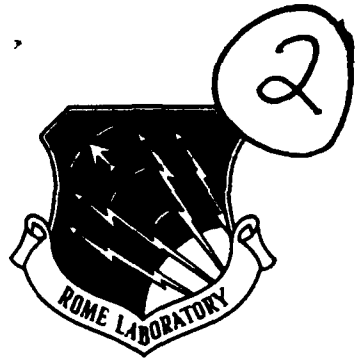


RL-TR-93-24
Final Technical Report
April 1993

THE BOOTSTRAPPED ALGORITHM: FAST ALGORITHM FOR BLIND SIGNAL SEPARATION

New Jersey Institute of Technology

Yeheskel Bar-Ness, Dr. A. Dinc, and Dr. H. Messer-Yaron

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

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

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-93-24 has been reviewed and is approved for publication.

APPROVED:

RICHARD N. SMITH
Project Engineer

FOR THE COMMANDER

JOHN A. GRANIERO
Chief Scientist for C3

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

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1993	Final Mar 91 - Jun 92

4. TITLE AND SUBTITLE
THE BOOTSTRAPPED ALGORITHM: FAST ALGORITHM FOR BLIND SIGNAL SEPARATION

5. FUNDING NUMBERS
C - F30602-88-D-0025, Task C-1-2456
PE - 62702F
PR - 4519
TA - 42
WU - P7

6. AUTHOR(S)
Yeheskel Bar-Ness, Dr. A. Dinc, and Dr. H. Messer-Yaron

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
New Jersey Institute of Technology
Dept of Electrical & Computer Engineering
University Heights
Newark NJ 07102

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Laboratory (C3BA)
525 Brooks Road
Griffiss AFB NY 13441-4505

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
RL-TR-93-24

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)

An adaptive algorithm termed "Bootstrapped Algorithm" is proposed and analyzed, and its performance is evaluated in this report. Using this algorithm as an interference canceler results in structures which are unique in that they are composed of multi-cancelers; each uses the outputs of other cancelers as references (desired signals) or in other forms to further improve performance. As a result of such unique structures, the bootstrapped algorithm is shown to perform as a "Signal Separator" rather than as an interference canceler. Clearly, because it does not require a reference signal in the form of decision feedback or training sequence, it is sometimes justifiabley referred to by the term "Blind Separator."

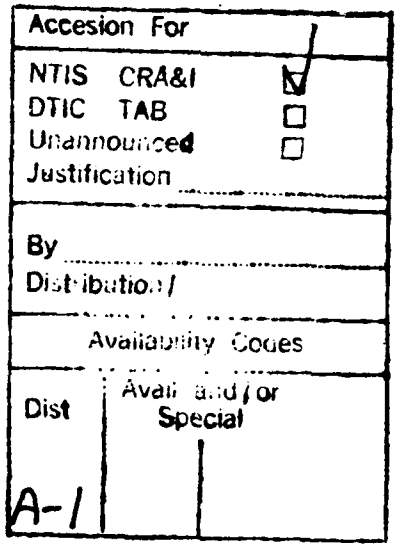
14. SUBJECT TERMS
Algorithm, Signal Processing, Interference Cancellation

15. NUMBER OF PAGES
252

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

TABLE OF CONTENTS



ABSTRACT

An adaptive algorithm termed "Bootstrapped Algorithm" is proposed and analyzed, and its performance is evaluated in this report. Using this algorithm as an interference canceler results in structures which are unique in that they are composed of multi-cancelers; each uses the outputs of other cancelers as references (desired signals) or in other forms to further improve performance. As a result of such unique structures, the bootstrapped algorithm is shown to perform as a "Signal Separator" rather then as an interference canceler. Clearly, because it does not require a reference signal in the form of decision feedback or training sequence, it is sometimes justifiably referred to by the term "Blind Separator."

Due to its many advantages, one might consider many applications for such blind signal separators ranging from neural networks and pattern recognition, direction finding and general signal estimation to co-channel interference cancellation and cross polarization interference suppression. After we present a thorough examination of the steady state analysis of the two-inputs two-outputs structures without noise, we consider the bootstrapped signal separator with noise present as a signal estimator and discuss its properties.

Application to cross-polarization cancellation in M-QAM dual-polarized transmission is thoroughly discussed. Improvement of error probability by using such a canceler is quantified. Because of its importance for neural network applications, extension of the bootstrapped algorithm to multi-inputs multi-outputs is also examined.

Convergence properties of the two-inputs two-outputs and the multi-inputs multi-outputs cases were also studied. Finally some preliminary results of the algorithm for wide-band signal cases were obtained.

It was found that the bootstrapped algorithm has many useful properties which make it an excellent candidate for use as a signal separator or interference canceler when other algorithms have some difficulties. In some cases, it clearly outperforms other algorithms.

1 Introduction

Reported below are the results of a study carried out at the Center for Communications and Signal Processing Research, NJIT, between March 1, 1991 and June 30, 1992. The aim of the research is to propose, analyze and evaluate the performance of a fast algorithm, termed "Bootstrapped Algorithm." The bootstrapped algorithm for interference cancellation was first proposed by the principal investigator in 1981 [1] and later used for cancelling cross polarization in satellite communication [2] and in the Microwave Terrestrial Radio Link [3-5]. Other possible applications to tactical communications are included in [6].

The bootstrapped interference canceler is unique in that it is composed of two separate cancelers each using the output of the other canceler as its reference (desired signal) input. As a result of such structure, it is shown to perform as a "Signal Separator" rather than an interference canceler. Clearly, as such, it does not (like regular LMS) require a reference signal in the form of decision feedback or training sequence. Hence sometimes, the term "Blind Separator" is used.

Three different structures are proposed in Fig. 1; (1) Backward/Backward (BB) (2) Forward/Forward (FF) and (3) Forward/Backward (FB). These are the three possible interconnections among different noise cancelers. The adaptive weights for these cancelers can be controlled by minimizing output powers or minimizing the absolute value of the cross correlation between any two outputs and hence sometimes for the two-input two-output, we used the names "Power-Power" (PP), "Power-Correlator" (PC) and "Correlator-Correlator" (CC) (see Figs. 2, 3 and 4).

To understand the behaviour of this newly proposed algorithm, one must discuss the steady state and show that under accepted conditions it converges to its steady state which represents the desirable signal separation. Effect of noise on the behaviour of the algorithm is also of interest. Looking at the algorithm as an estimator of the desired signal, one might also be interested in examining the properties of this estimator under noise.

Due to many advantages, one might consider many applications for such a blind signal separator, ranging from neural networks and pattern recognition, direction finding and general signal estimation to co-channel interference cancellation and cross polarization interference suppression. In this report, after studying the principle property of the algorithm without and with noise, we study the performance of the Backward/Backward (BB)- Power-Power bootstrapped cross-polarization canceler and compare it to that of the Diagonalizer and LMS cancelers. Performance study of the other two structures. i.e. Forward/Forward (FF) - Correlator-Correlator and Forward/Backward - Correlator-Power is followed. It is also compared with that of the BB/PP canceler.

Clearly, for cross-polarization cancellation of a dual polarized channel, we are interested in separating two orthogonal signals and obtaining at the outputs, signals as clean as possible with low error rates. Hence, we examined the two-input two-output case. Extension to the multidimensional case is also of interest particularly, as for example, in neural network applications. This was first done with the BB structure and then when using the FB and FF structure.

The convergence properties of the bootstrapped algorithm with two-inputs, two-outputs and multi-inputs, multi-outputs are also examined. Finally some preliminary

study of the algorithm for wide-band signal is started.

It was found that the bootstrapped algorithm has many useful properties which make it an excellent candidate for use as a signal separator or interference canceler when other algorithms have some difficulties. In some cases, it clearly outperforms other algorithms.

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

2 Technical Summary

2.1 Bootstrapped Algorithm - Two inputs - Two outputs

2.1.1 Bootstrapped Adaptive Separation of Superimposed Signals - Steady State Analysis

Consider the two inputs (in complex envelope notation)

$$\begin{aligned} v_1(t) &= s_1(t) + bs_2(t) \\ v_2(t) &= cs_1(t) + s_2(t) \end{aligned} \tag{1}$$

where b and c are complex values and $|b|^2$ and $|c|^2$ are the input signal to interference ratios. $s_1(t)$ and $s_2(t)$ are zero mean uncorrelated stationary complex processes. If we process these two inputs by using Widrow's noise canceler,[7] then the two outputs will follow the "power-inversion" relation with respect to power ratios at the inputs to the weighted elements (see Fig. 1 of appendix A). Particularly, when $|b|$ and $|c|$

are less than unity, as in a cross polarization interference case, this canceler is useless. A novel way to obtain a high signal to interference ratio at both output ports is to use a bootstrapping approach. In this approach, two cancellation paths and two summations are used to obtain the system outputs, and an adaptive algorithm is employed to optimize the signal-to-interference power ratio at the two output ports simultaneously.

Three different configurations are possible for a bootstrapped algorithm; these are: (see Fig. 1)

1. The Backward/Backward (BB)
2. The Forward/Forward (FF)
3. The Forward/Backward (FB)

Such interconnections between the two "noise cancelers" help improve the performance of each one and result in high quality signal separation.

In order to control the adaptive weights, one needs optimization criteria. For the two inputs two outputs case, the following are possible,

1. Power-Power
2. Power-Correlation
3. Correlation-Correlation

In the first, the weights are controlled by minimizing the output powers, respectively. With the third, the absolute values of the cross correlation between one output and a modified version of the second output, and versa is minimized. The second criterion is a combination of both. See. for example, Fig. 2 wherein the backward/backward configuration is used and the weights are controlled by minimizing the powers at the output.

To understand the operation of this particular separator, let the power ratio of the two signals at point No.3 be such that $s > n$ (even if only slightly greater). Point No.3, being the input to the weighted element of the β processor, (terminal No.1, is the other input) will result (because of the power-inversion) in $n > s$ at point No.4 and output $v_p(t)$. But point No.4, being the input to the weighted element of the α processor, (terminal No.2 is its other input) will result in s still greater than n at point No.3 and hence at output port $v_q(t)$. This process of bootstrapping will continue resulting in a very high snr at one port and a very high nsr at the other. The snr and nsr, at these ports respectively, will be upper bounded by values depending on the noise, the impurities of the system and control errors. Consequently, the power-power canceler of Fig. 2 acts as a high quality power separator. In fact, ideal separation occurs only in a noise and impurity free case. The operation of the other configurations are discussed in detail in Appendix A.

One can show that for the backward/backward configuration, power or correlation minimization are equivalent and either or both of them can be used to control the two weights. The forward/forward configuration's weights can only be controlled by de-correlation. The feedback arm of the forward/backward configuration can be controlled by either minimizing the corresponding power output or by de-correlation of the output signals. The forward arm of this configuration can be controlled only by de-correlation (see Fig 5 and 6).

Appendix A of this report contains a detailed study of the steady state analysis of the bootstrapped adaptive signal separator. This analysis concentrates on three configuration - optimization structures.

1. Backward/Backward - Power/Power which for simplicity is termed "Power-Power"

2. Forward/Backward - Power/Correlation which is termed "Power/Correlator"

3. Forward/Forward - Correlator/Correlator which is termed Correlator/Correlator

After obtaining the optimum complex weights of the three different bootstrapping configurations, we examine the conditions under which, utilizing an adaptive algorithm, these values are attainable. It is shown that crucial to some of the analytical results is the assumption that the input de-polarization is always much less than one. The optimal signal power outputs, for the different configurations, are derived and the optimal signal-to-interference power ratios at the two different output ports are calculated and compared. Also calculated and compared are the input-output transmission ratios and the cancellation factors at the two separate ports. Finally, some other questions related to the subject of this paper, namely steady state analysis, are raised.

Finally, we make the following remark in relation to performance comparison of the three bootstrapping configurations considered in this report: while the symmetric power-power and correlation-correlation schemes produce the same output signal powers and power ratios, the asymmetric power-correlation scheme has a slightly different signal power and power ratio at its Q port. These differences are only of a second order and might exist only if, due to some system or input impurities (Gaussian noise, quantization error, etc.), the ideal optimal conditions are not reached. The same conclusion is in effect when comparing the input-output transmission or cancellation factors.

Examination of Figures 2 through 4 reveals the fact that the three configurations proposed require different levels of hardware complexity; some need correlators (harder to implement) other require power measurements. The correlation-correlation scheme is expected to be the most complex while the power-power scheme the least. Also, different signal paths (through the system's circuitry) are expected to result in different system delays with the different configurations, and hence different bandwidth limitations. The question of a possible trade-off between complexity and band-

width is raised. This will be addressed in a subsequent report.

Gaussian noise effects on the system performance and the analytical results of this paper should also be considered. In particular, it is important to investigate the limit on cancellation depth and hence on the quality of power separation that this noise might cause.

The above items are problems which are related to the subject of appendix A namely steady state analysis; but they will be reported separately, in appendix B. Finally the whole subject of dynamic analysis must be investigated for the three configurations.

2.1.2 Bootstrapped Adaptive Separation of Superimposed Signals - Effect of Thermal Noise on Performance

The effect of thermal noise on performance is of great importance. The fact that bootstrapped algorithms result in total separation of two superimposed signals when noise is not present needs to be re-examined when thermal noise is added. This problem is dealt with in appendix B. Again, the different configurations mentioned in the previous section are considered. From the study of the general case of multi-inputs, multi-outputs, the analysis in this appendix is then concentrated on a specific two inputs-two outputs case. In doing so, the effect of the thermal noise on separation performance in terms of the output desired-signal to overall noise ratio is presented. It was concluded that unless signals in the control loops are somewhat *discriminated*, the separation problem can *not* be solved. On the other hand, perfect separation is obtained for infinite signal-to-noise ratio (no noise) even if the discrimination is slight.

It is interesting to note that the bootstrapped algorithm, besides being a signal separator, can provide an estimate of the model parameters (i.e the ratio the signals which are superimposed at the input). As such, the question is what kind of estimator are we dealing with? Without noise the optimal weights converge to the true unbias

estimate of the model parameter. Additive thermal noise at the separator inputs causes *bias* in these estimates. This bias is inversely proportional to the input SNR and increases with the signal cross coupling. The bias causes degradation in the separation performance, but this degradation is small if $|\Delta|SNR >> 1$, where $\Delta = 1 - bc$, is the coupling factor (see eqn. (1)). There exist many practical important separation problems for which $|\Delta|SNR >> 1$ (as in the cross polarization example of section IV. of appendix B). Another question of interest when using the bootstrapped algorithm is, what happens if only one signal is present ? This case may be realistic if one of the signals to be separated is temporarily absent. Discussion of this in appendix B concludes that the bootstrapped algorithm is only applicable for separation of signals and not for noise reduction, etc. (as in the LMS algorithm).

2.2 Bootstrapped Algorithm - Application to Cross Polarization Cancellation

2.2.1 "Power-Power" Bootstrapped Cross-Pol Canceler for M-ary QAM Signals - Performance Evaluation and Comparison with LMS and Diagonalizer

M-QAM dual-polarized transmission became an important method for frequency re-use, particularly in microwave radio communication. However, the orthogonally polarized waves suffer degradation due to cross polarization interference. Different canceler structures were proposed to mitigate the effect of cross-polarization. Among theses are the Diagonalizer and the LMS [8]. The bootstrapped algorithm is another possible signal separator for dual polarized signal.

In appendix C, we study the performance of the "power-power" structure of the bootstrapped algorithm discussed in section 2.1 and compared to other cancelers. As dual polarized signals, we use M-ary QAM modulated signals and as a channel we use non-dispersive fading channel where the depth of fading varies. To deal with realistic

conditions, we also add thermal noise to the two inputs of the canceler. Hence, it was necessary to re-examine the parameter of the canceler with digital data (considered random and uncorrelated).

As a performance measure, we used the error probability. For this we obtained the Chernoff upper bound. To examine the tightness of these we derive approximates to the error probabilities using the Gauss Quadrature method. In obtaining our results, we assumed two kinds of output compensation: amplitude only compensation (automatic gain control, AGC) and amplitude together with phase compensation (AGC and output equalizer).

Many results of simulation and computer numerical calculations are included in appendix C. From these we conclude that, as expected, 16 QAM performance is much better than 64 QAM. It is also shown that adding phase compensation to the canceler output adds very little to the performance obtained when only amplitude compensation is included.

From comparing the results obtained with the moment generating method to the corresponding Chernoff bound, we concluded that these bounds are sufficiently tight. Comparing the results when different numbers of moments are used, and the concluded tightness of the Chernoff bound, we infer that approximately 10 moments are sufficient for deriving a good approximation for the average probability of error using the Gauss quadrature rule.

The performance of these three cancelers is compared numerically in Fig. 7 and Fig. 8 for the 16 QAM with r=-10 dB and -5 dB, respectively. Since the diagonalizer is shown to be useless for the 64 QAM, we only compare the LMS and power-power configuration with 64 QAM in Fig. 9 and 10. It is also clear from these figures that the LMS canceler outperforms the diagonalizer.

The LMS canceler outperforms the diagonalizer particularly when the cross coupling is high. Therefore, in the following figures, we compare the difference in perfor-

mance between the power-power and LMS when using 16 QAM and 64 QAM. This is done in Figs. 11, 12 and 13 when r=-15 dB, -10 dB and -5 dB, respectively. Except for a very low signal-to-noise ratio, the power-power outperforms the LMS with 64 QAM, more than with 16 QAM.

