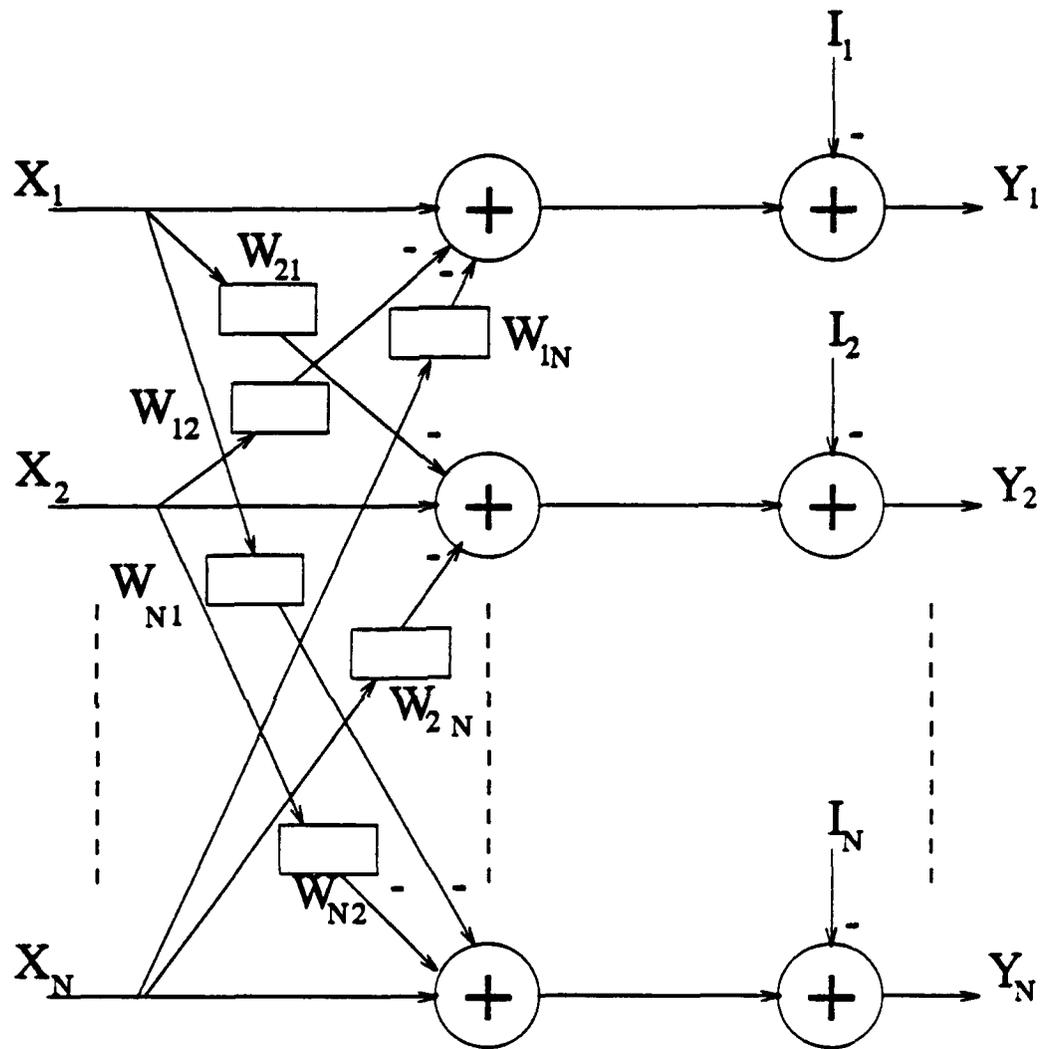


Figure 1 LMS canceller.

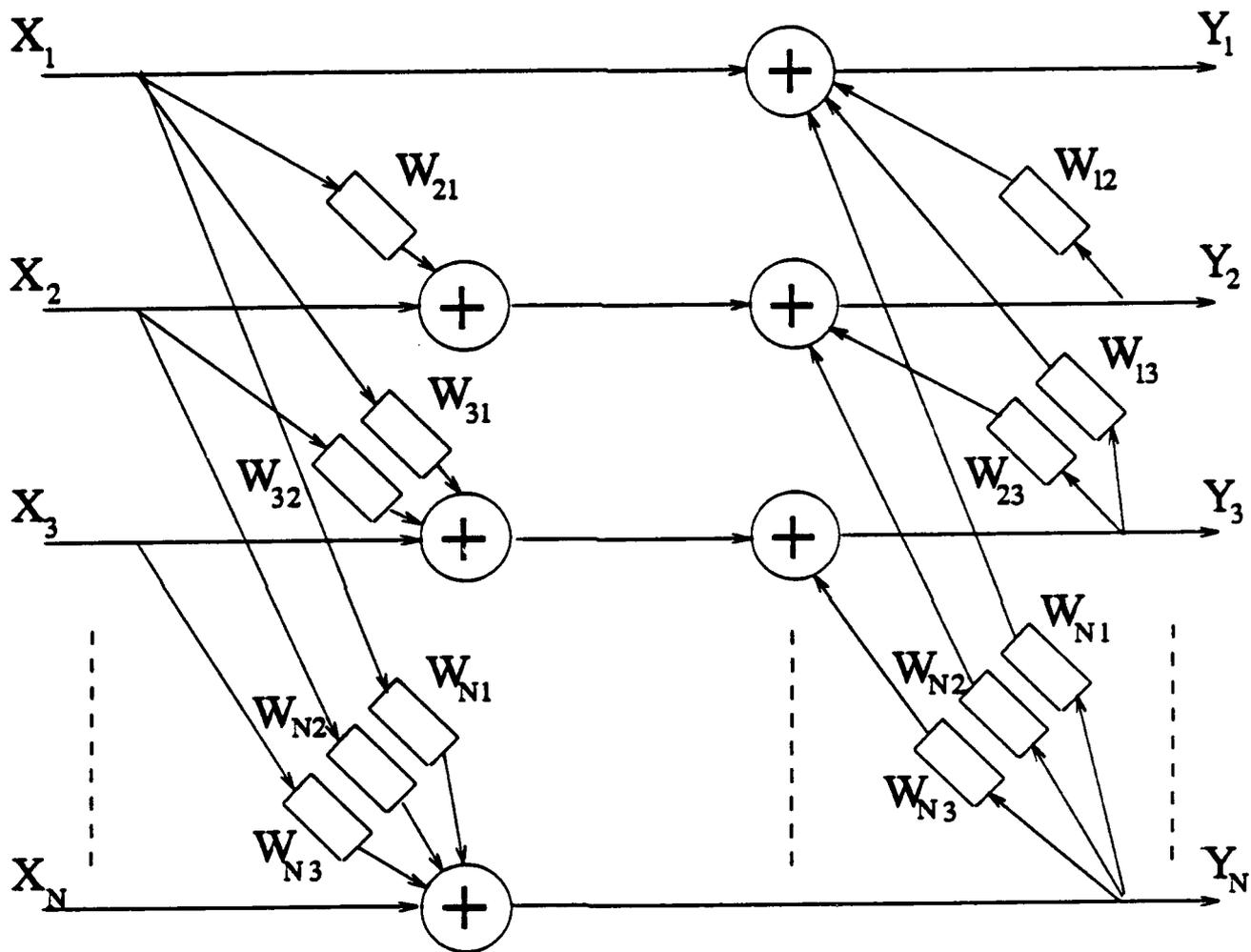


Figure 2 Forward/Backward Bootstrapped Canceller.

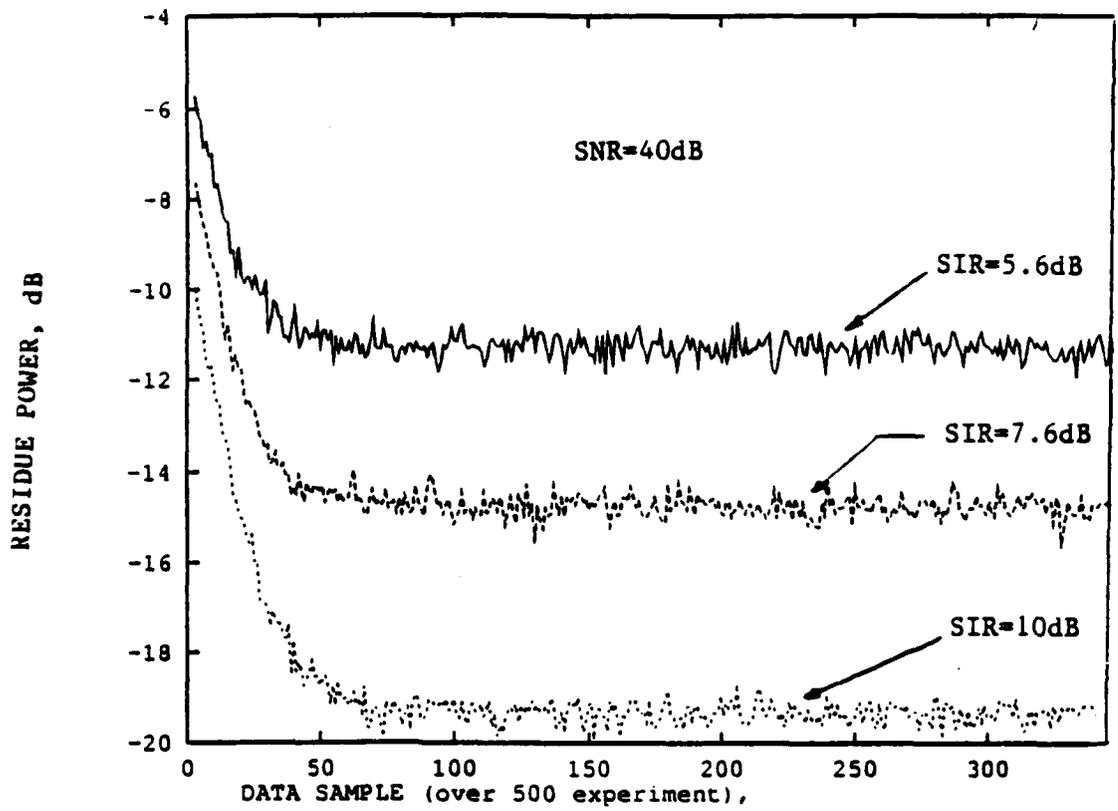


Figure 3 LMS, two-dimensional signal separator.

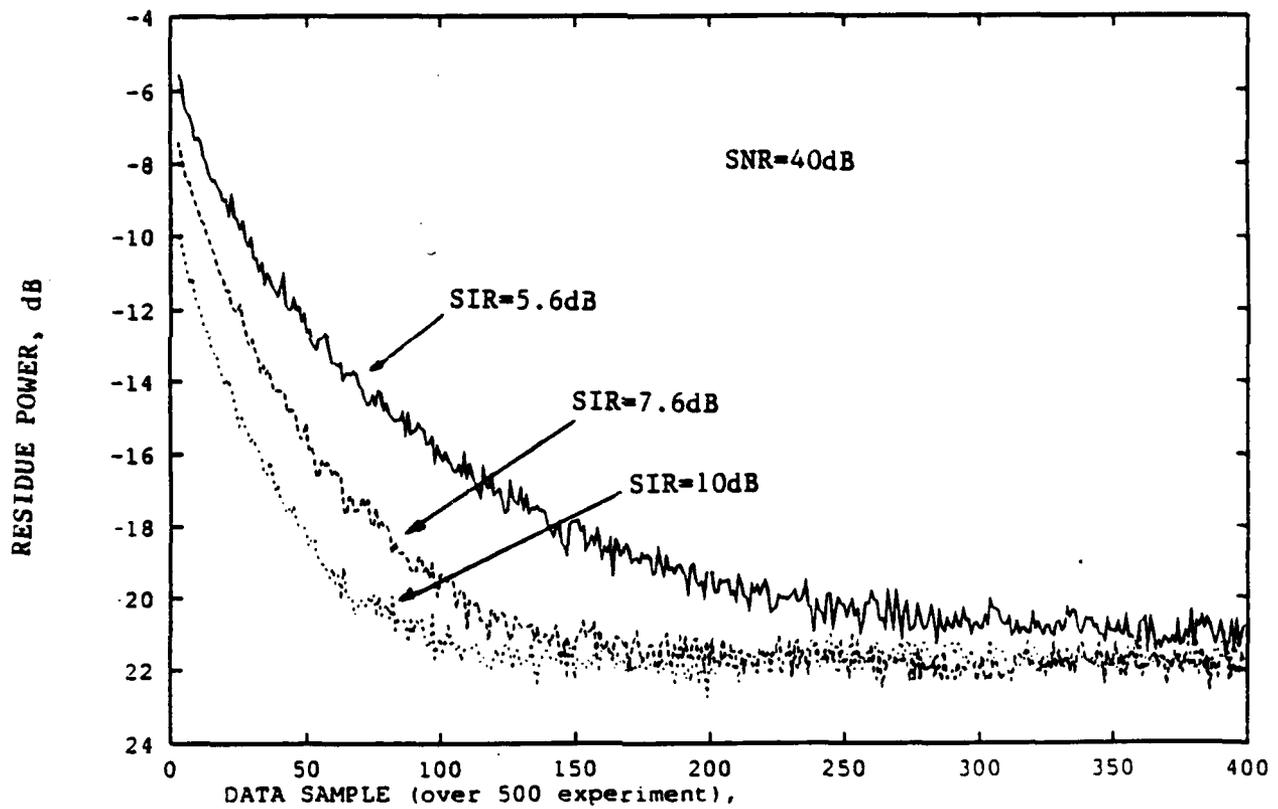


Figure 4 Output Y_1 of the Forward/Backward 2 dimensional signal separator.