2.2.2 Performance Comparison of Three Bootstrapped Cross-Pol Cancelers for M-ary QAM Signals

In appendix D we study the error probability performances of another two bootstrapped structures, namely the "Power-Correlator" and "Correlator-Correlator" when they are used to cancel cross polarization interference of dual channel M-ary QAM signals. As in dealing with the "Power-Power" structure, we use the Chernoff upper bound and Gauss Quadrature method to estimate the error probability as a function of signal-to-noise ratios.

From the results of this study, we conclude that the bootstrapped Power-Power structure outperforms the Correlator-Correlator and Power-Correlator structures, particularly when the cross coupling is high, such as -10 dB. Also, notice that the performance at the two outputs of the Power-Correlator structure is *not the same.* The performance at one output is the same as that of the Power-Power and the other output performance is close to the performance of the Correlator-Correlator structure.

In Fig. 14, we depict a performance comparison of the three bootstrapped structures with that of the LMS cross polarization canceler. From this figure, it is clearly shown that the Power-Power canceler scheme performs better than the others, while LMS performs better than that of the Correlator-Correlator and that of output-2 of the Power-Correlator structure.

2.3 Bootstrapped Algorithm - Multi-Inputs Multi-Outputs Separator

Beside the cross-polarization cancellation which requires only two inputs-two outputs structures and other applications such as in neural networks and signal separation of multi channel CDMA, there is a need for multi-inputs multi-outputs separators. Obviously, one can do with multi LMS cancelers when each delivers one signal and cancels all the others. This approach becomes very complex when the number of inputs and outputs becomes large. beside its inferiority in performance in comparison to the bootstrapped algorithms which is discussed in this section.

2.3.1 Multi-Inputs, Multi-Outputs Separator Using the Feedback/Feedback Bootstrapped Structure

In appendix E, we present the extension of multi-inputs, multi-outputs of the feedback/feedback structure (see Fig. 15). As an optimum criteria for controlling the weights, we used the minimization of correlation between one output and the cubic of the other outputs. Beside the complexity advantage of the bootstrapped structure to the LMS and the fact that the second needs supervisory reference while the first does not, we studied, in this part of the research, the speed of convergence property of the bootstrap in comparison to the LMS algorithm.

It was shown by simulation that the learning process of the bootstrap signal separator, though requiring fewer weights, is faster than the LMS algorithm. This is true especially at a high signal-to-interference (SIR) ratio. The learning process of the bootstrapped algorithm is almost independent of the SIR and converges within a reasonably small number of iterations. That is, the bootstrapped algorithm is almost independent of the eigenvalue spread of the input correlation matrix.

2.3.2 Convergence and Performance Comparison of Three Different Structures of Bootstrapped Blind Adaptive Algorithm for Multi-signal Co-Channel Separation

In appendix F, we present the multi-inputs multi-outputs extension of the other two bootstrapped structures, "Power-Correlator" and "Correlator-Correlator" (see Fig. 16 and 17). Performance of the separators is studied and compared by examining the depth of interference residue at each output and by considering the speed of convergence. Effect of compensation via AGC on the separator performance is also considered. As in the Power-Power structure discussed in the previous section, the optimum criteria for controlling the weights, we used minimization of correlation of the outputs with cubic non-linearities.

For channel inputs we used random bipolar independent sequences. Channel parameters were chosen to present different desired SIR ratios at the channel outputs. These parameters determine the canceler's input correlation matrix and hence the matrix eigenvalue spread. Thermal noise was also added to the output of the channel.

The optimum weights for all three separators were found analytically in the absence of noise. The signal separation process was shown via simulation by the outputs learning curves. These learning curves of the Power-Power separator is compared to the three outputs learning curves of the Power-Correlator separator when AGC is added to these outputs (see Fig 18). It was also shown that the three different bootstrapped separators converge to their steady state almost with the same speed for two or three signals. lowest for Power-Power and highest for Correlator-Correlator. One output of Power-Correlator results in residue similar to that of the Correlator-Correlator. However, adding AGC to the Correlator-Correlator outputs or to one of the outputs of the Power-Correlator which reduces the amount of residue, so that when AGC was added (when it was needed), all separators behaved similarly.

2.4 Bootstrapped Algorithm - Wideband Signal Separator

For superimposed wideband signals we suggest a bootstrap structure whose controlled elements are delays instead of complex weights or digital filters. In principle, the structure is the same as the backward/backward structure discussed in earlier sections (see Fig. 19). It is shown in appendix G that if the source locations are known then the proposed system provides a least square estimate of the source signals. This suggests that the backward/backward structure provides a simple implementation of the least square multisource estimator. Such an estimator, regularly implemented by a maximum likelihood approach, requires a complex set of software algorithms.

If the source locations are unknown, the algorithm suggested in appendix G converges to the least square solution, provided that some prior information about the source signal is available. Hence, we show that the bootstrapped algorithm with time delay control can be used for the separation of wideband superimposed signals.

First, the general wideband multi-source model of the problem is defined. It assumes having N point sources received by M omnidirectional sensors. This model is practical in passive sonar wherein the signals are wideband, noise -like random processes and the unknown source location is to be estimated. In active sonar, the signals are known and we are interested in estimating the source location. However, in wideband communication one is usually interested in the source signals themselves and not in their location.

Working in the frequency domain we first present the expression for the least-square estimate of the vector of the source signal. Using this expression we depict the direct block diagram that implements this estimator. Restricting our analysis to $M = N = 2$ it is shown that even when the directions of the two sources are known one needs to transfer function of the form of $\cos \omega\Delta$ or $1/\sin^2 \omega\Delta$ as well as delay (see Fig. 1 of appendix G). Our bootstrapped approach will remedy this problem and we only need delay elements and summers.

We show that with feedback structure the expression for the least square can easily be implemented (Fig. 19), if we know the location of the sources. If these locations are unknown then we must use an estimate of these locations. The bootstrapped algorithm can be used to estimate these locations. In fact, the controlled delays at the feedback path of the backward/backward structure will give us an estimate of the delay propagation of the signals' waves impinging on the array and hence directly related to these signals' directions of arrival. When the algorithm controlling these delays reaches its steady state, the outputs of the separator will each deliver sufficiently clean signals at only one output.

3 Conclusions and Recommendations

We have seen that the bootstrapped algorithm has properties which make it attractive for many applications of signal separation. Without noise the separation is ideal. Adding noise will degrade the performance however, for signal estimation this degradation is tolerable. Furthermore, the algorithm has the property to converge 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 signal power ratio and hence does not depend on the eigenvalues spread of the input correlation matrix. It is also important to note that the algorithm does not require a supervisory reference signal. Hence in this regard, it is a blind signal separator.

In this research we studied three structures of this algorithm and discussed possible optimization criteria for each. The case of two-inputs, two outputs was studied in detail. particularly as it is applied to the problem of separating two signals contaminated with cross-polarization interference. When handling M-ary QAM signal we estimate error probabilities without a canceler and with different cancelers, one of which was the bootstrap separator. It was shown that the bootstrapped structures always outperform other cancelers.

The bootstrapped structures were also extended to multi-inputs multi-outputs, and shown to have similar properties as the two-inputs two-outputs case. For optimization criteria we use de-correlation of outputs with cubic nonlinearity. As an application of multi-inputs we considered superimposed signal with uncorrelated digital data and for performance we again examined error probabilities.

Some work was also done with wideband signals. When using delay control in the bootstrapped algorithm, we could prove the very interesting result that this algorithm exhibits a simple hardware implementation of Least Squares estimator.

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

1. Two Input-Two Output Bootstrapped Algorithms

* Study performance of the bootstrapped structure for cross-pol cancellation of dispersive channels. Examine error probability and suggest co-pol compensation whenever needed.

* Further study the dependency of the bootstrapped algorithm on the equivalent system eigenvalues spread to facilitate comparison with LMS on one hand and eigen analysis on the other.

* Compare the complexity of this algorithm to those of other algorithms and draw conclusions regarding its implementation with software or hardware, digital or analog.

* Although bootstrapped algorithms do not require supervisory inputs, it is important to examine what improvement in performance one can get if such input signals are used.

* Study dynamic performance of the three structures and examine possible implementation in analog hardware.

2. Multi-Input Multi-Output Bootstrapped Algorithm

   * Further effort is needed to study error probabilities of multi-signal interference.

   * Analysis of these cases should be completed and conclusions regarding their performance, complexity and implementation should be drawn.

   * Applications of multi-signal co-channel cancellation should be sought.

   * Examine possible application for neural network implementation.

   * Study speed of convergence and depth of cancellation and their dependency on the number of signals being processed.

3. Wideband Application of Bootstrapped Algorithms

   * Further study the bootstrapped time delay estimation of two uncorrelated sources.

   * Extend this work to multi-input multi-output.

   * Examine implementation of such estimators for adaptive wideband interference canceler and for the signal separator.

   * Establish the requirement for stable robust control loops.

   * Examine the performance of these estimators and cancelers under different environmental conditions and with different wideband signal modulation.

   * Compare performance of these estimators to those of maximum likelihood and eigenanalysis estimators.

   * The above would be done using analysis and supported by the needed simulations.

4. The Bootstrapped Algorithm as a Blind Adaptive Equalizer

Blind equalization is a channel equalizer where the reference signal is obtained directly from the channel output instead of by using a pilot, training sequence, or decision feedback. Such an approach has the drawback of being non-adaptive and lacks the capability of following environmental variation. We suggest using the desired signal extracted from one bootstrapped algorithm as a supervisory signal for the blind equalizer. It is believed that such an approach will reduce many problems that current blind equalizers have.

References

[1] Bar-Ness, Y. and Rokach J.," Cross-Couple Bootstrapped Interference Canceler," AP-S Int. Sympos., Conference proceedings pp. 292-295, Los Angeles, CA. June 1981.

[2] Bar-Ness, Y., Carlin, J.W. and Steinberger, M.L. "Bootstrapping Adaptive Cross-Pol Cancelers for Satellite Communication," The Int. Conf. on comm., Conference proceedings paper no. 4F.5, Philadelphia, PA, June 1982.

[3] Carlin, J.W., Bar-Ness, Y., Gross, S., Steinberger, M.S. and Studdiford, W.E., "An IF Cross-Pol Canceler for Microwave Radio," Journal on Selected Area in Communication-Aduances in Digital Comm. by Radio, Vol. SAC-5, No. 3, pp. 502-514, April 1987.

[4] Bar-Ness, Y., Carlin, J.W., " Cross-Pol Canceler Architecture for Microwave Radio Applications." Int. Conf. on comm., Conference proceedings paper No. 52-5, Seattle, WA, June 1987.

[5] Bar-Ness, Y. "Effect of Number of Taps on Cross-Pol Canceler Performance for Digital Radio Systems," Globecom 87, Conference proceedings paper No. 31.7, Tokyo, Japan, Nov. 1987.

[6] Bar-Ness, Y. "Bootstrapped Algorithm for Interference Cancelation," AFCEA-IEEE Technical Conference on tactical Comm., Conference proceedings pp. 111-118, Fort Wayne, Indiana, May 1988.

[7] B.Widrow et al. "Adaptive Noise Cancelling: Principles and Applica tion," Proc. IEEE, Vol.63, No.12, December 1975, pp. 1692-1716.

[8] Kavehrad, M. "Performance of Cross-Polarized M-ary QAM Signals Over Non-dispersive Fading Channels," *AT&T Bell Lab. Tech. J.*, Vol. 63, pp. 499-521, March 1984.

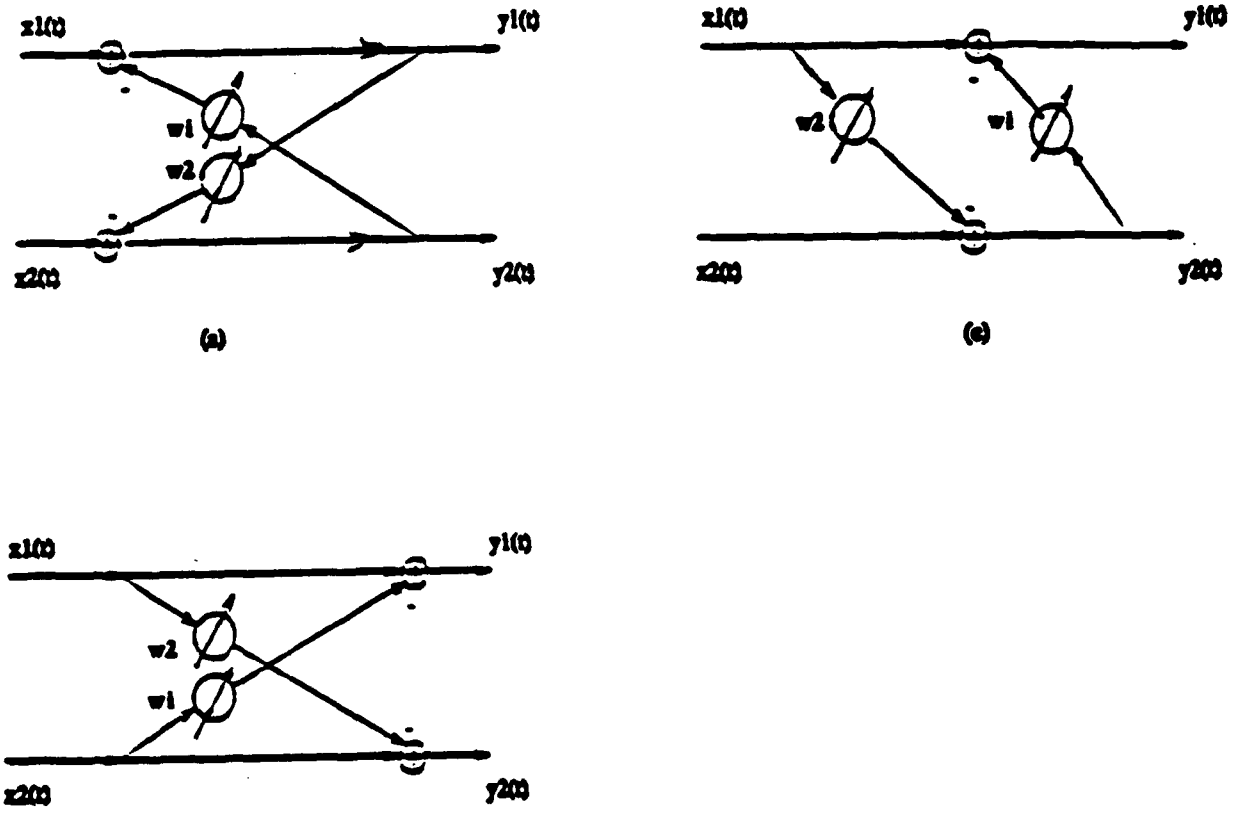


Fig.1 : The three configuration of the bootstrapped signal separator

a. Backward/Backward Structure
b. Forward/Forward Structure
c. Forward/Backward Structure

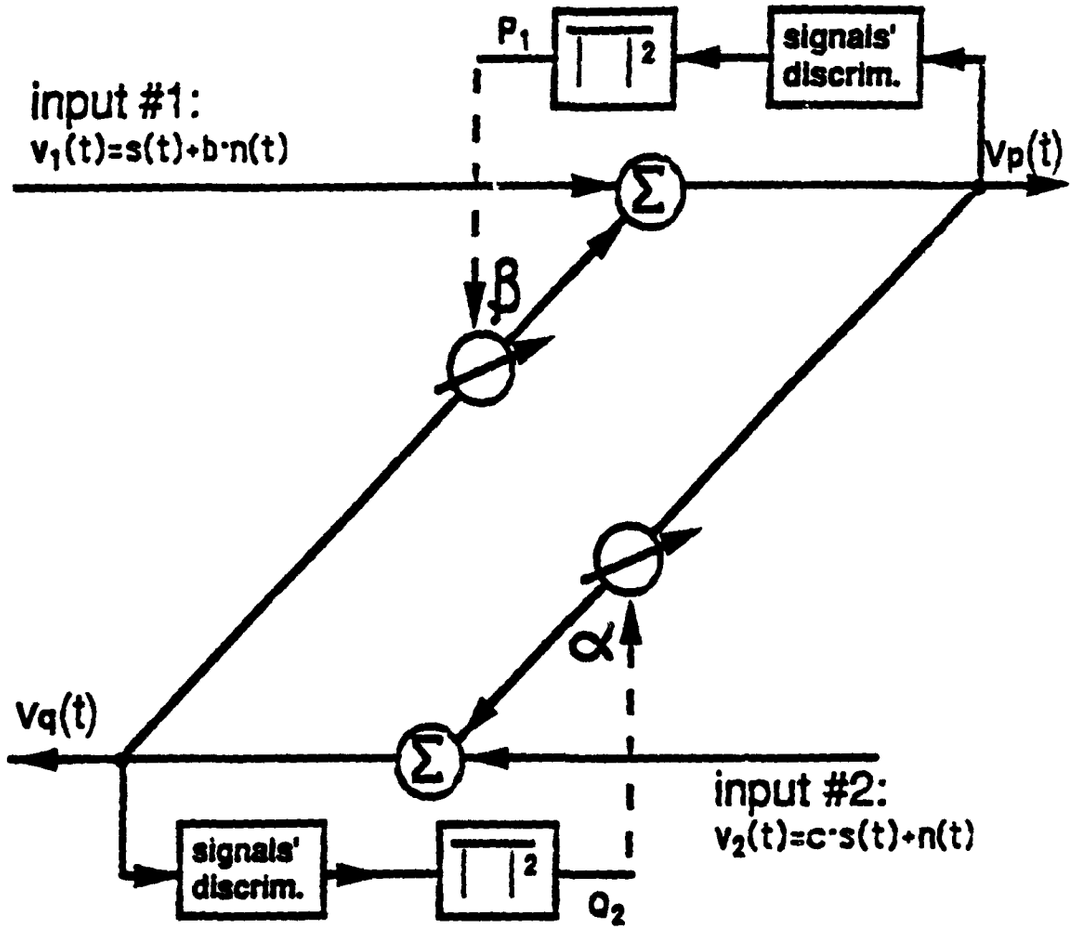


Fig. 2 Power-Power bootstrapped canceler

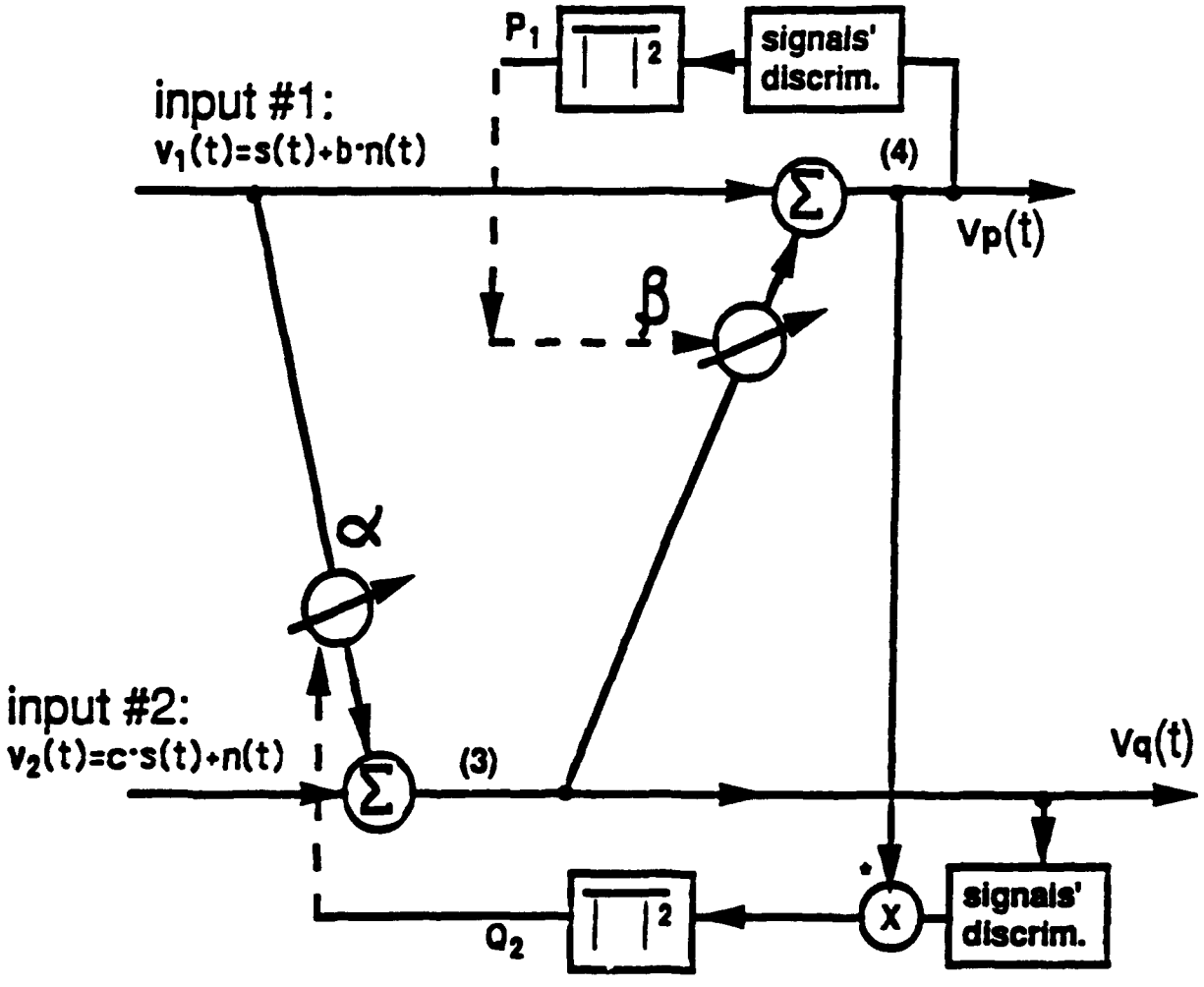


Fig. 3 Power-Correlation bootstrapped canceler

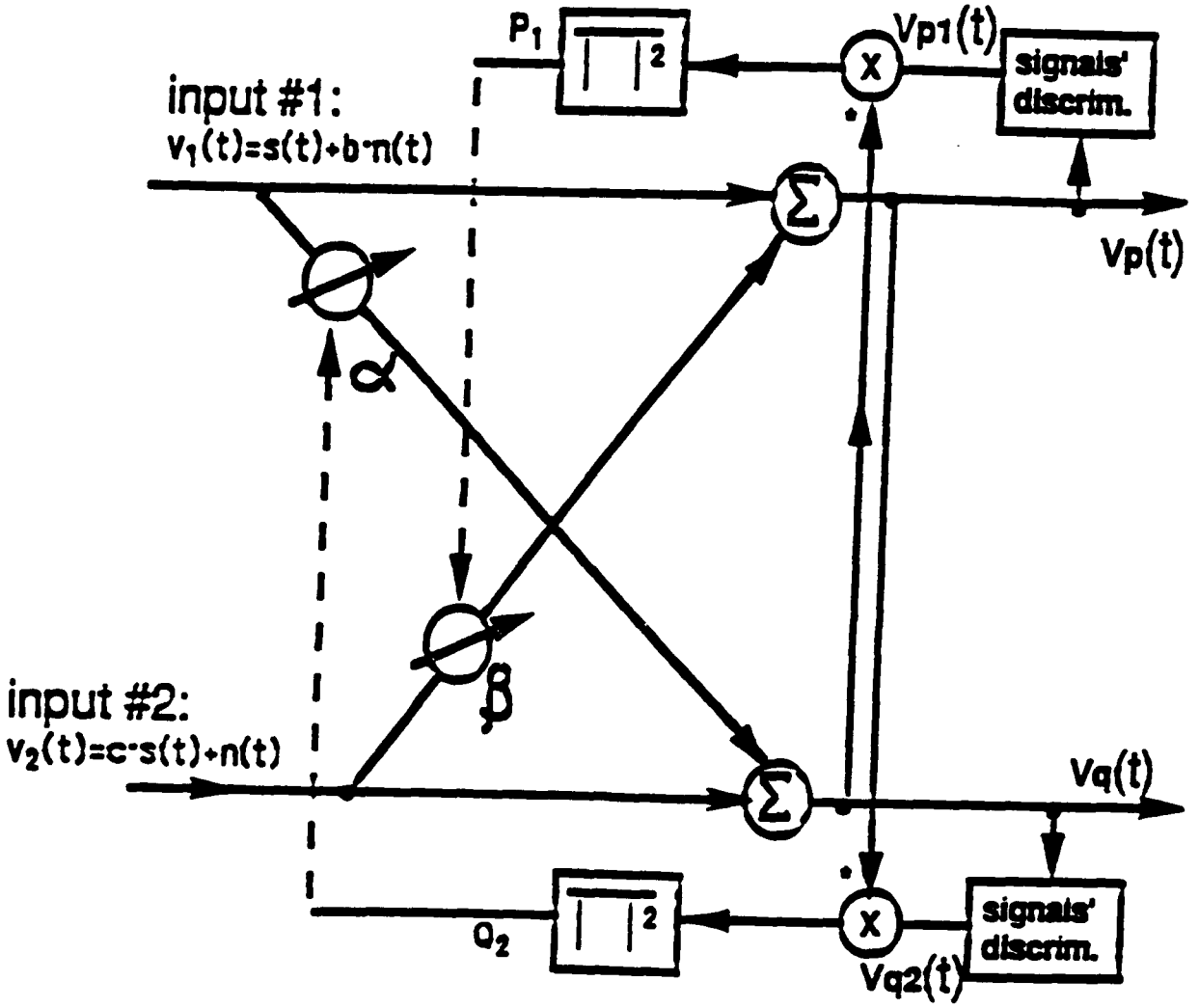


Fig. 4 Correlation-cororrelation bootstrapped canceler

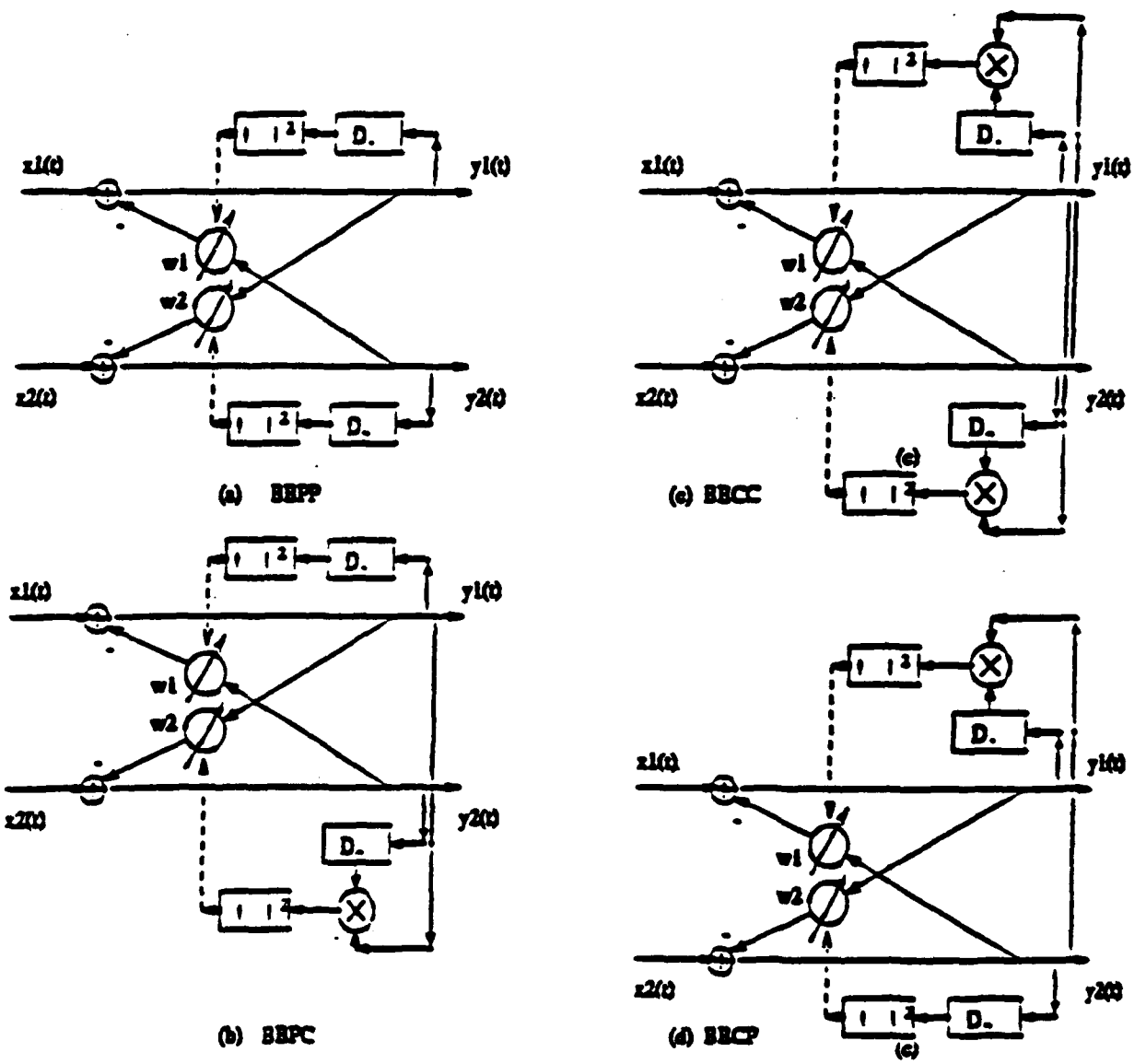


Fig. 5 Different Optimization Criteria for the Backward/Backward Structure

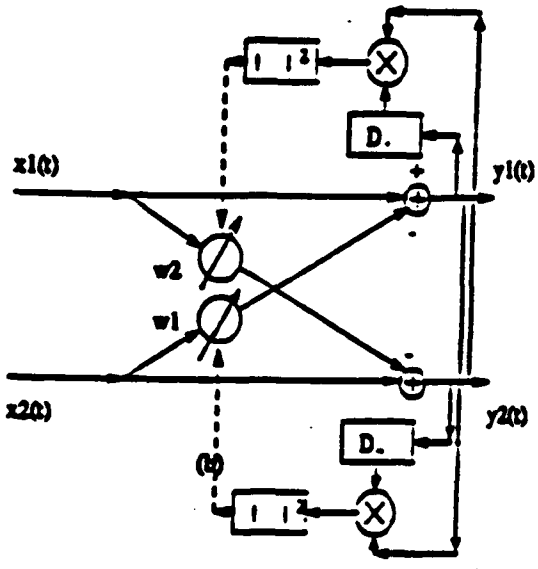
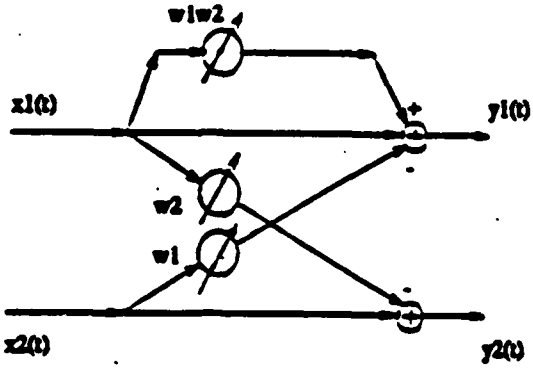