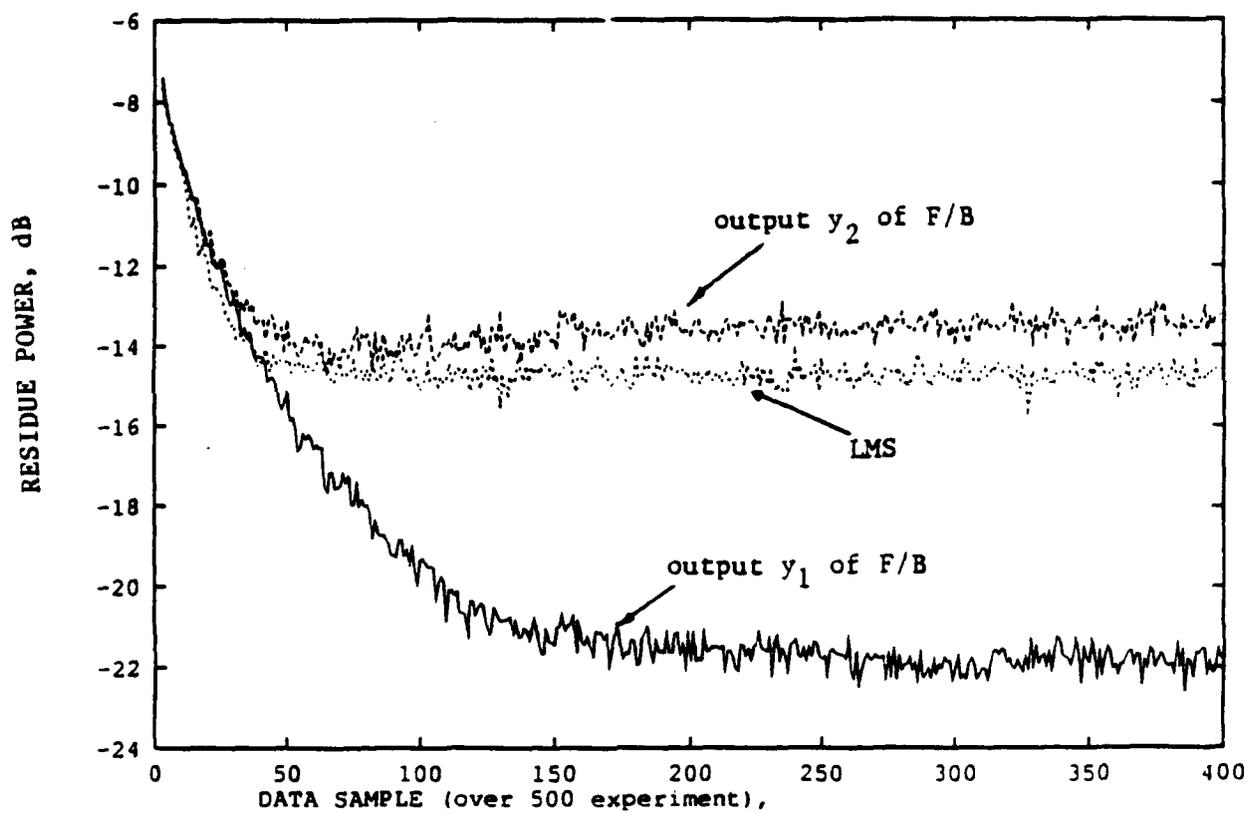


Figure 5 LMS and F/B separators performance comparison SNR = 40dB, SIR=7.6 dB.

APPENDIX H
BLIND DECISION FEEDBACK EQUALIZATION
USING THE DECORRELATION CRITERION ¹

by

Raafat E. Kamel and Yeheskel Bar-Ness

ABSTRACT

A new blind equalization algorithm is presented which is based on decorrelating the equalizer's output. The algorithm is used with a decision feedback structure. The performance of the new equalizer on nonrecursive channels is illustrated. The resulting equalizer is globally convergent.

¹This work will be presented at Mini Conference Globecom '93 Nov. 1993.

This work was partially supported by a grant from Rome Air Force Lab. (AFSC), Griffis Air Force Base, N.Y., under contract F30602-88-D-0025, Task C-2-2404.

I. INTRODUCTION

Adaptive equalization has been used to mitigate inter-symbol interference (*ISI*) in high rate data communication systems. Conventionally, the process is initiated by sending a training sequence which helps the equalizer adapt to the unknown channel characteristics. At the end of the training period, the equalizer operates in a "decision directed mode." That is, the detected data is used as a reference in the adaptation.

Blind equalizers are those which do not require a training sequence. Such equalizers are important in several data communication scenarios, such as multi-point networks [1] or fading channels where sending a training sequence is inappropriate.

It is important to emphasize that existing blind equalizers are the finite length linear FIR type, whose tap gains are updated by minimizing a cost function that merely depends on the channel response and the statistical properties of the transmitted sequence. Therefore, the quality of the resulting equalization depends directly on the level of ISI at the output of the channel.

Most of the previous research on blind equalization was devoted to designing cost functions for updating the equalizer's weights. The first known blind equalization algorithm was introduced by Sato [2]. The cost function used by Sato was generalized by Godard into a class of algorithms [1] which involved higher-order statistics of the transmitted sequence. Later it was found [3] that the Sato and Godard algorithms suffer from ill-convergence. They converge to local minima and, hence, are incapable of reducing ISI.

Clearly to eliminate ISI, using linear equalizers, one must achieve the channel inverse. Such an equalizer is called a zero-forcing (*ZF*) equalizer [4]. The *ZF* equalizer is known to enhance noise at frequencies where the channel spectrum has high attenuation. This is undesirable for channels that are subject to frequency selective fades, *e.g.*, radio channels. On the other hand a finite length linear equalizer can only approximate the inverse channel response, and for the non-minimum phase system the inverse filter is unstable.

The above factors have motivated researchers to use the decision feedback equalizer (*DFE*), depicted in Fig 1., which can compensate for amplitude distortion with minimal noise enhancement [5]. In fact, with this equalizer structure, the forward filter $C(z)$ need not approximate the inverse of the channel and, thus, it avoids noise enhancement. The feedback filter $D(z)$ is used to remove part of the ISI caused by previously detected symbols. In this paper we restrict ourselves to nonrecursive channels [moving-average (MA) models], since we are interested in frequency-selective radio channels, *i.e.*, radio links with multipath. Extension to a more general channel is also possible.

Conventional DFE uses a preamble during a training period and switches to a decision directed mode during data transmission. Using the decision directed mode in blind equalization is not advisable since, if the initial eye is closed (*i.e.*, the initial error rate is high), the equalizer can converge to equilibria that are far from removing ISI.

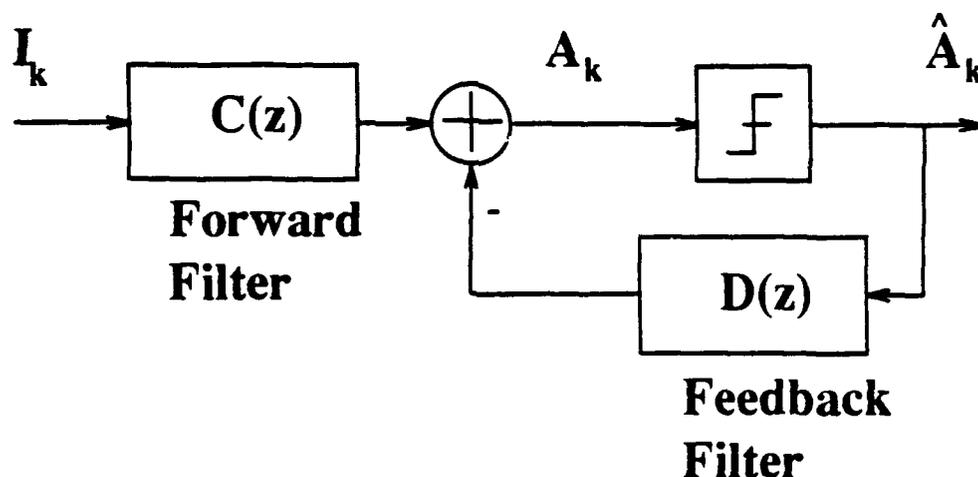


Figure 1 Decision Feedback Equalizer.

This paper discusses a blind equalization algorithm for use with the decision feedback equalizer. This algorithm is based on decorrelating the sequence at the input of the slicer. We show that this algorithm converges to the correct equilibrium despite error propagation. The paper does not discuss the effect of noise on the performance of the equalizer. In section 2 we introduce the criterion and algorithm used for the blind decision feedback equalizer. In

section 3 we discuss the dynamics of the proposed algorithm and show the effect of error propagation on the performance of the algorithm. An example is given in section 4 and conclusions in section 5.

II. THE DECORRELATION CRITERION AND ALGORITHM

The channel and equalizer model under consideration is shown in Fig. 2. The cascade of the transmit, channel, and receive filters is modelled as an FIR filter with impulse response

$$h(n) = 1 + \sum_{i=1}^N h_i \delta(n - i)$$

where $\delta(\cdot)$ is the kronecker delta. In the above equation we normalized relative to the first cursor (h_0). A more general model would merely differ by a gain. We also assume that the input I_k is a binary white sequence with a zero mean. The channel's output is thus given by

$$X_k = I_k + \sum_{i=1}^N h_i I_{k-i}$$

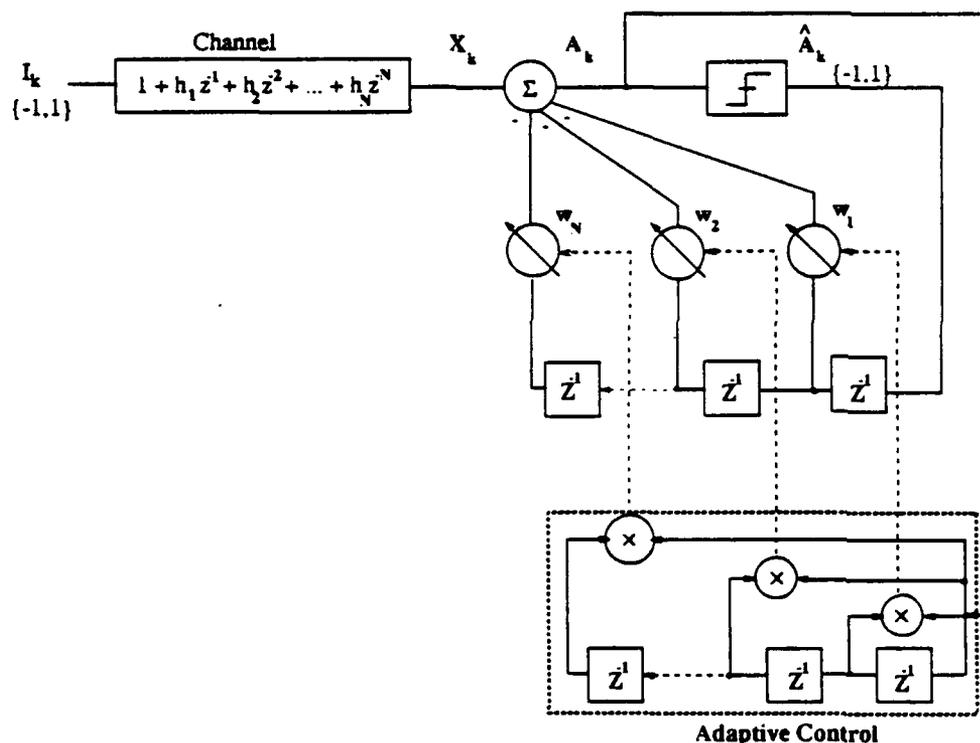


Figure 2 Channel and DFE model.

We assume the channel is slowly changing with time and the receiver has perfect carrier and timing information. Referring to Fig. 2, the input to the slicer of the decision feedback equalizer, A_k , is given by

$$\begin{aligned} A_k &= X_k - \hat{\mathbf{A}}'_{k-1} \mathbf{W} \\ &= I_k + \mathbf{I}'_{k-1} \mathbf{H} - \hat{\mathbf{A}}'_{k-1} \mathbf{W}, \end{aligned} \quad (1)$$

where X_k is the output of a moving average (MA) type channel of order $N + 1$ and is given by

$$X_k = I_k + \sum_{i=1}^N h_i I_{k-i},$$

$\hat{\mathbf{A}}_{k-1}$ is the vector of the past N decisions

$$\hat{\mathbf{A}}'_{k-1} = [\hat{A}_{k-1}, \hat{A}_{k-2}, \dots, \hat{A}_{k-N}].$$

The prime stands for transpose, \mathbf{I}_{k-1} is the vector of past transmitted information bits

$$\mathbf{I}'_{k-1} = [I_{k-1}, I_{k-2}, \dots, I_{k-N}],$$

where I_{k-i} takes values from the binary alphabet $\{-1, 1\}$ with equal probability. \mathbf{W} and \mathbf{H} are the equalizer and channel parameter vectors respectively;

$$\mathbf{W}' = [w_1, w_2, \dots, w_N]$$

$$\mathbf{H}' = [h_1, h_2, \dots, h_N].$$

Multiply eq. (1) by \mathbf{A}_{k-1} , the vector of past slicer's input, where

$$\mathbf{A}_{k-1} = [A_{k-1}, A_{k-2}, \dots, A_{k-N}]$$

to obtain

$$A_k \mathbf{A}_{k-1} = I_{k-1} \mathbf{A}_{k-1} + \mathbf{A}_{k-1} \mathbf{I}'_{k-2} \mathbf{H} - \mathbf{A}_{k-1} \hat{\mathbf{A}}'_{k-1} \mathbf{W}. \quad (2)$$

When taking the expectation of (2), the first term vanishes since

$$E\{A_{k-m}I_{k-n}\} = 0 \text{ for } m > n,$$

as A_k does not depend on the future data inputs. It can also be shown that

$$E\{A_{k-m}\hat{A}_{k-n}\} = 0 \text{ for } m > n.$$

Therefore,

$$\begin{aligned} & \begin{pmatrix} \overline{A_k A_{k-1}} \\ \overline{A_k A_{k-2}} \\ \vdots \\ \overline{A_k A_{k-N}} \end{pmatrix} \\ &= \begin{pmatrix} \overline{A_{k-1} I_{k-1}} & \overline{A_{k-1} I_{k-2}} & \cdots & \overline{A_{k-1} I_{k-N}} \\ 0 & \overline{A_{k-2} I_{k-2}} & \cdots & \overline{A_{k-2} I_{k-N}} \\ \vdots & & & \\ 0 & & \cdots & \overline{A_{k-N} I_{k-N}} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{pmatrix} \\ &- \begin{pmatrix} \overline{A_{k-1} \hat{A}_{k-1}} & \overline{A_{k-1} \hat{A}_{k-2}} & \cdots & \overline{A_{k-1} \hat{A}_{k-N}} \\ 0 & \overline{A_{k-2} \hat{A}_{k-2}} & \cdots & \overline{A_{k-2} \hat{A}_{k-N}} \\ \vdots & & & \\ 0 & 0 & \cdots & \overline{A_{k-N} \hat{A}_{k-N}} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{pmatrix}. \end{aligned} \tag{3}$$

The last entry of (3) can be written as

$$\overline{A_k A_{k-N}} = \overline{A_{k-N} I_{k-N}} h_N - \overline{|A_{k-N}|} w_N.$$

From (1) we get

$$\overline{A_{k-N} I_{k-N}} = \overline{I_{k-N}^2} = \sigma_I^2 = 1.$$

Therefore

$$\begin{aligned} \overline{A_k A_{k-N}} &= h_N - \overline{|A_{k-N}|} w_N \\ w_N &= \frac{h_N - \overline{A_k A_{k-N}}}{\overline{|A_{k-N}|}} \\ &= h_N - \overline{A_k A_{k-N}}, \end{aligned} \tag{4}$$

where it can be shown that

$$\overline{|A_{k-n}|} = 1 \quad \forall n.$$

It is clear from eqn. (4) that $w_N = h_N$ iff $\overline{A_k A_{k-N}} = 0$, i.e., A_k and A_{k-N} are decorrelated.

Next, consider the $(N-1)$ th entry

$$\begin{aligned} \overline{A_k A_{k-N+1}} &= \overline{A_{k-N+1} I_{k-N-1} h_{N-1}} \\ &\quad + \overline{A_{k-N+1} I_{k-N} h_N} - \overline{A_{k-N+1} \hat{A}_{k-N+1} w_{N-1}} \\ &\quad - \overline{A_{k-N+1} \hat{A}_{k-N} w_N}. \end{aligned}$$

But $\overline{A_{k-N+1} \hat{A}_{k-N+1}} = \overline{|A_{k-N+1}|} = 1$ and $\overline{A_{k-N+1} I_{k-N}} = \sigma_I^2 = 1$. Furthermore, if $w_N = h_N$ then

$$\begin{aligned} \overline{A_k A_{k-N+1}} &= h_{N-1} - w_{N-1} + \overline{A_{k-N+1} (I_{k-N} - \hat{A}_{k-N}) h_N} \\ &= h_{N-1} - w_{N-1} + \overline{A_{k-N+1} e_{k-N} h_N}, \end{aligned}$$

where $e_k \triangleq I_k - \hat{A}_k$. We can also show that

$$\overline{A_k A_{k-N+1}} = h_{N-1} - w_{N-1} + 2q_{k-N} h_N (h_1 + w_1), \quad (5)$$

where $q_k \triangleq P\{e_k \neq 0\}$. For the m th entry we have

$$\begin{aligned} \overline{A_k A_{k-m}} &= \overline{A_{k-m} I_{k-m} h_m} + \sum_{i=m+1}^N h_i \overline{A_{k-m} I_{k-i}} \\ &\quad - w_m \overline{|A_{k-m}|} - \sum_{i=m+1}^N w_i \overline{A_{k-m} \hat{A}_{k-i}}. \end{aligned}$$

Following the above argument and assuming $w_i = h_i$ for $i = m+1, \dots, N$ then

$$\begin{aligned} \overline{A_k A_{k-m}} &= h_m - w_m + \sum_{i=m+1}^N h_i \overline{A_{k-m} (I_{k-i} - \hat{A}_{k-i})} \\ &= h_m - w_m + \sum_{i=m+1}^N h_i \overline{A_{k-m} e_{k-i}}. \end{aligned}$$

It can also be shown that

$$\overline{A_k A_{k-m}} = h_m - w_m + 2q_{k-i} \sum_{i=m+1}^N h_i (h_{i-m} + w_{i-m}). \quad (6)$$

To summarize, we combine (4), (5) and (6) as

$$\begin{aligned} \overline{A_k A_{k-N}} &= h_N - w_N \\ \overline{A_k A_{k-N+1}} &= 2q_{k-N} h_N (h_1 + w_1) + h_{N-1} - w_{N-1} \\ \overline{A_k A_{k-N+2}} &= 2q_{k-N} h_N (h_2 + w_2) + \\ &\quad 2q_{k-N+1} h_{N-1} (h_1 + w_1) + h_{N-2} - w_{N-2} \\ &\quad \vdots \\ \overline{A_k A_{k-1}} &= 2q_{k-N} h_N (h_{N-1} + w_{N-1}) + \\ &\quad 2q_{k-N+1} h_{N-1} (h_{N-2} + w_{N-2}) \\ &\quad + \cdots + 2q_{k-2} h_2 (h_1 + w_1) + h_1 - w_1. \end{aligned} \quad (7)$$

The above equation is based on the assumption that the previous weights converged to the correct channel parameters. It is clear from (4) that $w_N = h_N$ iff $\overline{A_k A_{k-N}} = 0$. Further if q_{k-N} is zero, i.e., \hat{A}_{k-N} was correct, then $w_{N-1} = h_{N-1}$ iff $\overline{A_k A_{k-N+1}} = 0$. Similarly one can proceed and reason that if the $N - m$ previous decisions ($\hat{A}_{k-N+m}, m = 1, \dots, N - 1$) were correct then $w_m = h_m$ for $m = 1, \dots, N$ iff $\overline{A_k A_{k-m}} = 0$. This leads to $\mathbf{W} = \mathbf{H}$ together with $\hat{A}_{k-m}, m = 1, \dots, N$ being correct. We have from (1) that \hat{A}_k is the correct decision and, hence, following similar reasoning \hat{A}_{k+m} for $m > 0$ will be correct. This means that if we reach zero probability of error, the algorithm will continue to be at steady state of no ISI, provided that A_k is decorrelated with the previous N slicer inputs.

If, on the other hand, some of the previous decisions were erroneous, then the algorithm

is still in the transient state. The behavior of the algorithm in the transient state and its convergence to a zero probability of error will be discussed in the next section.

To control the weights in order to decorrelate A_k 's we use the steepest descent method. Any stochastic gradient algorithm can be expressed as

$$w_i^{k+1} = w_i^k + \mu f(\cdot) \text{ for } i = 1, \dots, N, \quad (8)$$

where μ is the constant of adaptation and $f(\cdot)$ is called the error function of the algorithm. The roots of the error function determine the steady state of the algorithm values to which the weights will converge.

The previous discussion suggests using $\overline{A_k A_{k-i}}$ as an error function for the algorithm. An appropriate error function in eqn. (8) would be $f(\overline{A_k A_{k-i}})$ such that $f(0) = 0$ i.e. $f(x)$ should have a root at zero. Since the roots of the error function represent the equilibrium points for the algorithm, some of which might not be the optimum, an error function with a distinct root at zero would be preferred. Therefore, a possible function would be

$$f(x) = x.$$

As a result, one can write eqn. (8) as

$$w_i^{k+1} = w_i^k + \mu \overline{A_k A_{k-i}} \text{ for } i = 1, \dots, N.$$

In a practical implementation one would estimate the expectation by its current realization, leading to the stochastic difference equation,

$$w_i^{k+1} = w_i^k + \mu A_k A_{k-i} \text{ for } i = 1, \dots, N. \quad (9)$$

It is clear from the above analysis that the algorithm in eqn. (9) will converge in the mean. i.e., the mean value of w_i will converge to the channel parameter h_i .

III. DYNAMIC BEHAVIOR

In this section we examine the dynamic behavior of the proposed equalizer. We use the probability of symbol error, q_k , at instant k as a performance index. A difference equation for q_k is developed, which can be solved together with the weight update equation to determine the probability of decision error as a function of time index k . The derivation given in this section can be extended to a general order N moving average type (MA) channel. However, for the sake of simplicity we consider an order 3 MA channel. (The same channel will be used in the simulation described in the next section.)

The channel output X_k at the k th instant is given by

$$X_k = I_k + h_1 I_{k-1} + h_2 I_{k-2}, \quad (10)$$

where h_1 and h_2 are the channel parameters. From (1) the slicer's input is given by

$$\begin{aligned} A_k &= X_{k-1} - w_1^{(k)} \hat{A}_{k-1} - w_2^{(k)} \hat{A}_{k-2} \\ &= I_{k-1} + h_1 I_{k-2} + h_2 I_{k-3} - w_1^{(k)} \hat{A}_{k-1} - w_2^{(k)} \hat{A}_{k-2}, \end{aligned} \quad (11)$$

where we have used the superscript k in the weights w_1 and w_2 to emphasize their dependence on time, since we are studying the transient response of the algorithm. Using eqn. (11) we will determine the probability of correct decision (q_k) as a function of the index k . Using the total probability theorem, one can write

$$\begin{aligned} q_k &= P\{\hat{A}_k \neq I_{k-1}\} \\ &= P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} \neq I_{k-2}, \hat{A}_{k-2} \neq I_{k-3}\} q_{k-1} q_{k-2} \\ &+ P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} \neq I_{k-2}, \hat{A}_{k-2} = I_{k-3}\} q_{k-1} p_{k-2} \\ &+ P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} = I_{k-2}, \hat{A}_{k-2} \neq I_{k-3}\} p_{k-1} q_{k-2} \\ &+ P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} = I_{k-2}, \hat{A}_{k-2} = I_{k-3}\} p_{k-1} p_{k-2}, \end{aligned} \quad (12)$$

where we have

$$\begin{aligned}
& P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} \neq I_{k-2}, \hat{A}_{k-2} \neq I_{k-3}\} \\
&= \frac{1}{2} \left(P \left\{ I_{k-2} > \frac{1 - (h_2 + w_2^{(k)})}{|h_1 + w_1^{(k)}|} \right\} \right. \\
&\quad \left. + P \left\{ I_{k-2} > \frac{1 + (h_2 + w_2^{(k)})}{|h_1 + w_1^{(k)}|} \right\} \right) \\
&\triangleq f_1(h_1, h_2, w_1^{(k)}, w_2^{(k)})
\end{aligned} \tag{13}$$

and:

$$\begin{aligned}
& P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} \neq I_{k-2}, \hat{A}_{k-2} = I_{k-3}\} \\
&= \frac{1}{2} \left(P \left\{ I_{k-2} > \frac{1 - (h_2 - w_2^{(k)})}{|h_1 + w_1^{(k)}|} \right\} \right. \\
&\quad \left. + P \left\{ I_{k-2} > \frac{1 + (h_2 - w_2^{(k)})}{|h_1 + w_1^{(k)}|} \right\} \right) \\
&\triangleq f_2(h_1, h_2, w_1^{(k)}, w_2^{(k)}).
\end{aligned} \tag{14}$$