Fig. 4



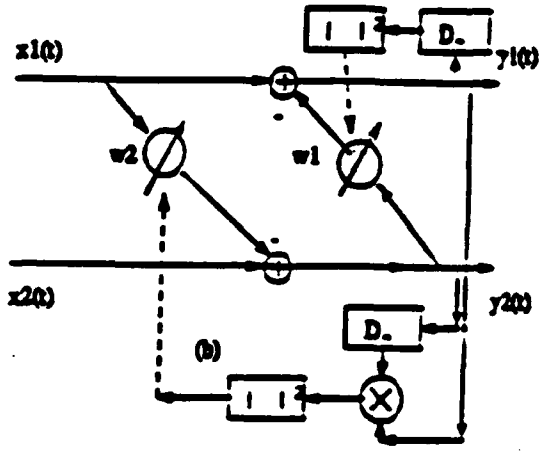


Fig. 6.a

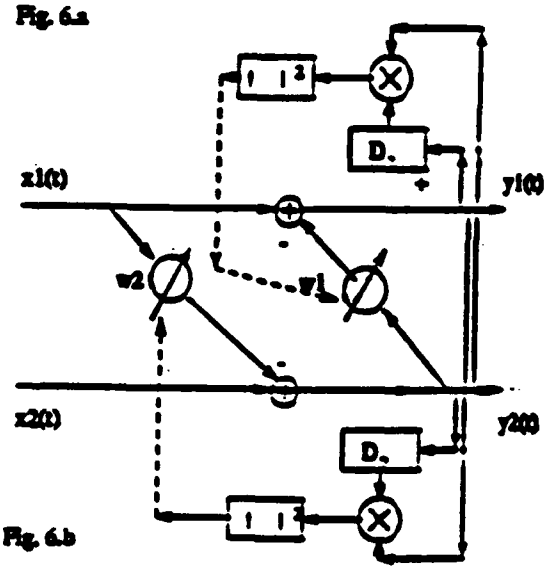


Fig. 6.b

Fig. 6 Optimization Criteria for Forward Forward/Forward and Backward/Forward

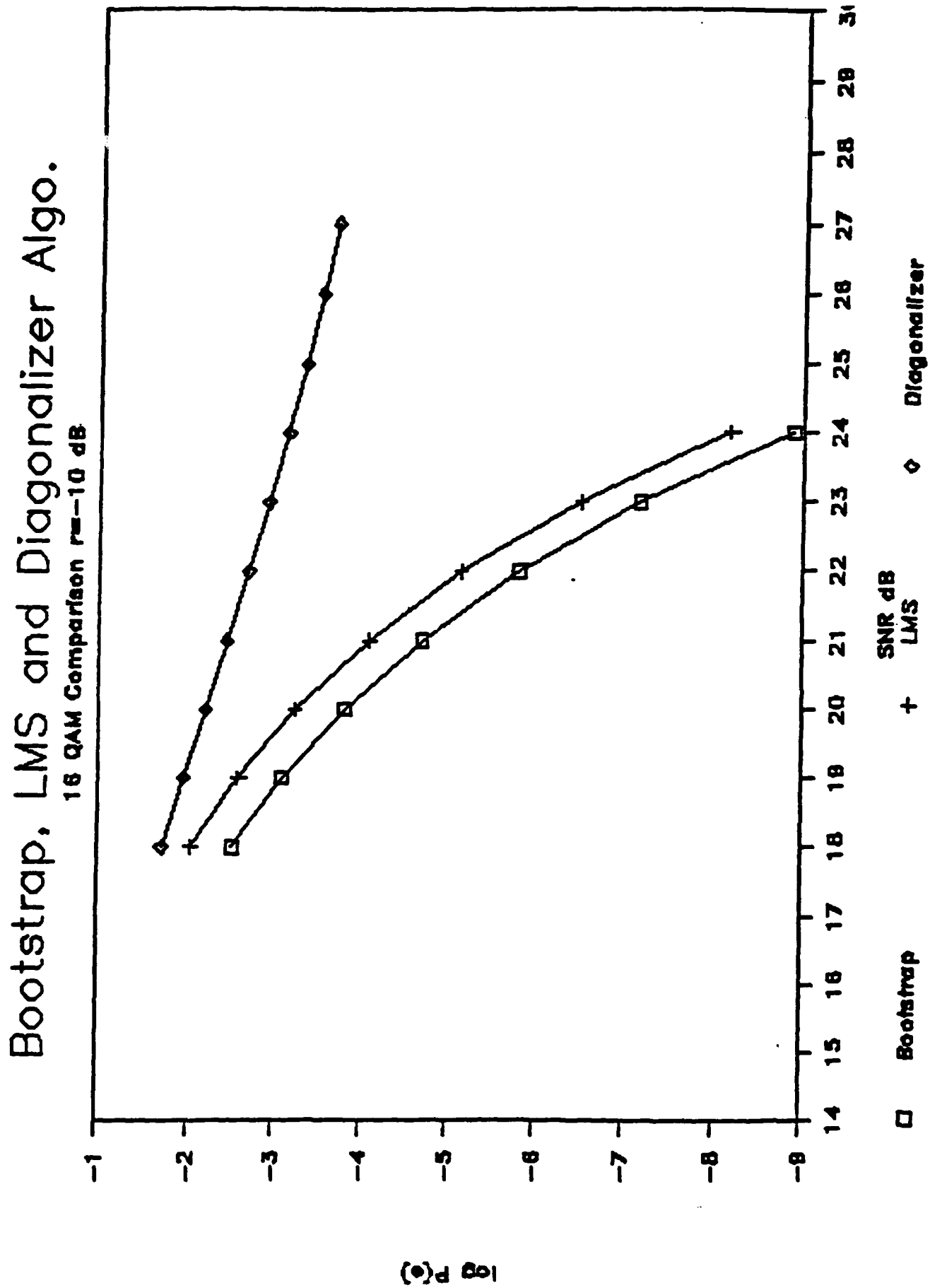


Fig. 7 Performance comparison of Bootstrap, LMS and Diagonalizer 16 QAM. r=-10 dB