Also,

$$\begin{aligned}
& P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} = I_{k-2}, \hat{A}_{k-2} \neq I_{k-3}\} \\
&= \frac{1}{2} \left(P \left\{ I_{k-2} > \frac{1 - (h_2 + w_2^{(k)})}{|h_1 - w_1^{(k)}|} \right\} \right. \\
&\quad \left. + P \left\{ I_{k-2} > \frac{1 + (h_2 + w_2^{(k)})}{|h_1 - w_1^{(k)}|} \right\} \right) \\
&\triangleq f_3(h_1, h_2, w_1^{(k)}, w_2^{(k)}),
\end{aligned} \tag{15}$$

and

$$\begin{aligned}
& P\{\hat{A}_k \neq I_{k-1} \mid \hat{A}_{k-1} = I_{k-2}, \hat{A}_{k-2} = I_{k-3}\} \\
&= \frac{1}{2} \left(P \left\{ I_{k-2} > \frac{1 - (h_2 - w_2^{(k)})}{|h_1 - w_1^{(k)}|} \right\} \right. \\
&\quad \left. + P \left\{ I_{k-2} > \frac{1 + (h_2 - w_2^{(k)})}{|h_1 - w_1^{(k)}|} \right\} \right) \\
&\triangleq f_4(h_1, h_2, w_1^{(k)}, w_2^{(k)}).
\end{aligned} \tag{16}$$

Substituting the above in equation (12)

$$\begin{aligned}
 q_k &= q_{k-1}q_{k-2}f_1(h_1, h_2, w_1^{(k)}, w_2^{(k)}) \\
 &+ q_{k-1}p_{k-2}f_2(h_1, h_2, w_1^{(k)}, w_2^{(k)}) \\
 &+ p_{k-1}q_{k-2}f_3(h_1, h_2, w_1^{(k)}, w_2^{(k)}) \\
 &+ p_{k-1}p_{k-2}f_4(h_1, h_2, w_1^{(k)}, w_2^{(k)}), \tag{17}
 \end{aligned}$$

where p_k is the probability of a correct decision. Eqn. (17) is a second order difference equation, which depends on the channel parameters h_1 and h_2 and on the current equalizer's weights $w_1^{(k)}$ and $w_2^{(k)}$. For the more general order N , channel eqn. (17) will take the general form

$$q_k = f(q_{k-1}, \dots, q_{k-N}, h_1, \dots, h_N, w_1^{(k)}, \dots, w_N^{(k)}) \tag{18}$$

The instantaneous probability of error may be computed recursively using eqn. (18), weights update eqn. (9), and the appropriate initial conditions for the probability of error. Eqn. (18) is highly nonlinear; therefore, only low-order channels are numerically tractable for showing the convergence of q_k to zero.

IV. ILLUSTRATIVE EXAMPLE

In studying the dynamic behavior of the blind decision feedback equalizer and examining the convergence of q_k in eqn. (17), we will consider two approaches.

In the first we will use the mean of the weights, *i.e.*, the expected values of $w_1^{(k)}$ and $w_2^{(k)}$. The controlling algorithm will be

$$\overline{w_i^{(k+1)}} = \overline{w_i^{(k)}} + \mu \overline{A_k A_{k-i}} \quad \text{for } i = 1, 2, \tag{19}$$

and by substituting from eqn. (6) we get

$$\begin{aligned}
 \overline{w_2^{(k+1)}} &= \overline{w_2^{(k)}} + \mu(h_2 - \overline{w_2^{(k)}}) \\
 \overline{w_1^{(k+1)}} &= \overline{w_1^{(k)}} + \mu((h_1 - \overline{w_1^{(k)}}) + 2q_{k-2}h_2(h_1 + \overline{w_1^{(k)}})).
 \end{aligned}$$

Thus, the result of numerical analysis will show the behavior of the algorithm in the mean sense. While in the other approach we use the stochastic control of eqn. (9) directly.

As an example, consider the non-minimum phase response of

$$H(z^{-1}) = 1 + 0.5z^{-1} - 1.44z^{-2},$$

and depict in Fig. 3 and Fig. 4 the probability of error q_k as a function of k . The initial probability of error used was $q_{-2} = q_{-1} = \frac{1}{2}$, ($q_0 = \frac{1}{4}$). In Fig. 3 we used eqn. (19) for updating the weights, while in Fig. 4 we applied the stochastic estimate of the correlation as in eqn. (9)

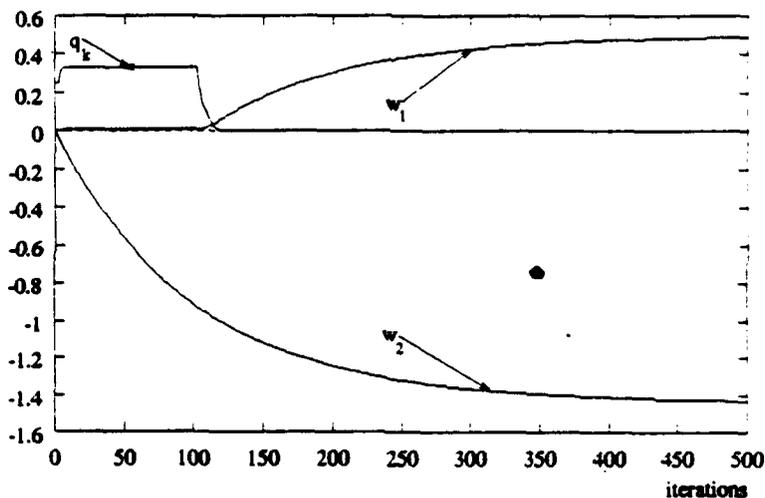


Figure 3 Probability of error for eqn. (19)

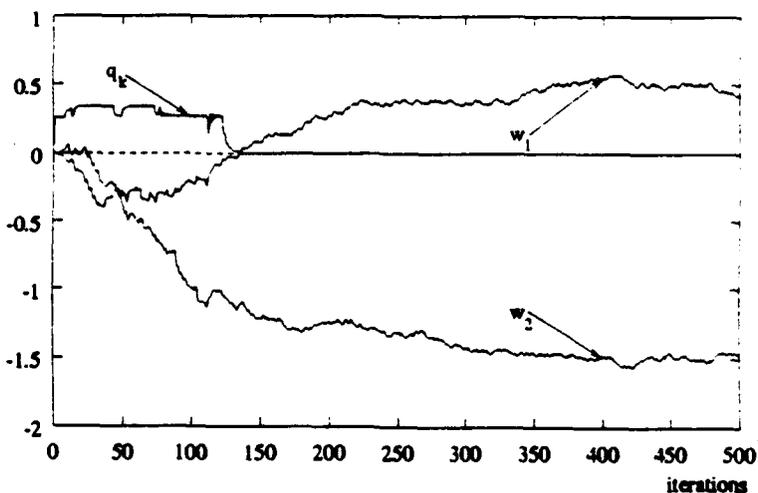


Figure 4 Probability of error for eqn. (9)

With both approaches we notice that q_k converges to zero after a certain number of iterations. Therefore, all the q terms in eqn. (7) vanish and the weights w_i will converge to h_i iff $\overline{A_k A_{k-i}}$ for $i = 1, 2$. From Figs. 3 and 4 one would note that the probability of error q_k approaches zero even before the equalizer convergence.

The channel above was used in a computer simulation. The weights of the feedback filter w_1 and w_2 are depicted in Fig. (4), the adaptation constant used was $\mu = 0.01$. In this simulation the weights were initialized to zero. Fig. (4) shows that the equalizer's weights converge to the channel parameters (in this case $w_1 = 0.5$ and $w_2 = -1.44$). Next, by varying the initial settings of the equalizer weights, we show that the algorithm always converges to the right point (0.5,-1.44) regardless of the initial condition. Fig. (5) shows the trajectories for the different initializations and shows that the decorrelation algorithm is globally convergent for the channel under consideration.

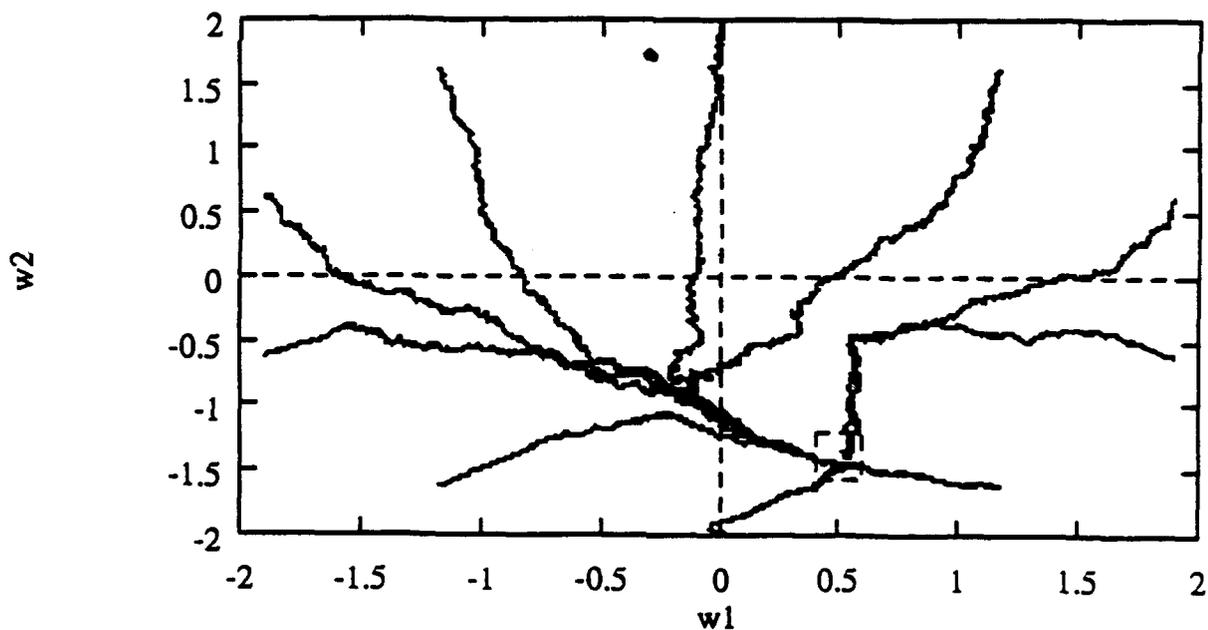


Figure 5 Admissibility

V. CONCLUSIONS

In this paper we introduced a new criterion and an algorithm for blind equalization. The algorithm was used in conjunction with a decision feedback equalizer. It decorrelates the data sequence at the input of the slicer. It was shown to converge to the optimum irrespective of the initial error rate.

Because of the feedback structure the equalizer does not suffer from noise enhancement as the linear equalizer does. A simulation example of a dispersive non-minimum phase channel was given to illustrate the convergence of the algorithm. With an adaptation constant μ of 0.01 the algorithm converges after 200 iterations. The simulation also shows that the algorithm converges to the optimum point regardless of the initial setting.

VI. APPENDIX

Claim 1 The probability density function $f_{A_k}(\cdot)$ of the random variable A_k defined in eq. (1) is an even function.