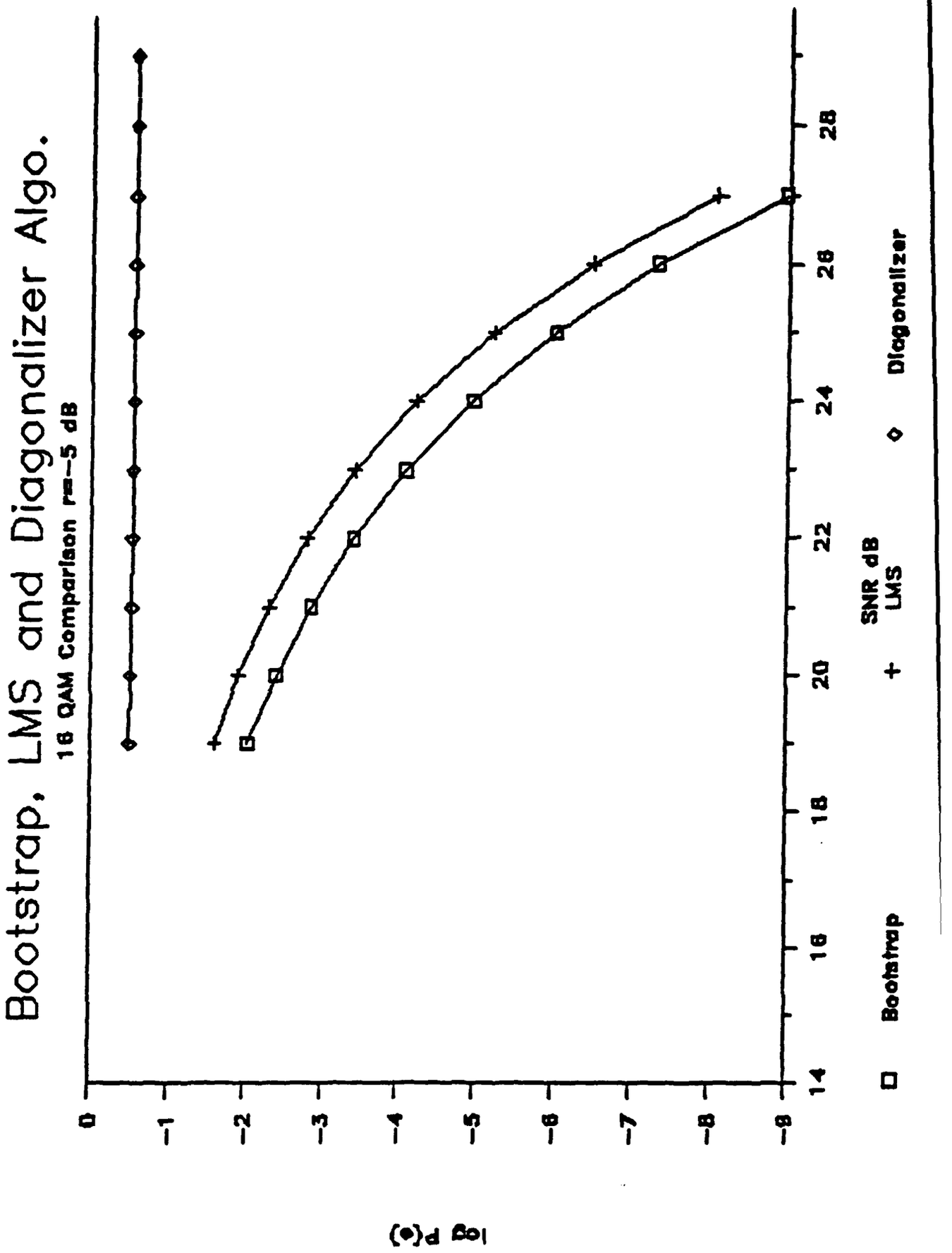


Fig. 8 Performance comparison of Bootstrap, LMS and Diagonalizer 16 QAM, r=-5 dB

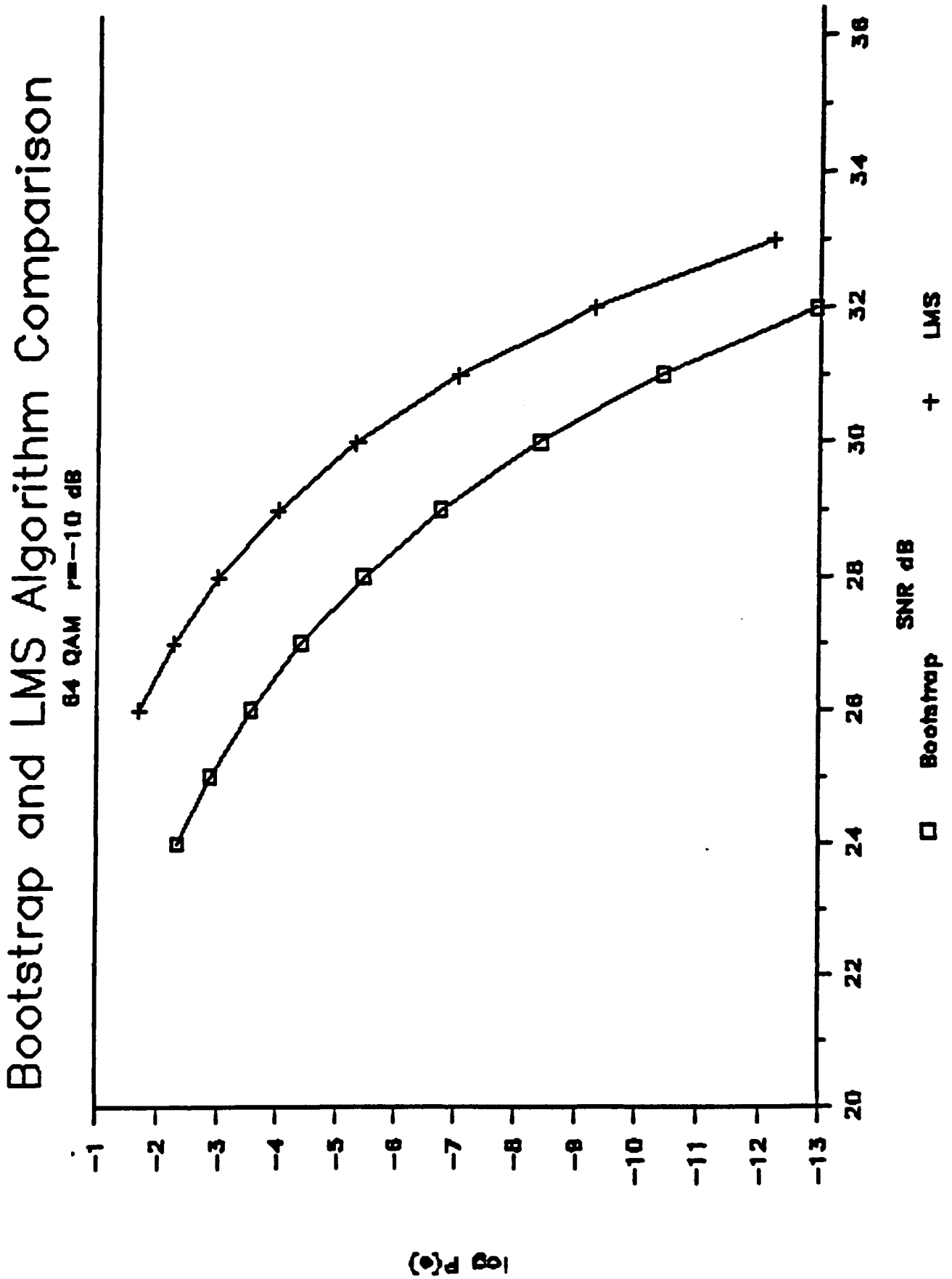


Fig. 9 Performance comparison of Bootstrap and LMS. 64 QAM, r=-10 dB

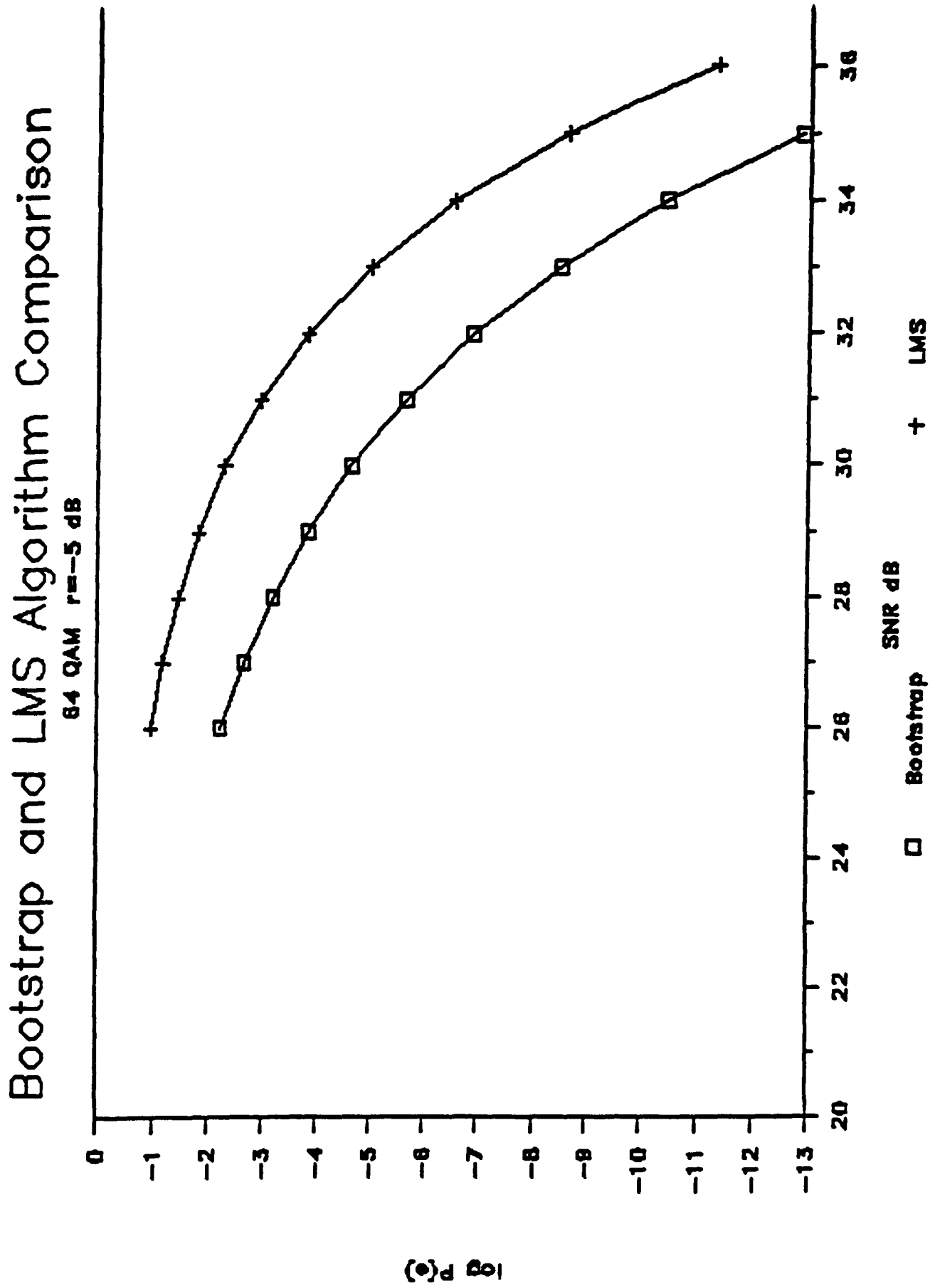


Fig. 10 Performance comparison of Bootstrap and LMS, 64 QAM. r=-5 dB

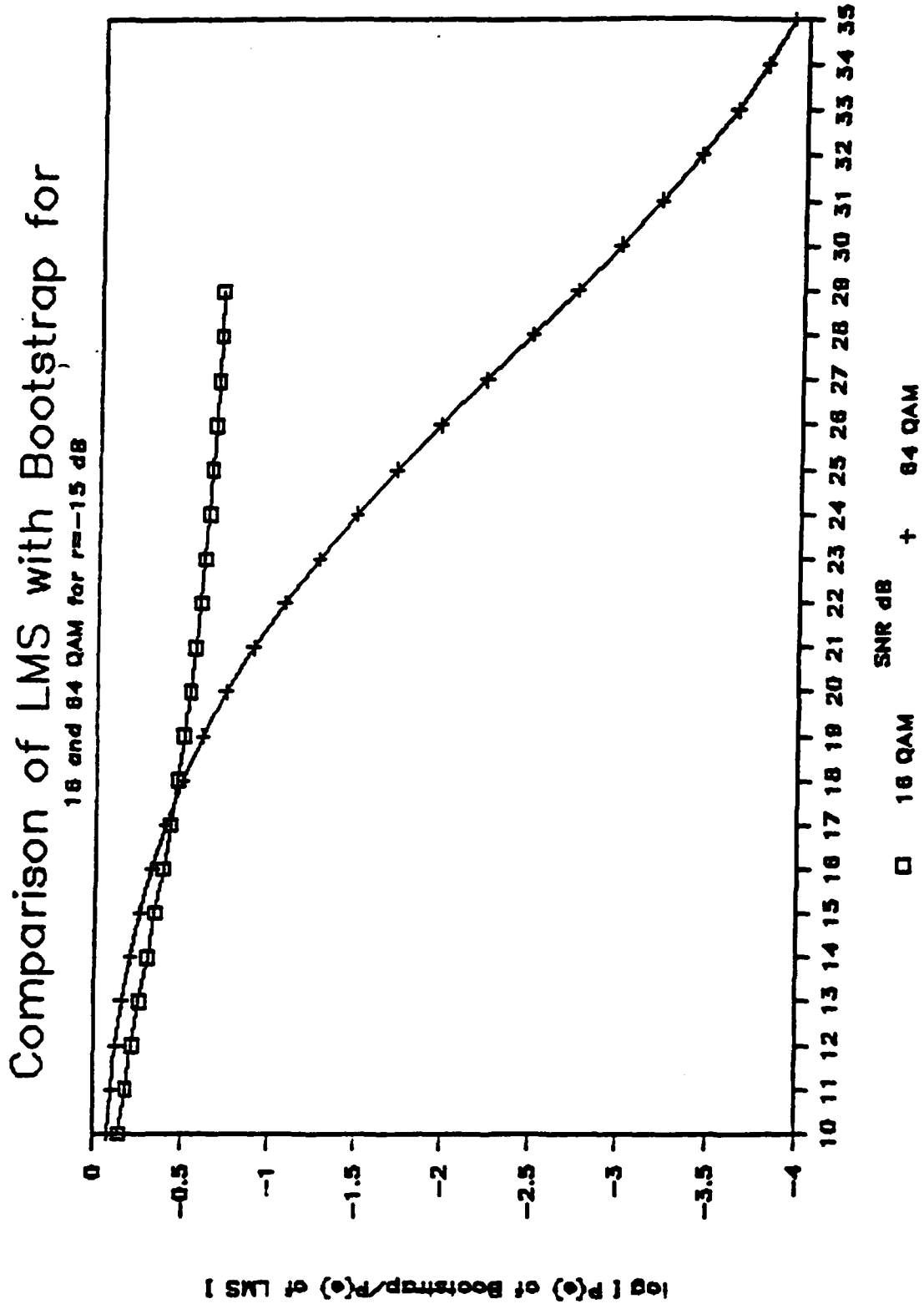


Fig. 11 Bootstrap and LMS Performance Difference 64 QAM, r=-15 dB

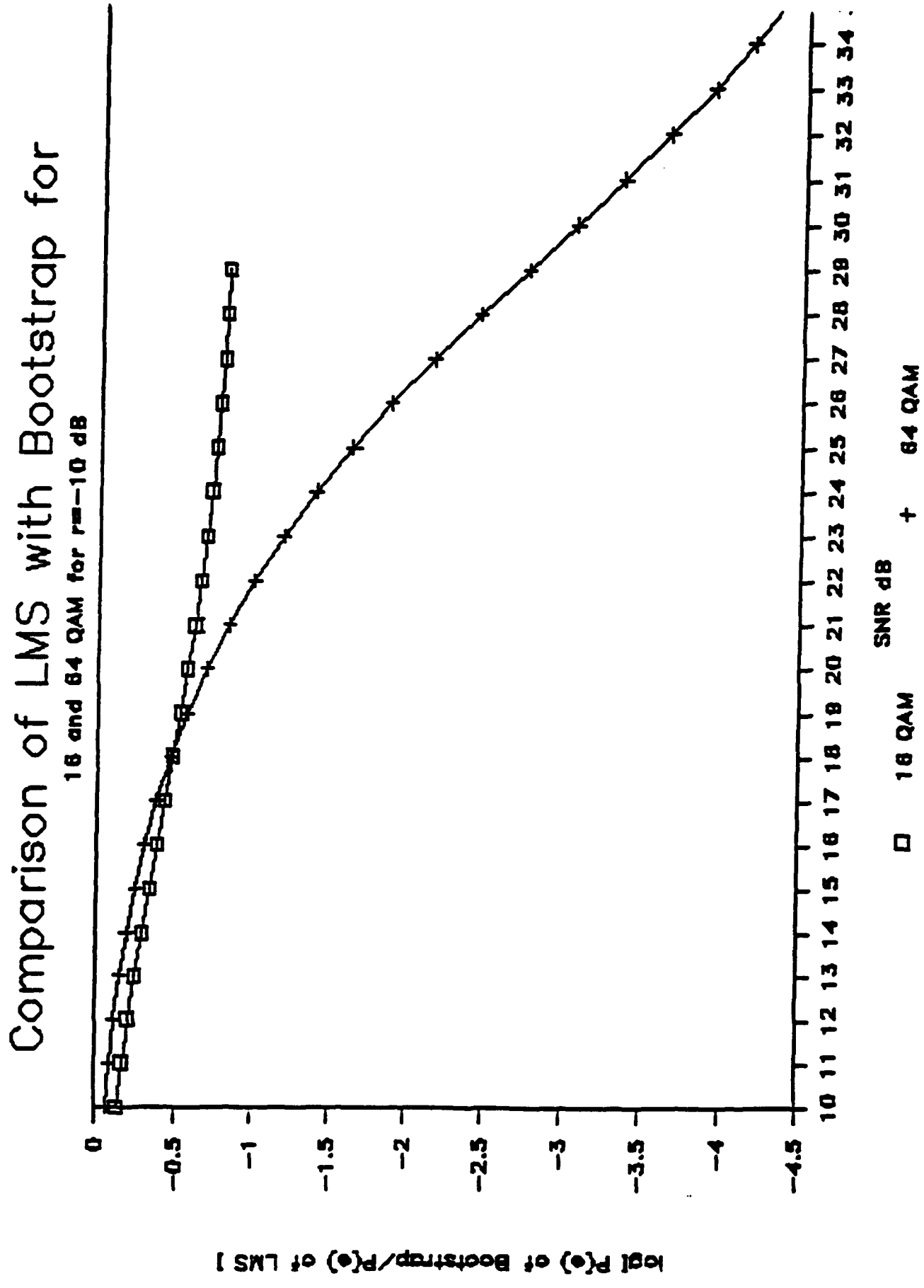


Fig. 12 Bootstrap and LMS Performance Difference 64 QAM, r=-10 dB

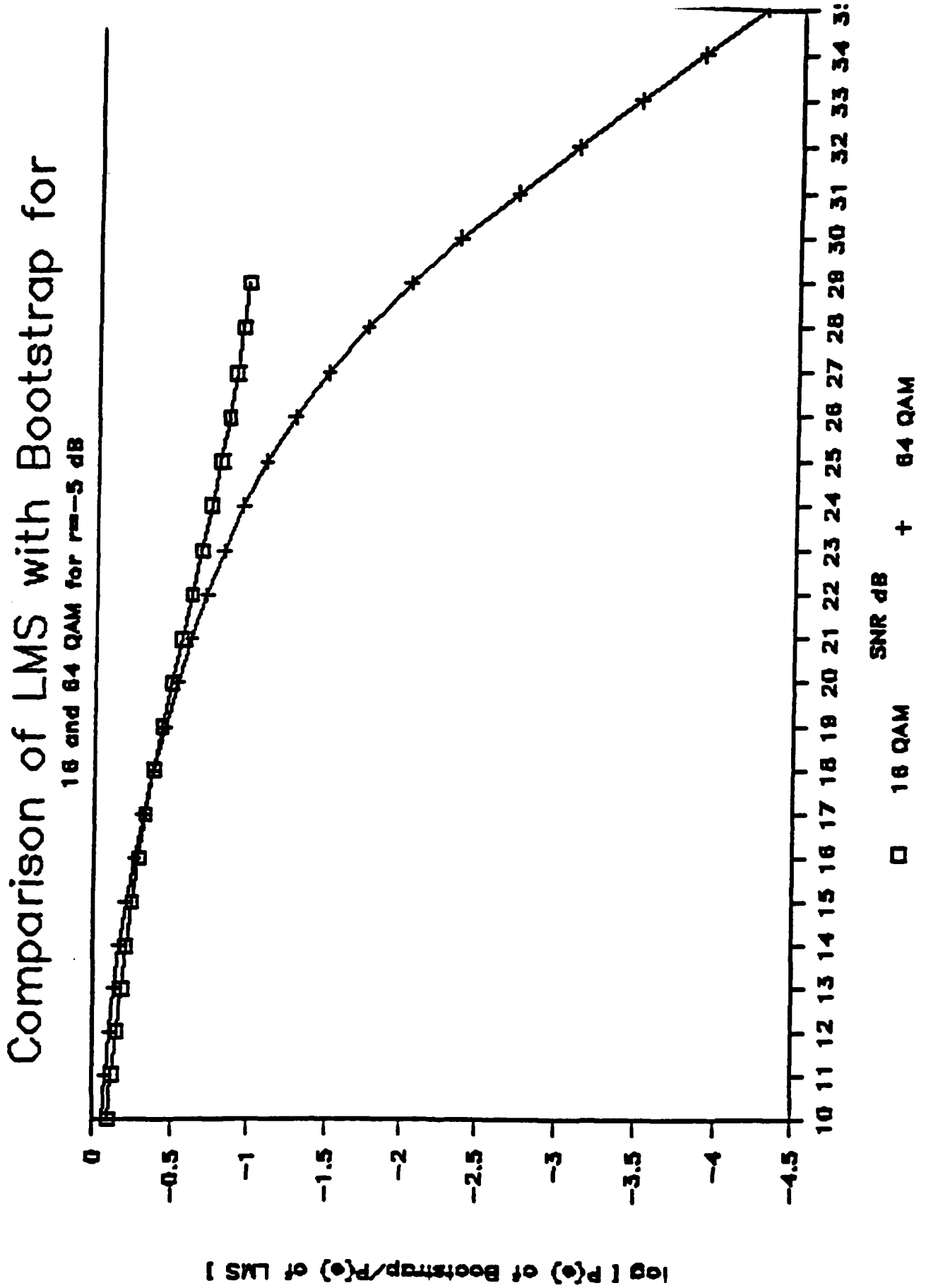


Fig. 13 Bootstrap and LMS Performance Difference 64 QAM. r=-5 dB

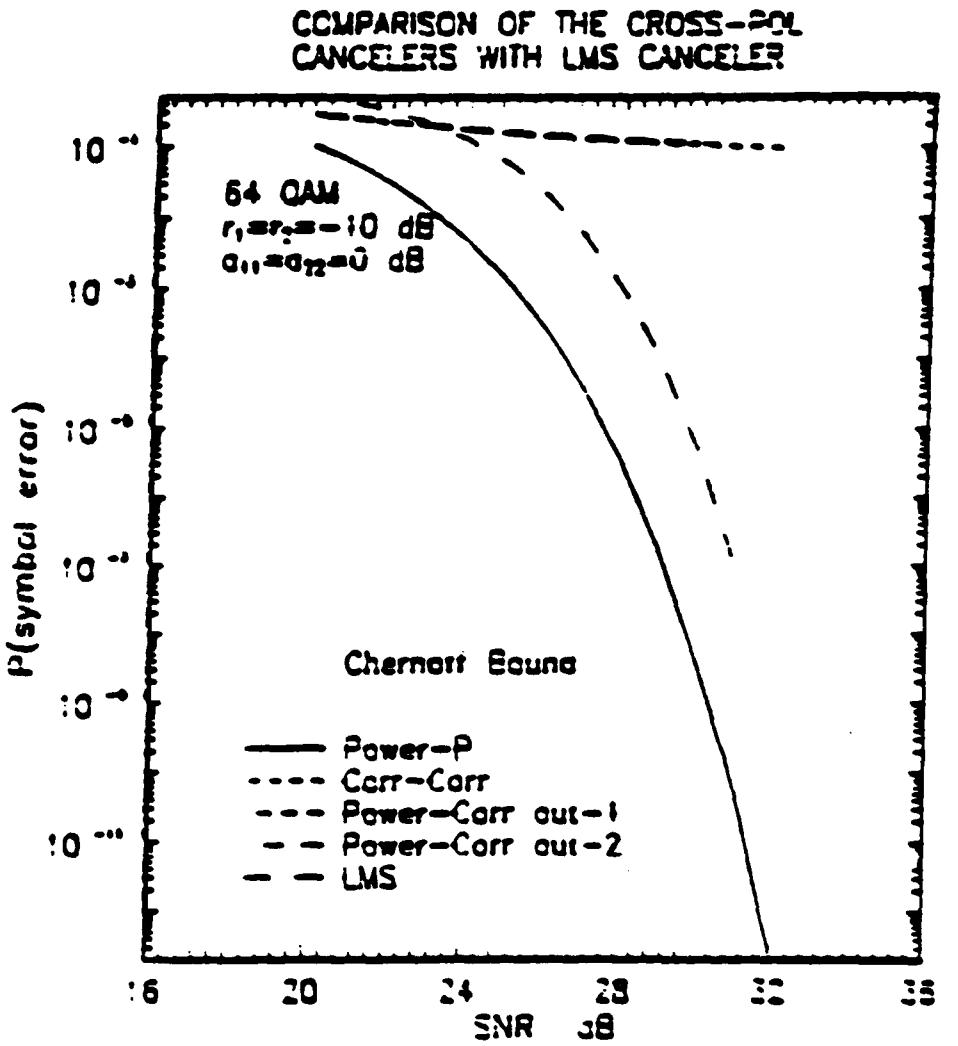


Fig. 14 Comparison of the Cross-pol Cancelers with LMS Canceler

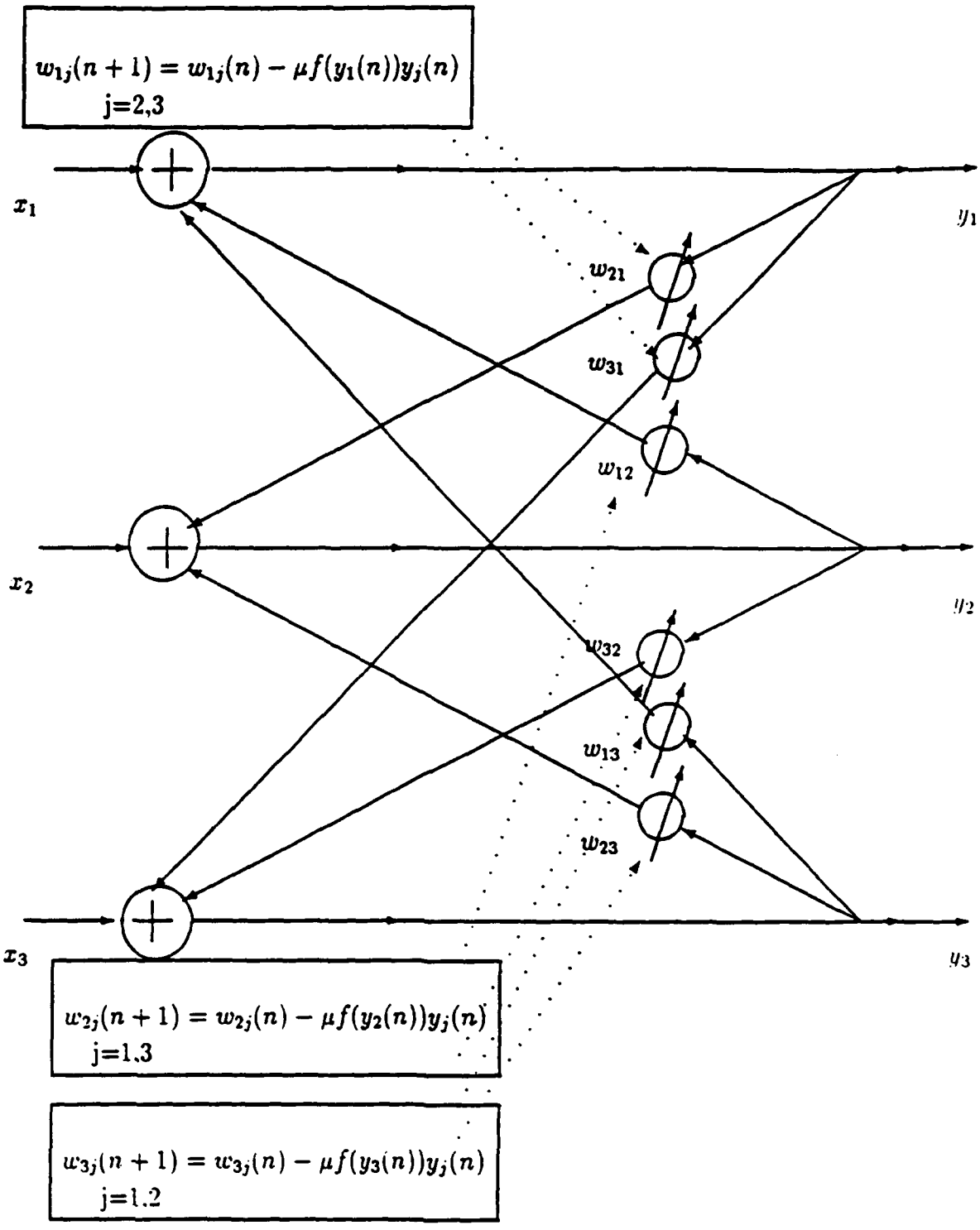


Fig. 15 Multi-Inputs Multi-Outputs Backward/Backward Bootstrapped Separator

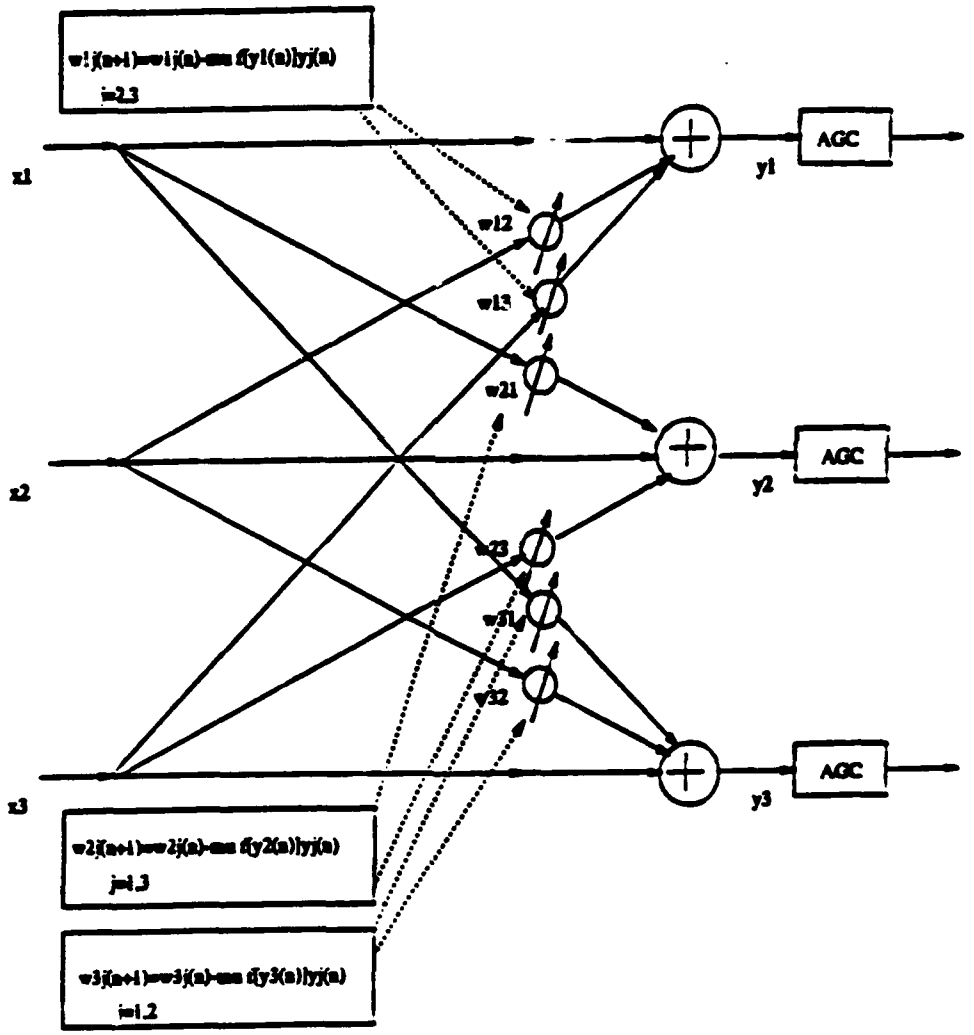


Fig. 16 Multi-Inputs Multi-Outputs Forward/Forward Separator

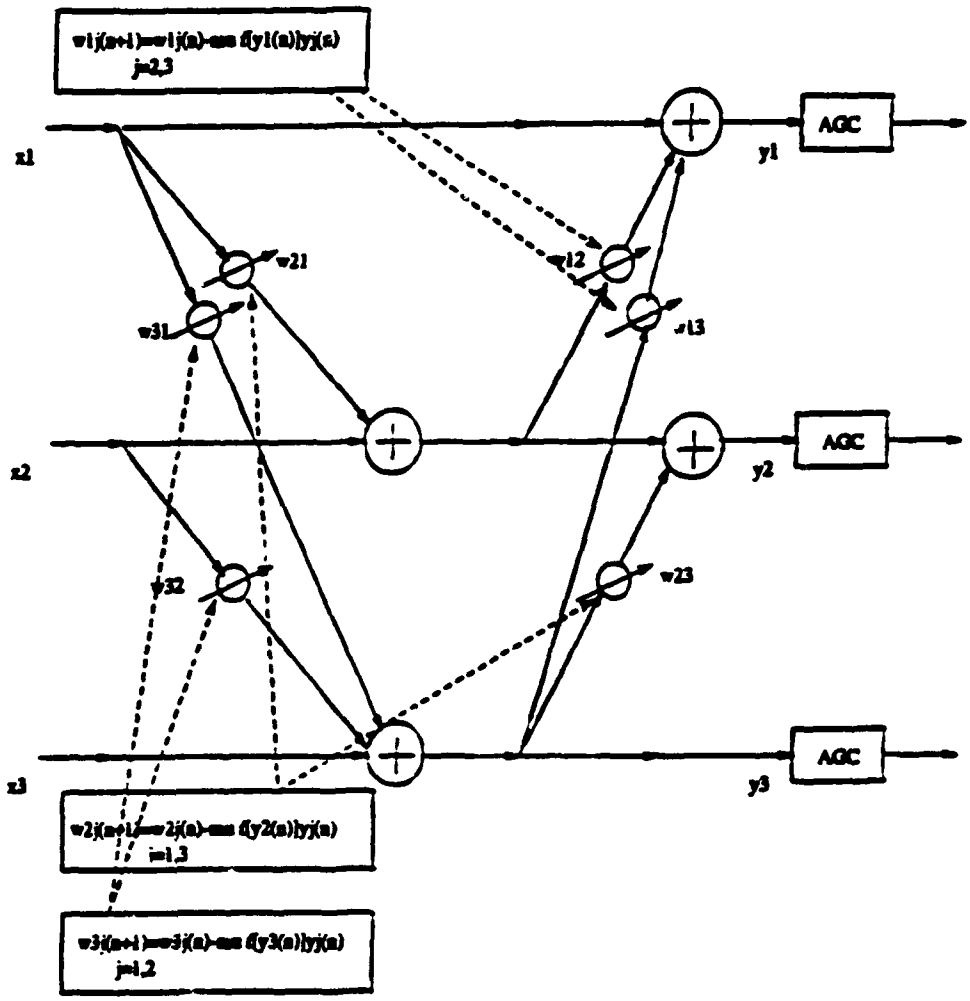


Fig. 17 Multi-Inputs Multi-Outputs Backward/Forward Separator

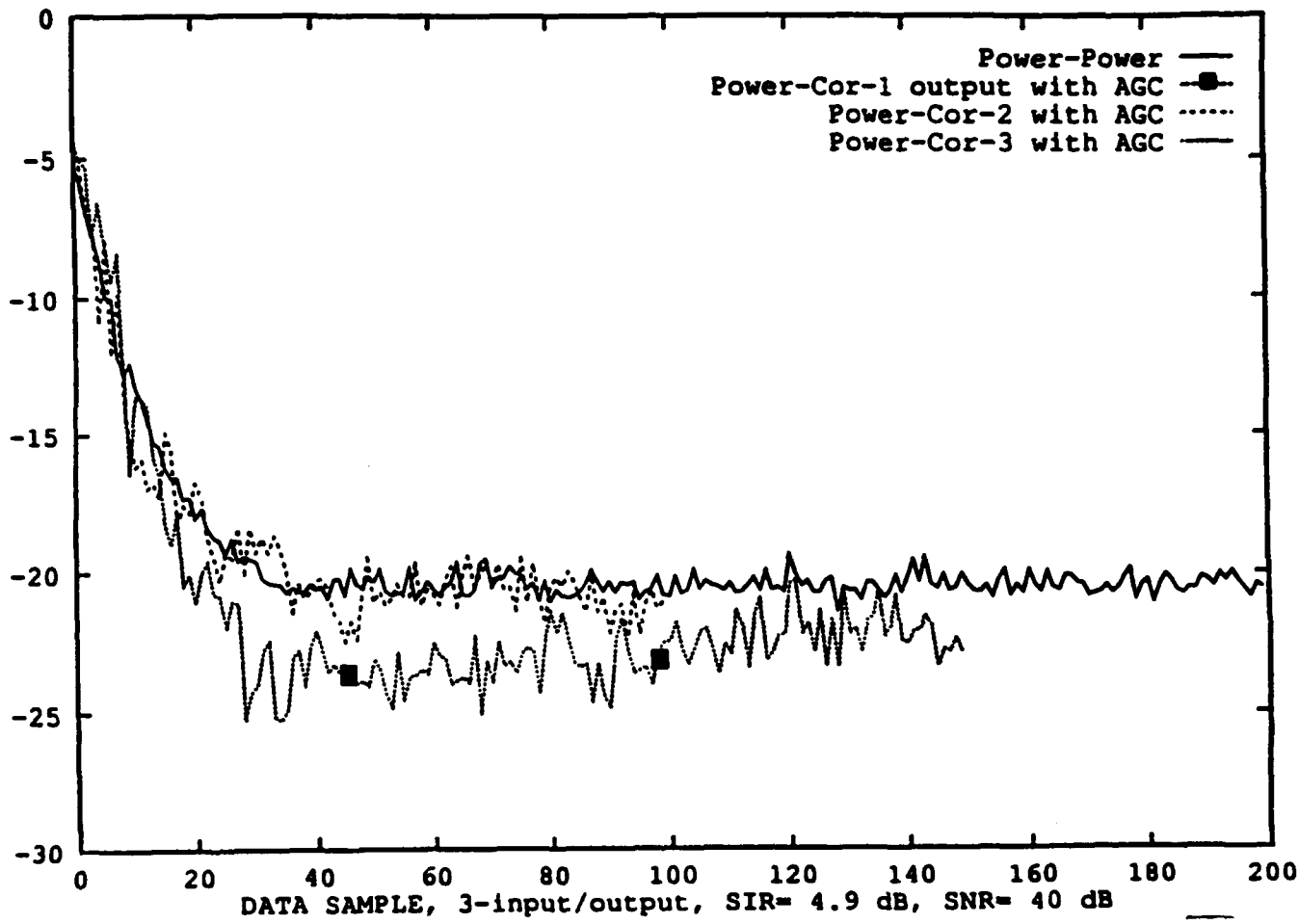


Fig. 18 Convergence Comparison of the Bootstrapped Signal Separators for three-inputs/three-outputs

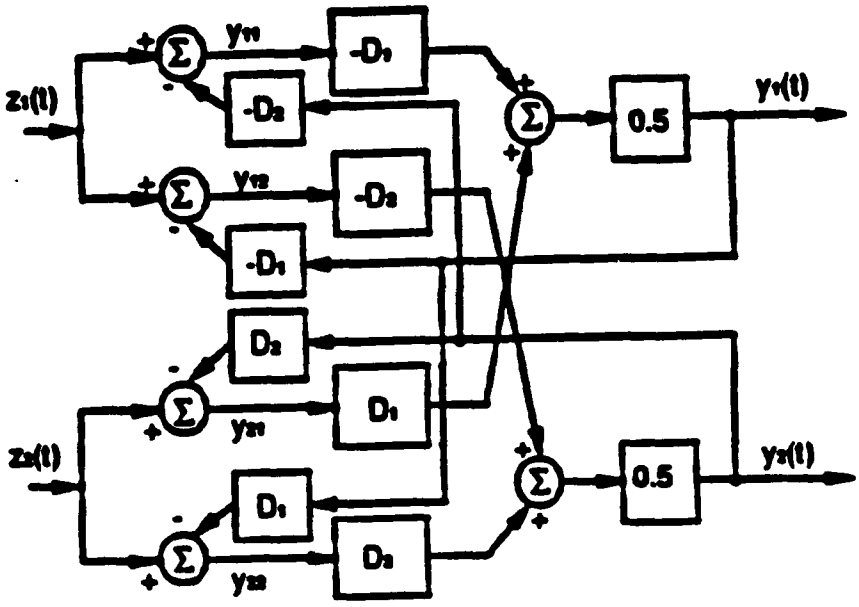


Fig. 19 Bootstrapped Implementation of the Least-Squares Separator.

Appendix A

Bootstrapped Adaptive Separation of Two Superimposed Signals - Steady State Analysis

Y. Bar-Ness

Abstract

The algorithm presented here differs from other interference cancellation systems in that it is a power separator rather than an interference canceler. That is, each of the two input signals interferes with the other and the function of the canceler is to remove that interference from both input signals rather than just one. A novel way to obtain high signal to interference ratio at both output ports is to use the bootstrapping approach.

Three bootstrapped configurations are discussed and analyzed herein. They particularly differ in the criteria used to obtain the optimal complex weight in the cancellation path.

Steady state parameters, such as optimal complex weights, optimal signal power outputs and signal-to-interference power ratios are derived for each of the different configurations. Also calculated are input-output-signal transmission and cancellation factors. In performing the analysis, emphasis is on application for cross-polarization cancellation.

1 Introduction

System degradations in dual orthogonally polarized communication systems may be reduced by the use of a cross-pol interference canceler. A cross pol canceler differs from other interference cancellation systems [1,2] in that it is a power separator rather than an interference canceler. That is, each of the two input signals interferes with the other and the function of the canceler is to remove that interference from both input signals rather than just one. A novel way to obtain a high signal to interference ratio at both output ports is to use a bootstrapping approach [3,4]. In this approach two cancellation paths and two summations are used to obtain the two system outputs, and an adaptive algorithm is employed to optimize the signal-to-interference power ratio at the two output ports simultaneously.

The purpose of this appendix is to report a detailed steady state analysis of the three bootstrapped configurations, two of which were first discussed in [3] and [4][1]. A summary of the results of this appendix is previously reported in the literature [5]. Independently, during the mid-eighties, a group of European researchers addressed the same separation problem and, under the name of "blind signal separation" developed a similar adaptive system [6]. More recently, a third group of signal processing researchers have applied similar ideas to speech signal separation [7].

The three configurations suggested in the appendix differ particularly in the criteria used to obtain the optimal complex weight in the cancellation path. The criteria used minimize either the interfering signals power at the two output ports, the correlation between the two signals at the two output ports, or simultaneously the interfering signal power at one port and the correlation between the two output signals at ports. Correspondingly they will be termed: 1) power-power, power-correlation and correlation-correlation cross pol cancelers. Consequently, the three configurations differ in the topology of their two cancellation paths, and the manner in which

[1]This work in reference [3] did not consider the cross polarization problem.