Proof

The input to the slicer in eq. (1) A_k is given by

$$\begin{aligned} A_k &= X_k - \sum_{i=1}^N w_i \hat{A}_{k-i} \\ &= I_k + \sum_{i=1}^N (h_i I_{k-i} - w_i \hat{A}_{k-i}). \end{aligned} \quad (\text{A.1})$$

If we denote the set of all correct decisions by A' and the set of all incorrect decisions by A'' .
i.e.,

$$\begin{aligned} A' &= \{\hat{A}_i : \hat{A}_i = I_i\} \\ A'' &= \{\hat{A}_i : \hat{A}_i = -I_i\}, \end{aligned}$$

then the input of the slicer in (A.1) can be written as

$$A_k = I_k + \sum_{i: \hat{A}_{k-i} \in A'} (h_i - w_i) I_{k-i} + \sum_{i: \hat{A}_{k-i} \in A''} (h_i + w_i) I_{k-i}. \quad (\text{A.2})$$

From the above equation one can see that A_k can be expressed as a sum of independent random variables. Therefore, the probability density function (pdf) of A_k is the convolution of the individual pdfs, thus,

$$f_{A_k} = f_{I_k} * \text{Conv}_{i: \hat{A}_{k-i} \in A'} f_{(h_i - w_i) I_{k-i}} * \text{Conv}_{i: \hat{A}_{k-i} \in A''} f_{(h_i + w_i) I_{k-i}}, \quad (\text{A.3})$$

where $\text{Conv}_{i: \hat{A}_{k-i} \in A'}$ and $\text{Conv}_{i: \hat{A}_{k-i} \in A''}$ are the convolution of the probability density functions of the corresponding random variables in the summations of eq. (A.3). Since I_k 's are random variables taking values of -1 and 1 with equal probabilities we have

$$\begin{aligned} f_{I_k}(x) &= \frac{1}{2} (\delta(x+1) + \delta(x-1)) \\ f_{(h_i - w_i) I_{k-i}}(x) &= \frac{1}{2} (\delta(x + h_i - w_i) + \delta(x - h_i + w_i)) \\ f_{(h_i + w_i) I_{k-i}}(x) &= \frac{1}{2} (\delta(x + h_i + w_i) + \delta(x - h_i - w_i)). \end{aligned}$$

The convolution equation in (A.3) can be transformed into a product form by using the Fourier Transform

$$\mathcal{F}_{A_k} = \mathcal{F}_{I_k} \cdot \prod_{i: \hat{A}_{k-i} \in A'} \mathcal{F}_{(h_i - w_i) I_{k-i}} \cdot \prod_{i: \hat{A}_{k-i} \in A''} \mathcal{F}_{(h_i + w_i) I_{k-i}}, \quad (\text{A.4})$$

where \mathcal{F}_X is the Fourier Transform of the pdf of the random variable X . Therefore, we have

$$\begin{aligned} \mathcal{F}_{I_k}(\omega) &= \cos(\omega) \\ \mathcal{F}_{(h_i - w_i) I_{k-i}}(\omega) &= \cos((h_i - w_i)\omega) \\ \mathcal{F}_{(h_i + w_i) I_{k-i}}(\omega) &= \cos((h_i + w_i)\omega). \end{aligned}$$

Now we consider the product terms in eq. (A.4). The first term,

$$\prod_{i: \hat{A}_{k-i} \in A'} \mathcal{F}_{(h_i - w_i) I_{k-i}}(\omega) = \frac{1}{2^{|A'| - 1}} \sum_{a_i} \cos(a_i \omega),$$

where $|A'|$ is the cardinality of the set A' and a_i s represent all the possible sums and differences among all $(h_i - w_i)$ such that $\hat{A}_{k-i} \in A'$.

Similarly, for the other product term in eq. (A.3) one can write,

$$\prod_{i: \hat{A}_{k-i} \in A''} \mathcal{F}_{(h_i + w_i)I_{k-i}}(\omega) = \frac{1}{2^{|A''|-1}} \sum_{b_i} \cos(b_i \omega),$$

where $|A''|$ is the cardinality of the set A'' and b_i s represent all the possible sums and differences among all $(h_i + w_i)$ such that $\hat{A}_{k-i} \in A''$.

As a result eq. (A.4) can be written as

$$\begin{aligned} \mathcal{F}_{A_k}(\omega) &= \cos(\omega) \cdot \frac{1}{2^{|A'|-1}} \sum_{a_i} \cos(a_i \omega) \cdot \frac{1}{2^{|A''|-1}} \sum_{b_i} \cos(b_i \omega) \\ &= \cos(\omega) \cdot \frac{1}{2^{N-2}} \sum_{a_i} \sum_{b_i} \cos(a_i \omega) \cos(b_i \omega) \text{ since } |A'| + |A''| = N \\ &= \cos(\omega) \cdot \frac{1}{2^{N-1}} \sum_{a_i} \sum_{b_i} \cos((a_i + b_i)\omega) + \cos((a_i - b_i)\omega) \\ &= \cos(\omega) \cdot \frac{1}{2^{N-1}} \sum_{c_i} \cos(c_i \omega), \end{aligned}$$

where c_i represents all the possible pairwise sums and differences of a_i s and b_i s. Further, one can write

$$\mathcal{F}_{A_k} = \frac{1}{2^N} \sum_{c_i} \cos((c_i + 1)\omega) + \cos((c_i - 1)\omega). \quad (\text{A.5})$$

Taking the inverse transform of eq. (A.5), we can write the pdf of A_k as

$$f_{A_k}(x) = \frac{1}{2^{N+1}} \sum_{c_i} (\delta(x - c_i - 1) + \delta(x + c_i + 1) + \delta(x - c_i + 1) + \delta(x + c_i - 1)). \quad (\text{A.6})$$

Therefore, the pdf of A_k is an even function, and it also exhibits half symmetry about ± 1 .

Claim 2

$$\mathbf{E}\{A_{k-m} \hat{A}_{k-n}\} = 0 \quad \text{for } m > n$$

Proof:

Consider the joint cumulative distribution function (CDF) of A_{k-m}, \hat{A}_{k-n} viz $F_{A_{k-m}, \hat{A}_{k-n}}(x, y)$,

$$\begin{aligned}
 F_{A_{k-m}, \hat{A}_{k-n}}(x, y) &= P\{A_{k-m} \leq x, \hat{A}_{k-n} \leq y\} \\
 &= P\{A_{k-m} \leq x, \hat{A}_{k-n} \leq y \mid \hat{A}_{k-n} = I_{k-n}\}p \\
 &\quad + P\{A_{k-m} \leq x, \hat{A}_{k-n} \leq y \mid \hat{A}_{k-n} = -I_{k-n}\}q \\
 &= P\{A_{k-m} \leq x, I_{k-n} \leq y \mid \hat{A}_{k-n} = I_{k-n}\}p \\
 &\quad + P\{A_{k-m} \leq x, -I_{k-n} \leq y \mid \hat{A}_{k-n} = -I_{k-n}\}q,
 \end{aligned}$$

where p is the probability of correct decision and q is the probability of incorrect decision.

$$\begin{aligned}
 F_{A_{k-m}, \hat{A}_{k-n}}(x, y) &= P\{A_{k-m} \leq x \mid \hat{A}_{k-n} = I_{k-n}\}P\{I_{k-n} \leq y \mid \hat{A}_{k-n} = I_{k-n}\}p \\
 &\quad + P\{A_{k-m} \leq x \mid \hat{A}_{k-n} = -I_{k-n}\}P\{-I_{k-n} \leq y \mid \hat{A}_{k-n} = -I_{k-n}\}q \\
 &= P\{A_{k-m} \leq x, \hat{A}_{k-n} = I_{k-n}\}P\{I_{k-n} \leq y \mid \hat{A}_{k-n} = I_{k-n}\} \\
 &\quad + P\{A_{k-m} \leq x, \hat{A}_{k-n} = -I_{k-n}\}P\{-I_{k-n} \leq y \mid \hat{A}_{k-n} = -I_{k-n}\}
 \end{aligned} \tag{A.7}$$

since A_{k-m} is independent of I_{k-n} for $m > n$. By definition

$$\begin{aligned}
 P\{I_{k-n} \leq y \mid \hat{A}_{k-n} = I_{k-n}\} &= \int_{-\infty}^y \delta(\mu - 1)P\{I_{k-n} = 1 \mid \hat{A}_{k-n} = I_{k-n}\} \\
 &\quad + \delta(\mu + 1)P\{I_{k-n} = -1 \mid \hat{A}_{k-n} = I_{k-n}\}d\mu.
 \end{aligned} \tag{A.8}$$

Now, from eq. (1) we write

$$A_{k-n} = I_{k-n} + Y_{k-n},$$

where

$$Y_{k-n} \triangleq \sum_{i=1}^N (h_i I_{k-n-i} - w_i \hat{A}_{k-n-i}).$$

From the definition of \hat{A}_k ,

$$P\{\hat{A}_{k-n} = I_{k-n} \mid I_{k-n} = 1\} = P\{\text{sgn}(A_{k-n}) = I_{k-n} \mid I_{k-n} = 1\}$$

$$= P\{\text{sgn}(1 + Y_{k-n}) = 1\}$$

$$= P\{Y_{k-n} > -1\}.$$

Similarly,

$$P\{\hat{A}_{k-n} = I_{k-n} \mid I_{k-n} = -1\} = P\{Y_{k-n} < 1\}.$$

However from Claim 1 of this appendix the pdf of A_k and, hence, of A_{k-n} and Y_{k-n} is even.

This leads to

$$P\{\hat{A}_{k-n} = I_{k-n} \mid I_{k-n} = 1\} = P\{\hat{A}_{k-n} = I_{k-n} \mid I_{k-n} = -1\}. \quad (\text{A.9})$$

Now, using Bayes law,

$$\begin{aligned} P\{I_{k-n} = 1 \mid \hat{A}_{k-n} = I_{k-n}\} &= \frac{P\{\hat{A}_{k-n} = I_{k-n} \mid I_{k-n} = 1\}P\{I_{k-n} = 1\}}{P\{\hat{A}_{k-n} = I_{k-n}\}} \\ P\{I_{k-n} = -1 \mid \hat{A}_{k-n} = I_{k-n}\} &= \frac{P\{\hat{A}_{k-n} = I_{k-n} \mid I_{k-n} = -1\}P\{I_{k-n} = -1\}}{P\{\hat{A}_{k-n} = I_{k-n}\}}. \end{aligned}$$

Therefore, by using eq. (A.9) we get

$$P\{I_{k-n} = 1 \mid \hat{A}_{k-n} = I_{k-n}\} = P\{I_{k-n} = -1 \mid \hat{A}_{k-n} = I_{k-n}\} = \frac{1}{2}.$$

Hence, we can write

$$\begin{aligned} P\{I_{k-n} \leq y \mid \hat{A}_{k-n} = I_{k-n}\} &= \frac{1}{2} \int_{-\infty}^y (\delta(\mu - 1) + \delta(\mu + 1)) d\mu \\ &= F_{I_{k-n}}(y). \end{aligned} \quad (\text{A.10})$$

Similarly, it can be shown that

$$\begin{aligned} P\{-I_{k-n} \leq y \mid \hat{A}_{k-n} = -I_{k-n}\} &= \frac{1}{2} \int_{-\infty}^y (\delta(\mu - 1) + \delta(\mu + 1)) d\mu \\ &= F_{I_{k-n}}(y). \end{aligned} \quad (\text{A.11})$$

Substituting eqs. (A.10) and (A.11) in (A.7), we get

$$\begin{aligned} F_{A_{k-m}\hat{A}_{k-n}}(x, y) &= P\{A_{k-m} \leq x, \hat{A}_{k-n} = I_{k-n}\}F_{I_{k-n}}(y) \\ &\quad + P\{A_{k-m} \leq x, \hat{A}_{k-n} = -I_{k-n}\}F_{I_{k-n}}(y) \\ &= F_{A_{k-m}}(x)F_{I_{k-n}}(y). \end{aligned}$$

Therefore, the joint pdf of $A_{k-m}\hat{A}_{k-n}$ is given by

$$f_{A_{k-m}\hat{A}_{k-n}} = f_{A_{k-m}}(x)f_{I_{k-n}}(y).$$

Hence,

$$\begin{aligned} E\{A_{k-m}\hat{A}_{k-n}\} &= E\{A_{k-m}\}E\{I_{k-n}\} \\ &= 0. \end{aligned}$$

Claim 3

$$\overline{I_{k-m}\hat{A}_{k-n}} = 0 \quad n \neq m$$

Proof:

We have

$$\begin{aligned} F_{\hat{A}_{k-n}I_{k-m}}(x, y) &= P\{\hat{A}_{k-n} \leq x, I_{k-m} \leq y\} \\ &= P\{I_{k-m} \leq y, \hat{A}_{k-n} \leq x \mid \hat{A}_{k-n} = I_{k-n}\}P\{\hat{A}_{k-n} = I_{k-n}\} \\ &\quad + P\{I_{k-m} \leq y, \hat{A}_{k-n} \leq x \mid \hat{A}_{k-n} = -I_{k-n}\}P\{\hat{A}_{k-n} = -I_{k-n}\} \\ &= P\{I_{k-m} \leq y, I_{k-n} \leq x \mid \hat{A}_{k-n} = I_{k-n}\}p_{k-n} \\ &\quad + P\{I_{k-m} \leq y, I_{k-n} \leq x \mid \hat{A}_{k-n} = -I_{k-n}\}q_{k-n}. \end{aligned}$$

I_{k-m} independent of I_{k-n} for $m \neq n$

$$\begin{aligned} F_{\hat{A}_{k-n}I_{k-m}}(x, y) &= P\{I_{k-m} \leq y \mid \hat{A}_{k-n} = I_{k-n}\}P\{I_{k-n} \leq x \mid \hat{A}_{k-n} = I_{k-n}\}p_{k-n} \\ &\quad + P\{I_{k-m} \leq y \mid \hat{A}_{k-n} = -I_{k-n}\}P\{-I_{k-n} \leq x \mid \hat{A}_{k-n} = -I_{k-n}\}q_{k-n}. \end{aligned}$$

Using eqs. (A.9) and (A.10), we get

$$\begin{aligned} F_{\hat{A}_{k-n}I_{k-m}}(x, y) &= P\{I_{k-m} \leq y, \hat{A}_{k-n} = I_{k-n-1}\}F_{I_{k-n-1}}(x) \\ &\quad + P\{I_{k-m} \leq y, \hat{A}_{k-n} = -I_{k-n-1}\}F_{I_{k-n-1}}(x) \\ &= F_{I_{k-m}}(y)F_{I_{k-n}}(x). \end{aligned} \tag{A.12}$$

Therefore, from eq. (A.12), we can conclude that

$$\begin{aligned} \overline{\hat{A}_{k-n} I_{k-m}} &= \overline{I_{k-m} I_{k-n}} \\ &= 0 \end{aligned} \tag{A.13}$$

Claim 4

$$E\{|A_k|\} = 1 \quad \text{for every } n \tag{A.14}$$

Proof:

The pdf $f_{|A_k|}$ of the random variable $|A_k|$ can be expressed as

$$\begin{aligned} f_{|A_k|}(x) &= \begin{cases} f_{A_k}(x) + f_{A_k}(-x) & x \geq 0 \\ 0 & x < 0 \end{cases} \\ &= \begin{cases} 2f_{A_k}(x) & x \geq 0 \\ 0 & x < 0 \end{cases} \end{aligned}$$

since $f_{A_k}(x)$ is an even function.