the adaptive feedback information is derived and hence in their hardware complexity. The use of either of these criteria would lead to the customary power-inversion result. However, each of the above bootstrapping arrangements results in power separation (high signal-to-interference ratio at both output ports) through the use of a discrimination technique.

In this appendix we obtain the optimal complex weights of the three different bootstrapping configurations. and examine the conditions under which, utilizing an adaptive algorithm, these optimal values are attainable. It is shown that crucial to some of the analytical results is the assumption that the input de-polarization is always much less than one. The optimal signal power outputs, for the different configurations, are derived and the optimal signal-to-interference power ratios at the two different output ports are calculated and compared. Also calculated and compared are the input-output transmission ratios and the cancellation factors at the two separate ports. Finally some other questions related to the subject of this appendix, namely steady state analysis. are raised. These will be investigated and reported separately.

Consider the two inputs (in complex envelope notation)

$$v_1(t) = s(t) + bn(t)$$

$$v_2(t) = cs(t) + n(t) \tag{1}$$

where b and c are complex values and $|b|^2$ and $|c|^2$ are the input signal to interference ratios. s(t) and n(t) are zero-mean uncorrelated stationary complex processes. We will not indicate the time-dependence of variables in the interest of brevity unless it is necessary. We now process these two inputs, *separately* using, the simple "noise cancelling" scheme proposed by Widrow [1] (see Figure 1). Here we consider each input as a "reference signal" (in Widrow's terminology) and sum it with a weighted version of the other signal to form an output. Two such outputs $v_q(t)$ and $v_p(t)$ are

formed simultaneously. The corresponding weights α and β (complex valued) can adaptively be adjusted so that the output powers Q and P are minimized, respectively. From Figure 1 we have

$$v_q(t) = v_2(t) + \alpha v_1(t) \tag{2}$$

$$Q = \overline{|v_q(t)|^2} = |c+\alpha|^2\overline{|s|^2} + |1+b\alpha|^2\overline{|n|^2} \tag{3}$$

where the over-bar stands for the expected value. Therefore, $\overline{|s|^2}$and $\overline{|n|^2}$ designate the powers of the two incoming signals respectively. Thus, one can show that,

$$\frac{dQ}{d\alpha} = 2(c+\alpha)\overline{|s|^2} + 2(1+b\alpha)b^*\overline{|n|^2} \tag{4}$$

where the derivative of a real function with respect to (w.r.t.) a complex variable $\alpha(\alpha = \alpha_R + j\alpha_I)$ is defined by; $d(.)/dt = d(.)/d\alpha_R + jd(.)/d\alpha_I$. [2] The asterisk denotes complex conjugate and $j = \sqrt{-1}$. Equating (4) to zero we obtain the optimal value of α

$$\alpha_{op} = -\frac{c\overline{|s|^2} + b^*\overline{|n|^2}}{\overline{|s|^2} + |b|^2\overline{|n|^2}} \tag{5}$$

Substituting in (3) we obtain the power of the s signal Q_s and of the n signal Q_n. These, following simple manipulation, are given by,

$$Q_s = \left|-\frac{b^*(1-cb)\overline{|n|^2}/\overline{|s|^2}}{1+|b|^2\overline{|n|^2}/\overline{|s|^2}}\right|^2 \overline{|s|^2} \tag{6}$$

$$Q_n = \left|\frac{(1-cb)}{1+|b|^2\overline{|n|^2}/\overline{|s|^2}}\right|^2 \overline{|n|^2} \tag{7}$$

and the n signal-to-s signal power ratio is

$$Q_n/Q_s = \frac{\overline{|s|^2}}{|b|^2\overline{|n|^2}} \tag{8}$$

[2] In obtaining the complex derivative of a real function $A = |d+e\delta|^2$ w.r.t. to the complex variable δ we used the simple general rule; $dA = 2e^*(d+e\delta)$

$$v_p(t) = v_1(t) + \alpha v_2(t) \tag{9}$$

$$P = \overline{|v_p(t)|^2} = |b + \beta|^2\overline{|n|^2} + |1 + \beta c|^2\overline{|s|^2} \tag{10}$$

so that

$$\beta_{op} = -\frac{b\overline{|n|^2} + c^*\overline{|s|^2}}{\overline{|n|^2} + |c|^2\overline{|s|^2}} \tag{11}$$

The power of the s signal and the n signal are, respectively;

$$P_n = \left|-\frac{c^*(1-cb)\overline{|s|^2}/\overline{|n|^2}}{1 + |c|^2\overline{|s|^2}/\overline{|n|^2}}\right|^2 \overline{|n|^2} \tag{12}$$

$$P_s = \left|\frac{(1-cb)}{1 + |c|^2\overline{|s|^2}/\overline{|n|^2}}\right|^2 \overline{|n|^2} \tag{13}$$

and

$$P_s/P_n = \overline{|n|^2}/|c|^2\overline{|s|^2} \tag{14}$$

Like (8), equation (14) represents a "power inversion" relation with respect to the power ratio at the input to the weighted elements. In fact, if $|c| \neq |b|$ then we expect an improvement in power ratio of one signal to the other at one port, while a degradation of the power ratio of the second signal to the first at the second port. For example, if $|b| < |c|$ then Q_s/Q_n is greater than the s signal-to-n signal power ratio (snr) at input terminal No.1,while P_n/P_s is less than n signal-to-s signal power ratio (nsr) at input terminal No.2.

Similar results prevail if instead of the criterion of minimizing, separately, the output powers Q and P we would implement the criterion of minimizing the correlation of each output with the corresponding input. That is to adjust α and β so that, $\left|\overline{v_q(t)v_1^*(t)}\right|^2$ and $\left|\overline{v_p(t)v_2^*(t)}\right|^2$ are minimized respectively. (See Fig.1)

In conclusion, the "noise cancelling" scheme can not be used in cross polarization problems where high quality power separation is required, i.e. improvement in both snr at one port and nsr at the other. This motivates one to explore the bootstrapping approach wherein the general idea is to feedback the power-inverted "noise canceler" output as a "reference signal" for the other "noise canceler" and vice versa. Suitable feedback arrangements enhance the right power ratios resulting in power separation with high signal-to-interference ratios at both output ports.

2 Bootstrapping Techniques

2.1 Optimization Criterion

In this section we investigate three bootstrapping configurations in which the following optimization criterion are employed, these are:

1. Power-Power criterion
2. Power-Correlation criterion
3. Correlation-Correlation criterion

These configurations are depicted in Figures 2 through 4 respectively. The first configuration follows Bar-Ness and Rokach [4], the second was considered in unpublished communications by Steinberger, while the third resembles the cancellation network proposed by Brandwood [3]. Notice the discrimination networks in these configurations. Such a discriminator which depends on the properties of the signals may turn out to be a limiter, filter, etc. Its sole effect is to slightly emphasize one signal with respect to the other. Their necessity will be obvious later. In references [3] and [4] where narrow band signals were assumed, hard limiters were implemented as discrimination techniques. Computer simulations have shown that limiters are only effective in a narrow class of applications [5]. For satellite communication [5]

in which the frequency plans for the two polarizations are staggered, a frequency discrimination technique [8] with simple comb filter was used.

Consider the power-power canceler of Figure 2. Let the power ratio of the two signals at point No. 3 be such that $s > n$ (even if only slightly greater). Point No. 3 being the input to the weighted element of the β processor (terminal No. 1 is the other input) will result (because of power-inversion) in $n > s$ at point No. 4 and output. $v_p(t)$. But point No. 4 being the input to the weighted element of the α processor (terminal No. 2 is its other input) will result in s still greater than n at point No. 3 and hence at output port, $v_q(t)$. This process of bootstrapping will continue, resulting in a very high snr at one port and a very high nsr at the other. The snr and nsr, at these ports respectively, will be upper bounded by values depending on the noise, the impurities of the system and control errors. Consequently, the power-power canceler of Figure 2 acts as a high quality power separator. In fact ideal separation occurs only in a noise and impurity free case.

In the cross pol canceler of Figure 3, the cancellation weight β is controlled via a power criterion which minimizes the power P at the $v_p(t)$ port. The weight α is controlled by minimizing the correlation between the two outputs (i.e. $\left|\overline{v_q(t)v_p^*(t)}\right|^2$) thus, the name power-correlator canceler. Here, to obtain a perfect cancellation of the n signal at $v_p(t)$, it is required that the β processor has a clean sample of n signal at point No. 3 (because of power-inversion behavior). The correlation processor which controls α can operate with a sample of the signal n (at point No. 3) which is corrupted by the signal s, but needs a clean sample of the signal s (at the other correlator input) to generate its feedback signal. Since one of the two processors (the correlation) can in effect defer its need for a clean sample of the signal n, it makes sense to let that processor perform its cancellation first, making a clean sample of n available to the other processor (power processor), which in turn provides a clean sample of the signal s (at point No. 4) to the first processor (correlation processor) to generate a feedback

signal. Thus, although neither processor can function properly unless the other one does, both (in bootstrap operation) can operate properly together. Consequently, the power-correlator canceler of Figure 3 can perform as a power separator. It might seem complicated to comprehend such a qualitive argument, nevertheless analysis will support these claims.

Finally, the correlator-correlator canceler of Figure 4 operates as follows: The α processor can operate with a sample of the single n (at point No. 3) which is corrupted by the signal s, but needs a clean sample of the signal s (at point No. 4) to generate its feedback. Similarly with the β processor, where the signals n and s (at the point No. 3 and No. 4) are interchanged. Since initially neither points, No. 3 or No. 4 contains a clean sample of s or n. respectively, neither processor performs properly unless the other one does. However, if one processor starts its cancellation it results in a cleaner sample of the proper signal needed by the other processor and vice verse. This bootstrapping behavior finally results in the desired power separation at the output ports.

2.2 The Optimal Weights

The optimal weights for the different configurations are now obtained using the corresponding optimization criterion

2.2.1 The Power-Power Scheme

From Figure 2,

$$v_p(t) = v_1(t) + \beta v_q(t) \tag{15}$$

$$v_q(t) = v_2(t) + \alpha v_p(t) \tag{16}$$

By using (1) we get,

$$P = \overline{|v_p(t)|^2} = \frac{|1+\beta c|^2\overline{|s|^2} + |b+\beta|^2\overline{|n|^2}}{|1-\alpha\beta|^2} \tag{17}$$

where we used the fact that $s(t)$ and $n(t)$ are zero-mean uncorrelated processes. Taking the derivative of (17) we get

$$\frac{\partial P}{\partial \beta} = -\frac{2(1-\alpha\beta)[(1+\beta c)(\alpha+c)^*\overline{|s|^2} + (b+\beta)(1+\alpha b)^*\overline{|n|^2}]}{|1-\alpha\beta|^4} \tag{18}$$

From $\partial P/\partial\beta = 0$ we get after simple manipulation,

$$(\alpha+c)^*(1+\beta c)\overline{|s|^2} + (b+\beta)(1+\alpha b)^*\overline{|n|^2} = 0 \tag{19}$$

and

$$\beta_{op} = -\frac{(\alpha+c)^*\overline{|s|^2} + b(1+\alpha b)^*\overline{|n|^2}}{c(\alpha+c)^*\overline{|s|^2} + (1+\alpha b)^*\overline{|n|^2}} \tag{20}$$

provided $1-\alpha\beta \neq 0$. Thus, if $\alpha = -c$, then $\beta_{op} = -b$ and perfect cancellation will result. Also from Figure 2

$$Q = \overline{|v_q(t)|^2} = \frac{|1+\alpha b|^2\overline{|n|^2} + |c+\alpha|^2\overline{|s|^2}}{|1-\alpha\beta|^2} \tag{21}$$

Therefore, $\partial Q/\partial\alpha = 0$ implies

$$(1+b\alpha)(b+\beta)^*\overline{|n|^2} + (\alpha+c)(1+\beta c)^*\overline{|s|^2} = 0 \tag{22}$$

and

$$\alpha_{op} = -\frac{(b+\beta)^*\overline{|n|^2} + c(1+\beta c)^*\overline{|s|^2}}{b(b+\beta)^*\overline{|n|^2} + (1+\beta c)^*\overline{|s|^2}} \tag{23}$$

provided $1-\alpha\beta \neq 0$. Again if $b = -\beta$ $\alpha_{op} = -c$ and perfect cancellation results.

Comparing (19) and (22), we notice that their left hand side terms form a complex conjugate pair. Therefore, they constitute a degenerate set of equations which does not have a unique solution in α and β, so that difficulties are to be expected if α_{op}, β_{op} are sought via any search algorithm.

This difficulty can be remedied by using the signals' discrimination networks whose effect on the two signals is different in some way e.g. complex gain, etc. Such a network helps distinguish one signal from the other by putting emphasis on either one. For example, if discriminator No.1 is such that its output power is P_1 given by (17) with s_1 and n_1 replacing s and n, respectively, then (19) and (20) will change accordingly. Similarly using Q_2 to drive α, where Q_2 is given by (21) with s_2 and n_2 replacing s and n respectively, then (22) and (23) will change accordingly. If we have at least $s_1 \neq s_2$, or $n_1 \neq n_2$ then the set of Equations (19) and (22) is no longer degenerate. They are a set of two bilinear equations in α and β. In fact with the effect of the discriminator (17) and (21) can be rearranged to get

$$(\alpha + c)(1 + \beta c)^*(\overline{|s_1|^2}\,\overline{|n_2|^2}/\overline{|n_1|^2} - \overline{|s_2|^2}) = 0 \tag{24}$$

$$(\beta + b)(1 + \alpha b)^*(\overline{|n_2|^2}\,\overline{|s_1|^2}/\overline{|s_2|^2} - \overline{|n_1|^2}) = 0 \tag{25}$$

If based on some slight information about the signals, discriminators[3] can be found such that at their output we have $\overline{|s_1|^2}\,\overline{|n_2|^2} \neq \overline{|s_2|^2}\,\overline{|n_1|^2}$ then the only two solutions are $\alpha = -c$ and $\beta = -b$ or $\alpha = \frac{1}{b}$ and $\beta = -\frac{1}{c}$. The second pair is not valid if we constrain $|\alpha|$ and $|\beta|$ to be smaller than unity (b and c are assumed to have magnitude less than unity). The unique solution $\alpha_{op} = -c$ and $\beta_{op} = -b$ are the required values of α and β that renders perfect cancellation at both outputs.(see (17) and (21))

[3]Such discriminator has been shown to exist for satellite communication [5] microwave radio [10] and narrow-band tactical communication [9] application.

2.2.2 The Power-Correlation Scheme

From Figure 3,we have

$$v_p(t) = v_1(t) + \beta v_q(t) \tag{26}$$

$$v_q(t) = v_2(t) + \alpha v_1(t) \tag{27}$$

$$P = \overline{|v_p(t)|^2} = |(1 + \beta(\alpha + c)|^2\overline{|s|^2} + |b + \beta(1 + \alpha b)|^2\overline{|n|^2} \tag{28}$$

where we used the assumption that $s(t)$ and $n(t)$ are zero-mean uncorrelated processes. Equating the derivative of (28) with respect to β to zero we get

$$(\alpha + c)^*\overline{|s|^2} + \beta|\alpha + c|^2\overline{|s|^2} + \beta|1 + \alpha b|^2\overline{|n|^2} + b(1 + \alpha b)^*\overline{|n|^2} = 0 \tag{29}$$

and

$$\beta_{op} = -\frac{b(1 + \alpha b)^*\overline{|n|^2} + (\alpha + c)^*\overline{|s|^2}}{|1 + \alpha b|^2\overline{|n|^2} + |\alpha + c|^2\overline{|s|^2}} \tag{30}$$

Thus if $\alpha = -c$ then $\beta_{op} = -b/(1 + \alpha b)$ and a perfect cancellation of the n signal will result (see(28)). With α close to -c, we can assume that $|\alpha + c|^2 \ll |1 + \alpha b|^2\overline{|s|^2}/\overline{|n|^2}$[4] and hence

$$\beta_{op} = -\frac{b}{1 + \alpha b} - \frac{(\alpha + c)^*}{|1 + \alpha b|^2}\frac{\overline{|s|^2}}{\overline{|n|^2}} \tag{31}$$

Also from Figure 3

$$Q_2 = \left|\overline{v_q(t)v_p^*(t)}\right|^2 = |A|^2 \tag{32}$$

where as it can be easily shown. using the zero correlation property of $s(t)$ with $n(t)$ that

$$A = (\alpha + c)(1 + \beta(\alpha + c))^*\overline{|s|^2} + (1 + \alpha b)(b + \beta(1 + \alpha b))^*\overline{|n|^2} \tag{33}$$

Notice that in comparison to the corresponding terms in (17), (21) or (28) the terms in $|A|^2$ do not depend linearly on the independent variable α (see also Appendix A-1) and hence the general rule for the derivative in the footnote on page 4 is not valid.

[4]Since the signal propagate through the same path one can take without loss of generality that $\overline{|s|^2}$ and $\overline{|n|^2}$ are of the same order.

However, directly from (P-1) and (P-2) in Appendix A-1, we have

$$\partial Q_2/\partial \alpha_R = A^* \partial A/\partial \alpha_R + A \partial A^*/\partial \alpha_R \tag{34}$$

$$\partial Q_2/\partial \alpha_I = A^* \partial A/\partial \alpha_I + A \partial A^*/\partial \alpha_I \tag{35}$$

where the subscripts R and I designate the real and imaginary parts, respectively. From (33),

$$\partial A/\partial \alpha_R = \overline{|s|^2} + 2\beta^*(\alpha + c)_R \overline{|s|^2} + |b|^2 \overline{|n|^2} + 2\beta^*(b_R + \alpha_R |b|^2)\overline{|n|^2} \tag{36}$$

and

$$\partial A/\partial \alpha_I = j\overline{|s|^2} + 2\beta^*(\alpha + c)_I |s|^2 + j|b|^2 |n|^2 + 2\beta^*(-b_I + \alpha_I |b|^2) j\overline{|n|^2} \tag{37}$$

It is clear that $\partial A^*/\partial \alpha_R = (\partial A/\partial \alpha_R)^*$ and $\partial A^*/\partial \alpha_I = (\partial A/\partial \alpha_I)^*$. Equation (34) and (35) become simultaneously equal to zero if either one of the following conditions is satisfied.

1. $A = A^* = 0$, $\partial A/\partial \alpha_R \neq 0$ $(\partial A^*/\partial \alpha_R \neq 0)$

   and $\partial A/\partial \alpha_I \neq 0$ $(\partial A^*/\partial \alpha_I \neq 0)$

2. $\partial A/\partial \alpha_R = 0$ and $\partial A/\partial \alpha_I = 0$ $A \neq 0$

3. Neither of the terms above is zero, but the corresponding sums in (33) and (34) are zero.

Under the first condition we have from (33)

$$(\alpha + c)\overline{|s|^2} + \beta^* |a + c|^2 \overline{|s|^2} + \beta^* |1 + \alpha b|^2 \overline{|n|^2} + b^*(1 + \alpha b)\overline{|n|^2} = 0 \tag{38}$$

With α close to -c and $|\beta| < 1$ we have $\beta_R |\alpha + c|^2 \ll (\alpha + c)_R$ and $\beta_I |\alpha + c|^2 \ll (\alpha + c)_I$, and we can neglect the second term in (38) leaving

$$\alpha_{op} = -c - (1 + \alpha_{op} b)(b + \beta(1 + \alpha_{op} b))^* \overline{|n|^2}/\overline{|s|^2} \tag{39}$$

Again if $\beta = -b/(1 + \alpha_{op}b)$, then $\alpha_{op} = -c$ and a perfect cancellation of the s signal at $v_q(t)$ will follow. Finally we notice from the conclusion following (30) and (38) that the condition for the first leads to a result which is needed by the second equation and vice verse. Hence if these two equations are solved simultaneously, they lead to the stated α_{op} and β_{op}.

As in the power-power arrangement, Equations (29) and (38) constitute a degenerate set of equation which does not have a unique solution. Therefore signals' discrimination networks are needed if α_{op} and β_{op}, the solution of (29) and (38), are sought via a search algorithm.

Under the third condition above (i.e., when none of the terms A, $\partial A/\partial\alpha_R$ or $\partial A/\partial\alpha_I$ equals zero, but the corresponding sums of (34) and (35) do), it is possible to show by taking the required derivative that the optimum values of α_R and α_I are given by

$$\alpha_R = -\frac{c_R\overline{|s|^2} + b_R\overline{|n|^2}}{\overline{|s|^2} + |b|^2\overline{|n|^2}} - \frac{Real[A]}{2Real[AB]} \tag{40}$$

$$\alpha_I = -\frac{c_I\overline{|s|^2} - b_I\overline{|n|^2}}{\overline{|s|^2} + |b|^2\overline{|n|^2}} - \frac{I_m[A]}{2Real[AB]} \tag{41}$$

By examining the nature of the solution of the minimization problem, α, defined by (40) and (41), one can show that such a solution can not exist in practice if we restrict β to satisfy

$$|\beta_R|_{max} = |\beta_I|_{max} \leq 1/2\sqrt{2}(\overline{|s|^2}/\overline{|n|^2})^{1/2}_{min}(1/|1 + \alpha b|_{max}) \tag{42}$$

or if (using the stronger condition),

$$|\beta_R|_{max} = |\beta_I|_{max} \leq 1/2\sqrt{2}(\overline{|s|^2}/\overline{|n|^2})^{1/2}_{min}[1/(1 + |\alpha||b|_{max}] \tag{42a}$$

In which case this extremum point will be either a point of maximum or a saddle depending whether, or not, β_R and β_I are of the same order. Particularly if β is constrained to satisfy (42), then this extreme value of α will not be a minimum point.