Substituting from eq. (A.6)

$$f_{|A_k|} = \begin{cases} \frac{1}{2^{N+1}} \sum c_i (\delta(x - |1 + c_i|) + \delta(x - |1 - c_i|)) & x \geq 0 \\ 0 & x < 0. \end{cases}$$

The above equation is symmetric about $x = 1$, therefore the mean

$$E\{|A_k|\} = 1.$$

Claim 5

$$\text{For } m \neq n \quad \overline{\hat{A}_{k-m} \hat{A}_{k-n}} = 0$$

$$m = n \quad \overline{\hat{A}_{k-n}^2} = 1$$

Proof:

Assume $m < n$, then

$$\begin{aligned}
F_{\hat{A}_{k-m}\hat{A}_{k-n}} &= P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} \leq y\} \\
&= P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} \leq y \mid \hat{A}_{k-m} = I_{k-m}\}p_{k-m} \\
&\quad + P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} \leq y \mid \hat{A}_{k-m} = -I_{k-m}\}q_{k-m} \\
&= P\{\hat{A}_{k-n} \leq x, I_{k-m} \leq y \mid \hat{A}_{k-m} = I_{k-m}\}p_{k-m} \\
&\quad + P\{\hat{A}_{k-n} \leq x, -I_{k-m} \leq y \mid \hat{A}_{k-m} = -I_{k-m}\}q_{k-m}.
\end{aligned}$$

\hat{A}_{k-n} depends only on I_{k-m} for $m \geq n$, hence, it is independent of all I_{k-m} with $m < n$.

Therefore,

$$\begin{aligned}
F_{\hat{A}_{k-m}\hat{A}_{k-n}} &= P\{\hat{A}_{k-n} \leq x \mid \hat{A}_{k-m} = I_{k-m}\}P\{\hat{I}_{k-m-1} \leq y \mid \hat{A}_{k-m} = I_{k-m}\}p_{k-m} \\
&\quad + P\{\hat{A}_{k-n} \leq x, \mid \hat{A}_{k-m} = -I_{k-m}\}P\{-\hat{I}_{k-m} \leq y, \mid \hat{A}_{k-m} = -I_{k-m}\}q_{k-m} \\
&= P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} = I_{k-m}\}P\{\hat{I}_{k-m} \leq y \mid \hat{A}_{k-m} = I_{k-m}\} \\
&\quad + P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} = -I_{k-m}\}P\{-\hat{I}_{k-m} \leq y, \mid \hat{A}_{k-m} = -I_{k-m}\}.
\end{aligned}$$

By using eqs. (A.9) and (A.10), we get

$$\begin{aligned}
F_{\hat{A}_{k-m}\hat{A}_{k-n}} &= P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} = I_{k-m}\}F_{I_{k-m}}(y) \\
&\quad + P\{\hat{A}_{k-n} \leq x, \hat{A}_{k-m} = -I_{k-m}\}F_{I_{k-m}}(y) \\
&= F_{\hat{A}_{k-n}}(x)F_{I_{k-m}}(y).
\end{aligned} \tag{A.15}$$

Therefore,

$$\begin{aligned}
\overline{\hat{A}_{k-m}\hat{A}_{k-n}} &= \overline{\hat{A}_{k-n}I_{k-m-1}} \\
&= 0.
\end{aligned} \tag{A.16}$$

For $m > n$, a similar proof can be shown by conditioning on \hat{A}_{k-n} instead.

For $m = n$, since the pdf of A_{k-n} is even, it follows that

$$P\{\hat{A}_{k-n} = 1\} = P\{\hat{A}_{k-n} = -1\} = \frac{1}{2}.$$

From the above, it is straightforward to show that, for $m = n$,

$$\overline{\hat{A}_{k-n}^2} = .1$$

Claim 6

For $m + 1 < i < N$

$$\begin{aligned}\overline{A_{k-m}e_{k-i}} &= \overline{A_{k-m}(I_{k-i} - \hat{A}_{k-i})} \\ &= 2q_{k-i}(h_{i-m} + w_{i-m}),\end{aligned}$$

where q_{k-i} is the probability that $\hat{A}_{k-i} \neq I_{k-i}$.

Proof:

Note that

$$\overline{A_{k-m}e_{k-i}} = \overline{A_{k-m}I_{k-i}} - \overline{A_{k-m}\hat{A}_{k-i}}.$$

From eq. (1),

$$A_{k-m} = I_{k-m} + \sum_{j=1}^N h_j I_{k-m-j} - \sum_{j=1}^N w_j \hat{A}_{k-m-j}.$$

We consider each term separately:

$$\overline{A_{k-m}I_{k-i}} = \overline{I_{k-m}I_{k-i}} + \sum_{j=1}^N h_j \overline{I_{k-m-j}I_{k-i}} - \sum_{j=1}^N w_j \overline{\hat{A}_{k-m-j}I_{k-i}}.$$

The first term in the RHS is zero, since $i > m$. Similarly, the summation in the second term is $j = i - m$. Using the result in Claim 3 the terms in the second summation are all zero except for $j = i - m$. Therefore,

$$\overline{A_{k-m}I_{k-i}} = h_{i-m} - w_{i-m} \overline{\hat{A}_{k-i}I_{k-i}}. \quad (\text{A.17})$$

Now, consider

$$\overline{A_{k-m}\hat{A}_{k-i}} = \overline{I_{k-m}\hat{A}_{k-i}} + \sum_{j=1}^N h_j \overline{I_{k-m-j}\hat{A}_{k-i}} - \sum_{j=1}^N w_j \overline{\hat{A}_{k-m-j}\hat{A}_{k-i}}.$$

Using Claim 3, the first term is zero. Furthermore, using the same claim, the only non-zero term in the first summation is $j = i - m$. On the other hand, using Claim 5 the only non-zero term in the second sum is $j = i - m$.

$$\begin{aligned}\overline{A_{k-m}I_{k-i}} &= h_{i-m}\overline{\hat{A}_{k-i}I_{k-i}} - w_{i-m}\overline{\hat{A}_{k-i}^2} \\ &= h_{i-m}\overline{\hat{A}_{k-i}I_{k-i}} - w_{i-m}.\end{aligned}\tag{A.18}$$

Combining eqs. (A.17) and (A.18), we get

$$\begin{aligned}\overline{A_{k-m}e_{k-i}} &= (h_{i-m} + w_{i-m})\left(1 - \overline{\hat{A}_{k-i}I_{k-i}}\right) \\ &= 2q_{k-i}(h_{i-m} + w_{i-m})\end{aligned}\tag{A.19}$$

Claim 7

$$\begin{aligned}& \mathbf{P}\{(-1 + (h_1 + w_1^{(k)})I_{k-1} + (h_2 + w_2^{(k)})I_{k-2}) > 0\} \\ &= \frac{1}{2} \left(\mathbf{P}\{I_{k-2} > \frac{1 - (h_2 + w_2^{(k)})}{|h_1 + w_1^{(k)}|}\} + \mathbf{P}\{I_{k-1} > \frac{1 + (h_2 + w_2^{(k)})}{|h_1 + w_1^{(k)}|}\} \right)\end{aligned}\tag{A.20}$$

Proof:

Define \mathcal{P} as

$$\mathcal{P} \triangleq \mathbf{P}\{(-1 + (h_1 + w_1^{(k)})I_{k-1} + (h_2 + w_2^{(k)})I_{k-2}) > 0\}.$$

Then

$$\begin{aligned}\mathcal{P} &= \frac{1}{2} \left(\mathbf{P}\{(-1 + (h_1 + w_1^{(k)})I_{k-1} + (h_2 + w_2^{(k)})I_{k-2}) > 0\} \right. \\ &\quad \left. + \mathbf{P}\{(-1 + (h_1 + w_1^{(k)})I_{k-1} - (h_2 + w_2^{(k)})I_{k-2}) > 0\} \right).\end{aligned}$$

If $(h_1 + w_1^{(k)}) > 0$, then

$$\mathcal{P} = \frac{1}{2} \left(\mathbf{P}\{I_{k-1} > \frac{1 - (h_2 + w_2^{(k)})}{h_1 + w_1^{(k)}}\} + \mathbf{P}\{I_{k-1} > \frac{1 - (h_2 + w_2^{(k)})}{h_1 + w_1^{(k)}}\} \right).$$

If, on the other hand, $(h_1 + w_1^{(k)}) < 0$, then

$$\begin{aligned} \mathcal{P} &= \frac{1}{2} \left(P\left\{I_{k-1} < \frac{1 - (h_2 + w_2^{(k)})}{h_1 + w_1^{(k)}}\right\} + P\left\{I_{k-1} < \frac{1 - (h_2 + w_2^{(k)})}{h_1 + w_1^{(k)}}\right\} \right) \\ &= \frac{1}{2} \left(P\left\{I_{k-1} > -\frac{1 - (h_2 + w_2^{(k)})}{h_1 + w_1^{(k)}}\right\} + P\left\{I_{k-1} > -\frac{1 - (h_2 + w_2^{(k)})}{h_1 + w_1^{(k)}}\right\} \right). \end{aligned}$$

The last step follows since the pdf of I_{k-1} is an even function. Therefore, combining the above

$$\begin{aligned} &P\left\{(-1 + (h_1 + w_1^{(k)})I_{k-1} + (h_2 + w_2^{(k)})I_{k-2}) > 0\right\} \\ &= \frac{1}{2} \left(P\left\{I_{k-1} > \frac{1 - (h_2 + w_2^{(k)})}{|h_1 + w_1^{(k)}|}\right\} + P\left\{I_{k-1} > \frac{1 + (h_2 + w_2^{(k)})}{|h_1 + w_1^{(k)}|}\right\} \right). \end{aligned}$$

References

- [1] Godard, D. N., "Self-recovering equalization and carrier recovery in two dimensional data communication systems," *IEEE Trans. on Comm.*, 1980, COM-28, pp. 1867-1875.
- [2] Sato, Y., "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. on Comm.*, 1975, COM-23, pp. 679-682.
- [3] Johnson, C. R., "Admissibility in blind adaptive channel equalization," *IEEE Control Systems Magazine*, 1991, pp. 3-15 .
- [4] Lucky, R. W., "Automatic equalization for digital communications," *Bell System Technical Journal*, 1965, 44 pp. 547-588.
- [5] Qureshi, S. U. H., "Adaptive Equalization," *Proc. of IEEE*, 1984. 73(9), pp. 1349-1387.

APPENDIX I
ADAPTIVE TWO-STAGE DETECTION SCHEME
IN SYNCHRONOUS TWO-USER
CDMA SYSTEMS¹

by

Zoran Siveski and Yeheskel Bar-Ness

ABSTRACT

An adaptive two-stage scheme for a synchronous, two-user CDMA environment with unknown received signal energies is presented. It consists of a tandem of the matched filter front-end followed by the interference canceler whose weights are adjusted by an adaptive algorithm. The error probability was evaluated analytically, and it was shown that the receiver provides satisfactory performance in the near-far scenarios.

¹This work will be presented at Milcom '93.

This work was partially supported by a grant from Rome Air Force Lab. (AFSC, Griffiss Air Force Base, NY under contract F30602-88-D-0025, Task C-2-2404.