To exhibit numerically the meaning of this design constrain, we first notice using (30) with $\alpha \simeq -c$ that $|\beta_{op}| = |b|/|1 + \alpha b| \leq |b|/(1 - |\alpha||b|) \leq |b|_{max}/(1 - |\alpha||b|_{max})$. Therefore for any possible β_{op} to be within the value of (42a) it is sufficient to require that

$$|b|_{max}/(1 - |\alpha||b|_{max}) \leq 1/2\sqrt{2}(\overline{|s|^2}/\overline{|n|^2})^{1/2}_{min}[1/(1 + |\alpha||b|_{max})] \qquad (42b)$$

Assuming $(\overline{|s|^2}/\overline{|n|^2})_{min} \simeq 1$ then for $|\alpha| \simeq c| \simeq b|_{max}$, (41b) is satisfied if $|b|_{max} \leq 0.295$. Let the de-polarization coefficients $|b|$ and $|c|$ be at most 0.295 ($\simeq 10db$), then the upper bound on β_R and β_I given by (42a) equals 0.326.

Notice that for a wider useful range of β, $(\overline{|s|^2}/\overline{|n|^2})_{min}$ should be largest. If, for example, the expected $(\overline{|s|^2}/\overline{|n|^2})_{min}$ is less than the expected $(\overline{|n|^2}/\overline{|s|^2})$ then it is preferable to change the role of the two signals in the system by exchanging the power processor with the correlation processor and vise versa.

Under the second condition for making (34) and (35) equal to zero (namely when $\partial A/\partial\alpha_R$
$= 0$ and $\partial A/\partial\alpha_I = 0$ but $A \neq 0$) we notice that (37) can not be made to equal zero unless β is real. Similarly, (37) can not be made to equal zero unless β is purely imaginary. In these particular cases

$$\alpha_R = -\frac{1}{2\beta_R} - \frac{c_R\overline{|s|^2} + b_R\overline{|n|^2}}{\overline{|s|^2} + |b|^2\overline{|n|^2}} \qquad \beta \textit{ real} \qquad (43)$$

$$\alpha_I = -\frac{1}{\beta_I} - \frac{c_I\overline{|s|^2} - b_R\overline{|n|^2}}{\overline{|s|^2} + |b|^2\overline{|n|^2}} \qquad \beta \textit{ purely imaginary} \qquad (44)$$

respectively. It is shown in Appendix A-2 that these values of α_R and α_I are points of maximum for Q_2. Furthermore they will not occur in practice since $|\beta|$ is constrained to be small just as $|\alpha|$.

In conclusion, we observe that in the power-correlation scheme the optimal weight

(β_{op}) for the power processor must satisfy (30), while α_{op} for the correlation processor can be obtained solely by (30) provided certain design conditions and constraints are met. These conditions, can easily be met in practical cross polarization problems.

2.2.3 The Correlation-Correlation Scheme

From Figure 4 it is quite obvious that without the signals' discrimination networks, the control loops are indistinguishable. Notice that

$$v_p(t) = v_1(t) + \beta v_2(t) \tag{45}$$

$$v_q(t) = \alpha v_1(t) + v_2(t) \tag{46}$$

and by using (1) we get

$$P_1 = \left|\overline{v_{p1}(t)v_q(t)}\right|^2 = |A_1|^2 \tag{47}$$

where

$$A_1 = (b+\beta)(1+b\alpha)^*(\delta_{n1})^2\overline{|n|^2} + (c+\alpha)^*(1+c\beta)(\delta_{s1})^2\overline{|s|^2} \tag{47a}$$

and, $\delta_{n1} = n_1/n$ and $\delta_{s1} = s_1/s$, represent the effect of the discriminator networks.[5] Similarly

$$Q_2 = |\overline{v_{q2}(t)v_p(t)}|^2 = |A_2|^2 \tag{48}$$

where

$$A_2 = (b+\beta)^*(1+b\alpha)(\delta_{n2})^2\overline{|n|^2} + (c+\alpha)(1+c\beta)^*(\delta_{s2})^2\overline{|s|^2} \tag{48a}$$

and, $\delta_{n2} = n_2/n$ and $\delta_{s2} = s_2/s$. In the derivation of (47a) and (48a) we used the zero correlation property of $s(t)$ and $n(t)$. To derive the optimal weights, α_{op} and β_{op} we follow the same steps used in obtaining α_{op} of the correlation processor in the

[5] For simplicity the effect of discriminator was assumed to cause different real gain for the different signals.

previous scheme. That is, to find β_{op} of the correlation-correlation scheme, we must equate $\partial P_1/\partial\beta_I = 0$ and $\partial P_1/\partial\beta_R = 0$ simultaneously to zero. In Appendix A-3 we prove that $\partial P_1/\partial\beta_R = 0$ and $\partial P_1/\partial\beta_I = 0$ if and only if $A_1 = 0$. This implies, using (47a),

$$\beta_{op} = -\frac{b(1+b\alpha)^*(\delta_{n1})^2\overline{|n|^2} + (c+\alpha)^*(\delta_{s1})^2\overline{|s|^2}}{(1+b\alpha)^*(\delta_{n1})^2\overline{|n|^2} + c(c+\alpha)^*(\delta_{s1})^2\overline{|s|^2}} \tag{49}$$

Notice that if $\alpha = -c$ then $\beta_{op} = -b$ and we get a perfect cancellation of the n signal at $v_p(t)$. We can use a similar argument to find α_{op} from equating, simultaneously, $\partial Q_2/\partial\alpha_R$ and $\partial Q_2/\partial\alpha_I$ to zero, namely

$$\alpha_{op} = -\frac{c(1+c\beta)^*(\delta_{s2})^2\overline{|s|^2} + (b+\beta)^*(\delta_{n2})^2\overline{|n|^2}}{(1+c\beta)^*(\delta_{s2})^2\overline{|s|^2} + b(b+\beta)^*(\delta_{n2})^2\overline{|n|^2}} \tag{50}$$

Again if $\beta = -b$ then $\alpha_{op} = -c$ and we get a perfect cancellation of the s signal at $v_q(t)$. Equations (49)and (50) are the same as (20) and (23), respectively, which represent the optimal weights for the power-power scheme. Recall that the latter were obtained with the rather reasonable condition that $|1-\alpha\beta| \neq 0$.

To conclude this section, we introduce in Table 1 a summary of the equations representing the optimal weights of the three configurations together with the conditions under which these weights are attainable.

2.3 The Optimal Output Signals' Power and Power Ratios

Using the optimal weights derived in the previous section, we now find the optimal powers of the two signals at the two different output ports. This we do for the different cross-pol cancelers schemes by implementing the corresponding optimal weights.

2.3.1 The Power-Power Scheme

From (20) we get

$$1 + c\beta = \frac{(1+\alpha b)^*(1-cb)\overline{|n_1|^2}/\overline{|s_1|^2}}{c(\alpha+c)^* + (1+\alpha b)^*\overline{|n_1|^2}/\overline{|s_1|^2}} \tag{51}$$

and

$$b + \beta = \frac{(\alpha+c)^*(1-cb)}{c(\alpha+c)^* + (1+\alpha b)^*\overline{|n_1|^2}/\overline{|s_1|^2}} \tag{52}$$

We also introduce in (51) and (52) the effect of the discriminator network No.1 (see Figure 2). From (17) we get, by using (51), the s signal's power

$$P_s = \frac{1}{|1-\alpha\beta|^2}\left|\frac{(1+\alpha b)^*(1-cb)\overline{|n_1|^2}/\overline{|s_1|^2}}{c(\alpha+c)^* + (1+\alpha b)^*\overline{|n_1|^2}/\overline{|s_1|^2}}\right|^2 \overline{|s|^2} \tag{53}$$

and by using (52), the n signal's power

$$P_n = \frac{1}{|1-\alpha\beta|^2}\left|\frac{(\alpha+c)^*(1-cb)}{c(\alpha+c)^* + (1+\alpha b)^*\overline{|n_1|^2}/\overline{|s_1|^2}}\right|^2 \overline{|n|^2} \tag{54}$$

The signals' power ratio at this port is given by

$$P_s/P_n = \left|\frac{1+\alpha b}{\alpha+c}\right|^2 (\delta_{n1}/\delta_{s1})^2 \overline{|n|^2}/\overline{|s|^2} \tag{55}$$

where $\delta_{n1} = n_1/n$ and $\delta_{s1} = s_1/s$. From (55) we notice that the power-inversion relation has been enhanced, as a result of the bootstrapping arrangement, by the factor $|(1+\alpha b)/(\alpha+c)|^2$, which approaches infinity (perfect cancellation) if $\alpha = -c$. Further enhancement of the s-to-n signals' power ratio prevails if the discriminator network is chosen such that $\delta_{s1} < \delta_{n1}$. With the smallness condition assumed for $|b|$ and $|c|$ and the constraint imposed on $|\alpha|$, we can easily show that the input-output transmission for the s signal is approximately unity

$$P_s/\overline{|s|^2} \simeq 1 \tag{56}$$

provided that $|\alpha + c|$ is very small so that $|\beta + b|$ is approximately equal to zero. Similarly the input-output cancellation factor is

$$P_n/\overline{|n|^2} \simeq |\alpha + c|^2/(\overline{|n_1|^2}/\overline{|s_1|^2})^2 \tag{57}$$

Notice the improvement in cancellation if $\delta_{s1} < \delta_{n1}$.

Similar argument leads, for the second part, to

$$Q_n/Q_s = \left|\frac{1+\beta c}{\beta + b}\right|^2 (\delta_{s2}/\delta_{n2})^2 \overline{|s|^2}/\overline{|n|^2} \tag{58}$$

as the n-signal to s-signal ratio. Also the input-output transmission for the n-signal

$$Q_n/\overline{|n|^2} \simeq 1 \tag{59}$$

provided $|\beta + b|$ is very small so that $|\alpha + c|$ is approximately zero. The input-output cancellation factor of the s-signal,

$$Q_s/\overline{|s|^2} \simeq |\beta + b|^2/(\overline{|s_s|^2}/\overline{|n_2|^2})^2 \tag{60}$$

2.3.2 The Power-Correlation Scheme

Using the value of β_{op} (Equation 30) we obtain from (28) the optimal output powers at the P port. The s-signal optimal power,

$$P_s = \left|\frac{(1+\alpha b)^*(1-bc)\overline{|n_1|^2}/\overline{|s_1|^2}}{|\alpha + c|^2 + |1+\alpha b|^2 \overline{|n_1|^2}/\overline{|s_1|^2}}\right|^2 \overline{|s|^2} \tag{61}$$

where we also introduce the effect of the discriminator network on (30)
The other signal optimal power.

$$P_n = \left|\frac{(\alpha + c)^*(1-cb)}{|\alpha + c|^2 + |1+\alpha b|^2 \overline{|n_1|^2}/\overline{|s_1|^2}}\right|^2 \overline{|n|^2} \tag{62}$$

The signals' power ratio becomes

$$P_s/P_n = \left|\frac{1+\alpha b}{\alpha + c}\right|^2 (\delta_{n1}/\delta_{s1})^2 \overline{|n|^2}/\overline{|s|^2} \tag{63}$$

Furthermore, with the smallness condition assumed for the system inputs and parameters, $|\alpha + c|^2 \ll |1 + \alpha b|^2 \overline{|n_1|^2}/\overline{|s_1|^2}$, we have for the input-output transmission of the s-signal,

$$P_s/\overline{|s|^2} = \left|\frac{1 - bc}{1 + \alpha b}\right|^2 \simeq 1 \tag{64}$$

and the input-output cancellation factor of the n-signal is

$$P_n/\overline{|n|^2} \simeq \left|\frac{(\alpha + c)^*(1 - cb)}{|1 + \alpha b|^2}\right|^2 /(\overline{|n_1|^2}/\overline{|s_1|^2})^2$$

$$\simeq |\alpha + c|^2/(\overline{|n_1|^2}/\overline{|s_1|^2})^2 \tag{65}$$

The approximations made in (64) and (65) are true if we further assume, $|c| = |b| \ll 1$.

At the Q port of the correlator processor, the output power is obtained by substituting the optimal weight α_{op} from (39) we get for signals' power ratio

$$Q_n/Q_s = \left|\frac{1}{b + \beta(1 + \alpha_{op} b)}\right|^2 (\delta_{s2}/\delta_{n2})^2 \overline{|s|^2}/\overline{|n|^2} \tag{66}$$

and for the input-output transmission of the n-signal;

$$Q_n/\overline{|n|^2} = |1 - bc|^2 \simeq 1 \tag{67}$$

provided $|b| = |c| \ll 1$. The cancellation factor of the s-signal is

$$Q_s/\overline{|s|^2} = |b + \beta(1 + \alpha_{op} b)|^2/(\overline{|s_2|^2}/\overline{|n_2|^2})^2 \tag{68}$$

Notice that the output power ratio P_s/P_n (63) is the same as that obtained with the power-power scheme (55). Comparing the other port power ratio Q_n/Q_s (Equations (58)and (66)), we notice a slight difference which is a direct consequence of the fact that (39) is a second order equation in α, and that the steady state optimal weight of the β processor for the power-correlation scheme is not exactly the same as that of the β processor for the power-power scheme (compare (20) with (30)). Similar results are obtained when comparing the input-output transmission for one signal of the

input-output cancellation factor to the other signal (see (56), (57), (59) and (60) to the power-power scheme and (64), (65), (67), (75) to the power-correlation scheme). Notice, particularly, the slight difference that resulted from having $b + \beta(1 + \alpha_{op}b)$ (in (68)) for the power-correlation instead of $b + \beta$ (in (60)) for the power-power. However, both terms were found to reach an ideal zero value if the corresponding α processor approaches its ideal optimal value. Small differences might occur, however, because the different scheme's structure can result in different final ideal values of these terms.

2.3.3 The Correlation-Correlation Scheme

From Figure 4, we find the output power,

$$P = \overline{|v_p(t)|^2} = |b + \beta|^2\overline{|n|^2} + |1 + c\beta|^2\overline{|s|^2} \tag{69}$$

Except for the $|1 - \alpha\beta|^2$ factor, this is exactly the same as Equation (12) which represents the output power at the P-port of the power-power scheme. We have also shown in Section (2.3) that the optimal weights for the correlation-correlation scheme are the same as those for the power-power. Therefore. we expect to have the same terms (except for the division by $(1-\alpha\beta)^2$) for the s-signal output power P_s (Equation 53), the n-signal output power P_n (Equation 54) and $s - to - n$ power ratio P_s/P_n (Equation 55) at the P port and similarly for $Q_n.Q_s$ and Q_n/Q_s at the Q port. The same applies for the input-output transmissions and cancellation factors (Equations (56), (57), (59) and (60)).

In Table 2, we summarize the equations representing the s-signal and n-signal optimal output powers at the P and Q port of the three different configurations. In Table 3, we summarize the equations representing the power ratios P_s/P_n and Q_s/Q_n at the two ports, respectively. Also represented in Table 3 are the input-output transmission ratios and cancellation factors for the different configurations and the

separate ports.

3 Conclusion and Further Work

After demonstrating the fact that the simple "noise cancelling scheme" proposed by Widrow [1] cannot be used in the cross polarization problem, the bootstrapping approach was proposed. Three different optimization criterian, namely power-power, power-correlation and correlation-correlation criterian were considered. Correspondingly, three bootstrapping configurations were examined. For each, the optimal complex weights were found together with the conditions under which these optimal weights are the unique solutions for the optimization problem and hence will be attainable if an adaptive algorithm (such as the steepest decent) is used. Consequently, equations for the optimal signal power outputs and signal-to-interference power ratios at the two different output ports were derived. The input-output signal transmission and cancellation factors were calculated. It was shown that any of the bootstrapping arrangements achieves a power separation (high signal-to-interference ratio at both output ports). However, for a power separation a discriminator technique (which depends on the signals' properties and can take the form of a filter, limiter, etc.) must be used. The need for such discrimination networks was explained and their effect on obtaining unique steady state solutions was exhibited.

To conclude, we make the following remark in relation to performance comparison of the three bootstrapping configurations considered in this appendix: while the symmetric power-power and correlation-correlation schemes produce the same output signal powers and power ratios, the asymmetric power-correlation scheme has a slightly different signal power and power ratio at its Q port. These differences are only of a second order and might exist only if, due to some system or input impurities (Gaussian noise. quantization error, etc.), the ideal optimal conditions are not

reached. The same conclusion is in effect when comparing the input-output transmission or cancellation factors.

Examination of Figures 2 through 4 reveals the fact that the three configurations proposed require different levels of hardware complexity; some need correlators (harder to implement) others require power measurements. The correlation-correlation scheme is expected to be the most complex while the power-power scheme the least. Also different signal paths (through the system's circuitry) are expected to result in different system delays with the different configurations and hence, different bandwidth limitations. The question of a possible trade-off between complexity and bandwidth is raised. This will be addressed in a subsequent appendix.

Gaussian noise effects on the system performance and the analytical results of this appendix should also be considered. In particular it is important to investigate the limit on cancellation depth and hence on the quality of power separation that this noise might cause.

The above items are questions which are obviously related to the subject of this appendix, namely steady state analysis, but they will be reported separately in appendix B.

Finally, the whole subject of dynamic analysis must be investigated for the three configurations.

References

[1] B.Widrow et al, "Adaptive Noise Cancelling: Principles and Application," Proc. IEEE, Vol.63, No.12, December 1975, pp. 1692-1716.

[2] R.T.Compton. "The Power-Inversion Array: Concept and