I. INTRODUCTION

A conventional single-user detector implemented in the Code Division Multiple Access (CDMA) system consists of a bank of filters, each one matched to the signature sequence of the particular user. The sampled output of each matched filter, besides the desired signal, contains the residual interference from all other users. Careful choice of the signature sequences can reduce the amount of interference if the received signal energies are similar. However, the presence of a strong interference often makes it impossible to detect the weak user, a condition referred to as a near-far problem. In [1] a receiver that is optimum in the multiuser interference environment was proposed and shown to provide much improved performance. The improvement comes at the expense of high computational complexity. A class of suboptimum receivers that uses decorrelating detectors and which is based on the linear transformation of the sampled matched filters' outputs was considered in [2] and [3]. Another approach for suboptimum multi-user detectors with low complexity was proposed in [4] and [5], where, in order to perform detection of the desired user, tentative decisions on information bits of all other users are made. The estimate of the multiple access interference is then obtained and is subtracted from the desired signal. The performance of some of these is close to the performance of the optimum detector, particularly when the power of the interferers increases, they become indistinguishable. These schemes, however, have to perform an estimation of the received signal energies, knowledge of which is required for the detectors' proper operation.

In this paper a two-stage detector similar to one proposed in [5] is considered, except that the received signal energies are not assumed to be known, and therefore the canceler's weights are adjusted adaptively. A simple iterative algorithm is proposed in order to control the weights in the second stage of the receiver. The output error probability is computed and compared to the one of the conventional receiver, the decorrelating detector, and the detector in [5].

II. TWO-STAGE ADAPTIVE DETECTOR

In this paper we restrict ourselves to a two-user synchronous CDMA receiver shown in Fig. 1. It consists of two stages: a matched filter front-end followed by a dual interference canceler with the forward-forward structure. The received signal $x(t)$ can be written as:

$$x(t) = \sum_i b_1(i) \sqrt{A_1} s_1(t - iT) + b_2(i) \sqrt{A_2} s_2(t - iT) + n(t) \quad (1)$$

where $b_1, b_2 \in \{-1, +1\}$ are users' data bits, and A_1 and A_2 their respective energies, unknown to the receiver. The signature sequences $s_1(t)$ and $s_2(t)$ are known to the receiver, each having same duration T as a data bit. Denoting the normalized crosscorrelation parameter of the sequences as ρ we have:

$$\int_0^T s_j^2(t) dt = 1, \quad j = 1, 2 \quad \text{and} \\ \int_0^T s_1(t) s_2(t) dt = \rho < 1 \quad (2)$$

The additive noise $n(t)$ is Gaussian, with a zero mean and a power spectral density of $N_0/2$.

In the i th bit interval the sampled outputs of the first stage are:

$$x_1(i) = \sqrt{A_1} b_1(i) + \rho \sqrt{A_2} b_2(i) + n_1(i) \\ x_2(i) = \sqrt{A_2} b_2(i) + \rho \sqrt{A_1} b_1(i) + n_2(i) \quad (3)$$

where $n_1(i)$ and $n_2(i)$ are zero mean Gaussian random variables which can easily be shown to have variances of $N_0/2$, and the crosscorrelation $\rho N_0/2$. (For the sake of brevity, time index i is omitted from most of the expressions in the text.)

The first stage bit estimates are defined as:

$$\hat{b}_1 = \text{sgn}[x_1] \quad \text{and} \quad \hat{b}_2 = \text{sgn}[x_2] \quad (4)$$

yielding the two outputs of the second stage:

$$y_1 = \sqrt{A_1} b_1 + \rho \sqrt{A_2} b_2 - w_1 \hat{b}_2 + n_1$$

$$y_2 = \sqrt{A_2}b_2 + \rho\sqrt{A_1}b_1 - w_2\hat{b}_1 + n_2 \quad (5)$$

The control algorithm simultaneously minimizes the output powers $E\{y_1^2\}$ and $E\{y_2^2\}$. The optimum weights are obtained by using an iterative search:

$$w_j \leftarrow w_j - \mu \frac{dE\{y_j^2\}}{dw_j} \quad j = 1, 2 \quad (6)$$

where μ represents the convergence and stability rate constant. The optimum weights that minimize the output powers are the steady state values obtained from:

$$\begin{aligned} \frac{dE\{y_1^2\}}{dw_1} &= \frac{d}{dw_1} E\{(x_1 - w_1\hat{b}_2)^2\} \\ &= E\{(x_1 - w_1\hat{b}_2)(-\hat{b}_2)\} = 0 \end{aligned}$$

and

$$\begin{aligned} \frac{dE\{y_2^2\}}{dw_2} &= \frac{d}{dw_2} E\{(x_2 - w_2\hat{b}_1)^2\} \\ &= E\{(x_2 - w_2\hat{b}_1)(-\hat{b}_1)\} = 0 \end{aligned} \quad (7)$$

and are expressed as:

$$\begin{aligned} w_{10} &= \sqrt{A_1}E\{b_1\hat{b}_2\} + \rho\sqrt{A_2}E\{b_2\hat{b}_2\} + E\{n_1\hat{b}_2\} \\ w_{20} &= \sqrt{A_2}E\{b_2\hat{b}_1\} + \rho\sqrt{A_1}E\{b_1\hat{b}_1\} + E\{n_2\hat{b}_1\} \end{aligned} \quad (8)$$

The joint statistics appearing in the above expressions are evaluated in terms of the system's parameters and are presented in the Appendix. After substituting the optimum weights the canceler outputs become:

$$\begin{aligned} y_{10} &= \sqrt{A_1} [b_1 - E\{b_1\hat{b}_2\}\hat{b}_2] \\ &+ \rho\sqrt{A_2} [b_2 - E\{b_2\hat{b}_2\}\hat{b}_2] + E\{n_1\hat{b}_2\} + n_1 \\ y_{20} &= \sqrt{A_2} [b_2 - E\{b_2\hat{b}_1\}\hat{b}_1] \\ &+ \rho\sqrt{A_1} [b_1 - E\{b_1\hat{b}_1\}\hat{b}_1] + E\{n_2\hat{b}_1\} + n_2 \end{aligned} \quad (9)$$

From the above expressions one can conclude that if the bit estimates of the sampled correlators' outputs were almost perfect (i.e., $\hat{b}_1 \approx b_1$ and $\hat{b}_2 \approx b_2$), the decision variables at the canceler output become interference free (i.e., $y_{10} \approx \sqrt{A_1}b_1 + n_1$ and $y_{20} \approx \sqrt{A_2}b_1 + n_2$). The error performance of each user will be determined by their respective input signal-to-noise ratios, $SNR_1 = A_1/N_0$ and $SNR_2 = A_2/N_0$ only. This scenario requires a system with a rather low spectral efficiency (small ρ), and occurs whenever both $SNRs$ are large enough and are approximately equal.

If the power of one of the signals is large enough in comparison to the other such that, say $\rho^2 SNR_2 \gg SNR_1$, \hat{b}_2 will continue to be an almost perfect estimate of b_2 . Referring to (8), $w_{10} \approx \rho\sqrt{A_2}$, and the output 1 will again be interference free. However, the signal x_1 will now be dominated by the interference from user 2, and \hat{b}_1 will mostly be estimated as b_2 . By observing (8) again, one can see that it will cause $w_{20} \approx \sqrt{A_2}$, which is obviously larger than w_{10} . This results in the total cancelation of the desired signal 2 in the second output, which is expected because of the power inversion effect of the canceler. The decision variable of user 2 becomes $y_2 \approx \rho\sqrt{A_1}b_1 + n_2$, yielding to the disastrous output performance.

An easy and logical remedy to such a problem may be obtained by adding a constraint to the iterative algorithm, such that, at any bit interval,

$$w_j(i) = \min\{w_1(i), w_2(i)\} \quad j = 1, 2 \quad (10)$$

The restriction effectively prevents an increase of the weight that affects the signal with the very large input SNR ; the one that does not need the interference canceler in the first place. The effect of the constraint can be observed if, under the same assumptions described above, the larger of the weights, w_{02} , is replaced by the smaller one, $w_{01} \approx \rho\sqrt{A_2}$, resulting in the second output, $y_2 \approx (1 - \rho)\sqrt{A_2}b_2 + \rho\sqrt{A_1}b_1 + n_2$. This certainly will not be worse than the output that results when no constraint is imposed. The amount of improvement depends on the desired component to the residual interference ratio at the output, $[(1 - \rho)/\rho]^2 \Delta SNR$.

Another possible constraint considered was in "disabling" the canceling loop that contains

the larger weight element, such that

$$\max\{w_1(i), w_2(i)\} = 0 \quad (11)$$

In the above scenario, it means $w_2 = 0$, which corresponds to detecting the data bits of user 2, the one with the larger input SNR , directly from the correlator output.

III. ERROR PERFORMANCE

The error probability at the canceler's output is evaluated as follows. The bit estimates at the output are defined as:

$$\hat{b}_1 = \text{sgn}[y_1] \quad \text{and} \quad \hat{b}_2 = \text{sgn}[y_2] \quad (12)$$

The two-user output error probabilities $P_{e1}(\epsilon)$ and $P_{e2}(\epsilon)$ are the conditional error probabilities averaged over b_1, b_2 , and \hat{b}_2 , and over b_1, b_2 , and \hat{b}_1 , respectively. For user 1:

$$\begin{aligned} P_{e1} &= E_{b_1, b_2, \hat{b}_2} Pr\{\hat{b}_1 \text{ in error} | b_1, b_2, \hat{b}_2\} \\ &= \frac{1}{2} \sum_{b_2, \hat{b}_2} [Pr(n_1 > \sqrt{A_1} - \rho\sqrt{A_2}b_2 + w_{10}\hat{b}_2) \\ &\quad + Pr(n_1 < -\sqrt{A_1} - \rho\sqrt{A_2}b_2 + w_{10}\hat{b}_2)] \\ &= \frac{1}{2} \left[Pr(n_1 > \sqrt{A_1} + \rho\sqrt{A_2} - w_{10}, n_2 < \sqrt{A_2} + \rho\sqrt{A_1}) \right. \\ &\quad + Pr(n_1 > \sqrt{A_1} - \rho\sqrt{A_2} - w_{10}, n_2 < -\sqrt{A_2} + \rho\sqrt{A_1}) \\ &\quad + Pr(n_1 > \sqrt{A_1} + \rho\sqrt{A_2} + w_{10}, n_2 > \sqrt{A_2} + \rho\sqrt{A_1}) \\ &\quad \left. + Pr(n_1 > \sqrt{A_1} - \rho\sqrt{A_2} + w_{10}, n_2 > -\sqrt{A_2} + \rho\sqrt{A_1}) \right] \quad (13) \end{aligned}$$

Similarly, the error probability P_{e2} for user 2 is:

$$\begin{aligned} &\frac{1}{2} \left[Pr(n_2 > \sqrt{A_2} + \rho\sqrt{A_1} - w_{20}, n_1 < \sqrt{A_1} + \rho\sqrt{A_2}) \right. \\ &\quad \left. + Pr(n_2 > \sqrt{A_2} - \rho\sqrt{A_1} - w_{20}, n_1 < -\sqrt{A_1} + \rho\sqrt{A_2}) \right] \end{aligned}$$

$$\begin{aligned}
& + Pr(n_2 > \sqrt{A_2} + \rho\sqrt{A_1} + w_{20}, n_1 > \sqrt{A_1} + \rho\sqrt{A_2}) \\
& + Pr(n_2 > \sqrt{A_2} - \rho\sqrt{A_1} + w_{20}, n_1 > -\sqrt{A_1} + \rho\sqrt{A_2}) \Big] \quad (14)
\end{aligned}$$

When the constraint from either (10) or (11) is taken into account, the weights w_{10} and w_{20} in the above expressions should be modified accordingly.

Numerical results from the computation of the error probabilities are presented as functions of signal to noise ratios and the crosscorrelation coefficient. The error performance of the conventional receiver is also presented in order to illustrate its vulnerability to the near-far problem. In addition, the error curves for the decorrelating detector and the system presented in [5], where the exact knowledge of the received signal energies was assumed, are included for comparison.

The error probability curves are plotted versus the difference of the two input SNRs, $\Delta SNR = SNR_2 - SNR_1$, with SNR_1 kept constant. Three different crosscorrelation coefficient values, $\rho = 0.7$, $\rho = 0.5$, and $\rho = 1/3$ are considered; the first corresponding to the high bandwidth efficiency system. As mentioned previously, a constraint on the weights is also added to prevent possible cancelation of the larger desired signal component at one of the outputs.

In Fig. 2 we use $\rho = 0.7$ and ΔSNR in the range from -4 dB to 12 dB. As expected, the performance of user 1 is virtually the same regardless of the constraint strategy used. The error performance of user 2, with no weight constraints imposed, is very poor as could have been predicted. Marginal improvement occurs at high values of ΔSNR when the constraint from (10) is imposed. However, by using the constraint from (11) instead, excellent performance for both users is achieved, which is even better than the results in [5].

Fig. 3 is the same except for $\rho = 0.5$. Here we observe that the constraint from (10) provides better performance of user 2 than with the previous value of ρ . Again, as expected, the performance of user 1 is virtually the same as in [5]. This is a result of the factor $[(1 - \rho)/\rho]^2$ being greater for this case.

Fig. 4 depicts the case for $\rho = 1/3$. Without the constraint on the weights, there is a range of the values of ΔSNR for which the system performs very well. Notice the deterioration in performance of user 2 for very high ΔSNR due to the circumstances ($\rho^2 SNR_2 \gg SNR_1$) mentioned in the previous section.

With the constraint from (10) in place, the performance of user 2 matches the one obtained in [5]. There is a slight performance degradation for user 1 over a limited range of $\Delta SNRs$ when the constraint is imposed.

IV. CONCLUSION

A two-stage detector for multiple access systems that does not require knowledge of the received signals' energies was proposed and analyzed. It incorporates an interference canceler whose weights are obtained by an iterative algorithm. It was shown that the receiver has comparable error performance in near-far situations to similar receivers that assume the knowledge of the signal energies.

V. APPENDIX

$$\begin{aligned}
 E\{b_1 \hat{b}_1\} &= E_{b_1, b_2} \{b_1 \operatorname{sgn}[\sqrt{A_1} b_1 + \rho \sqrt{A_2} b_2 + n_1]\} \\
 &= \frac{1}{4} \sum_{b_1, b_2} b_1 \left[\operatorname{Pr}\{n_1 > -\sqrt{A_1} b_1 - \rho \sqrt{A_2} b_2\} \right. \\
 &\quad \left. - \operatorname{Pr}\{n_1 < -\sqrt{A_1} b_1 - \rho \sqrt{A_2} b_2\} \right] \\
 &= 1 - Q\left(\frac{\sqrt{A_1} + \rho \sqrt{A_2}}{\sqrt{N_0/2}}\right) - Q\left(\frac{\sqrt{A_1} - \rho \sqrt{A_2}}{\sqrt{N_0/2}}\right) \\
 E\{b_2 \hat{b}_2\} &= 1 - Q\left(\frac{\sqrt{A_2} + \rho \sqrt{A_1}}{\sqrt{N_0/2}}\right) - Q\left(\frac{\sqrt{A_2} - \rho \sqrt{A_1}}{\sqrt{N_0/2}}\right) \\
 E\{b_1 \hat{b}_2\} &= E_{b_1, b_2} \{b_1 \operatorname{sgn}[\sqrt{A_2} b_2 + \rho \sqrt{A_1} b_1 + n_2]\} \\
 &= \frac{1}{4} \sum_{b_1, b_2} b_1 \left[\operatorname{Pr}\{n_2 > -\sqrt{A_2} b_2 - \rho \sqrt{A_1} b_1\} \right.
 \end{aligned}$$

$$\begin{aligned}
& -Pr\{n_2 < -\sqrt{A_2}b_2 - \rho\sqrt{A_1}b_1\}] \\
& = Q\left(\frac{\sqrt{A_2} - \rho\sqrt{A_1}}{\sqrt{N_0/2}}\right) - Q\left(\frac{\sqrt{A_2} + \rho\sqrt{A_1}}{\sqrt{N_0/2}}\right) \\
E\{b_2\hat{b}_1\} & = Q\left(\frac{\sqrt{A_1} - \rho\sqrt{A_2}}{\sqrt{N_0/2}}\right) - Q\left(\frac{\sqrt{A_1} + \rho\sqrt{A_2}}{\sqrt{N_0/2}}\right) \\
E\{n_1\hat{b}_2\} & = E\{n_1 \operatorname{sgn}(\sqrt{A_2}b_2 + \rho\sqrt{A_1}b_1 + n_2)\} \\
& = \frac{1}{4} \sum_{b_1, b_2} \left[\int_{-\infty}^{\infty} \int_{-\sqrt{A_2}b_2 - \rho\sqrt{A_1}b_1}^{\infty} n_1 f_{n_1, n_2} dn_1 dn_2 \right. \\
& \quad \left. - \int_{-\infty}^{\infty} \int_{-\infty}^{-\sqrt{A_2}b_2 - \rho\sqrt{A_1}b_1} n_1 f_{n_1, n_2} dn_1 dn_2 \right] \\
& = \frac{1}{2} \sqrt{\frac{\rho^2 N_0}{\pi}} \left[e^{-\left(\frac{\sqrt{A_2} + \rho\sqrt{A_1}}{\sqrt{N_0}}\right)^2} + e^{-\left(\frac{\sqrt{A_2} - \rho\sqrt{A_1}}{\sqrt{N_0}}\right)^2} \right] \\
E\{n_2\hat{b}_1\} & = \frac{1}{2} \sqrt{\frac{\rho^2 N_0}{\pi}} \left[e^{-\left(\frac{\sqrt{A_1} + \rho\sqrt{A_2}}{\sqrt{N_0}}\right)^2} + e^{-\left(\frac{\sqrt{A_1} - \rho\sqrt{A_2}}{\sqrt{N_0}}\right)^2} \right]
\end{aligned}$$

where:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$$

and:

$$f_{n_1, n_2} = \frac{1}{\pi N_0 \sqrt{1 - \rho^2}} \exp\left[-\frac{n_1^2 + n_2^2 - 2\rho n_1 n_2}{N_0(1 - \rho^2)}\right]$$

VI. REFERENCES

- [1] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple access channels," *IEEE Trans. Inform. Theory*, Vol. IT-32, No. 1, pp. 85-96, Jan. 1986.
- [2] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code division multiple access channels," *IEEE Trans. Inform. Theory*, vol. IT-35, No. 1, pp. 123-136, Jan 1989.

- [3] R. Lupas and S. Verdu, "Near far resistance of multiuser detectors in asynchronous channels," *IEEE Trans Commun.*, vol. COM-38, No. 4, pp. 496-508, April 1990.
- [4] M. K. Varanasi and B. Aazhang, "Multistage detector in asynchronous code-division multiple access communications," *IEEE Trans. Commun.* vol. 38, pp. 509-519, April 1990.
- [5] M. K. Varanasi and B. Aazhang, "Near-optimum detector in synchronous code-division multiple-access system," *IEEE Trans. Commun.* vol. 39, pp. 725-736, May 1991.

Fig. 1 Two-stage receiver for two synchronous users

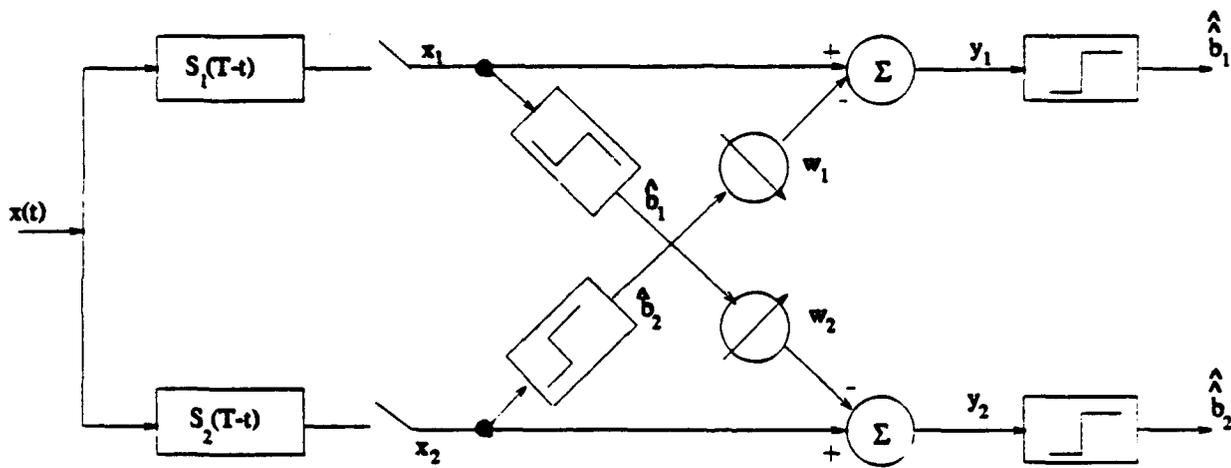


Fig. 2 Error probabilities for users 1 and 2 with $\rho = 0.7$ and $\text{SNR}_1 = 8\text{dB}$

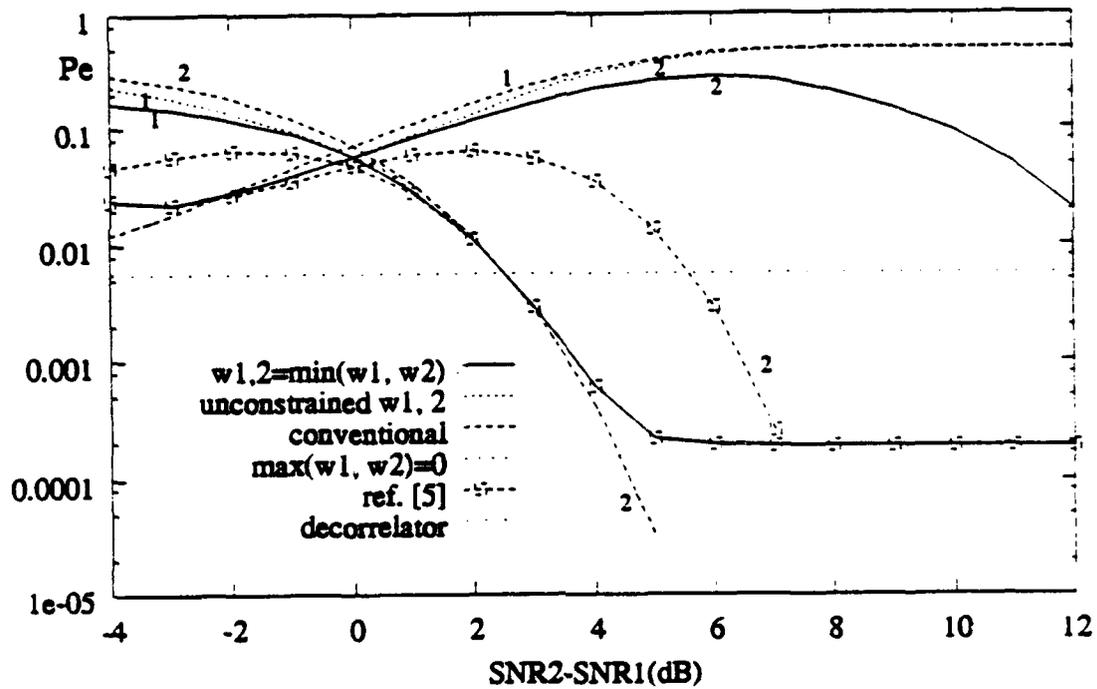


Fig. 3 Error probabilities for users 1 and 2 with $\rho = 0.5$ and $\text{SNR}_1 = 8\text{dB}$

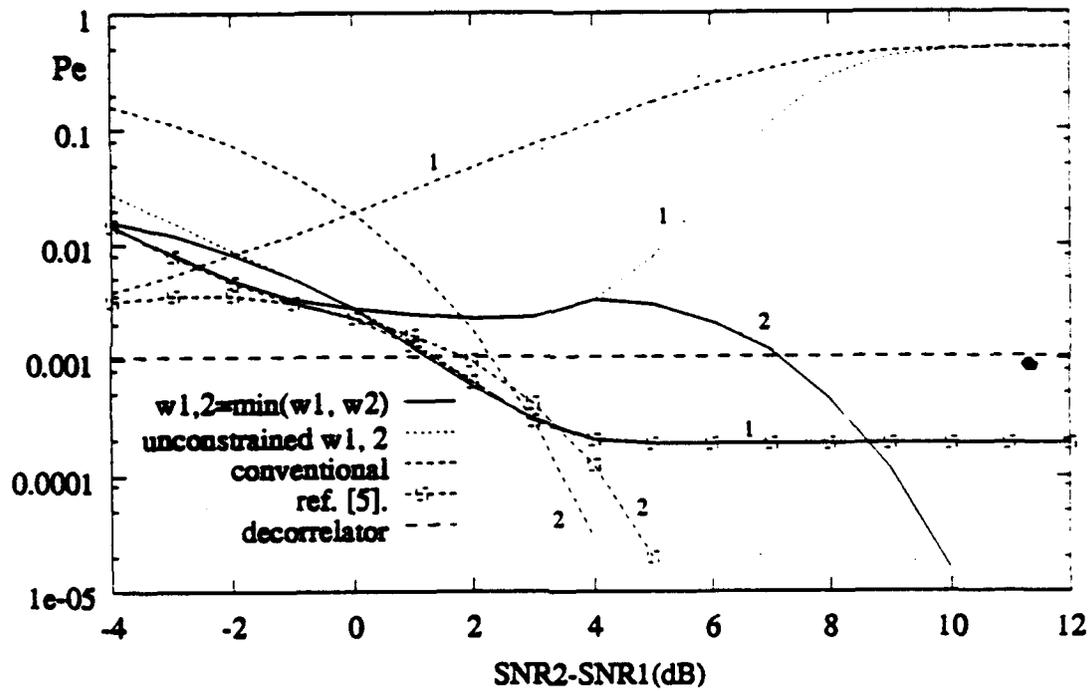


Fig. 4 Error probabilities for users 1 and 2 with $\rho=1/3$ and SNR1 = 8 dB

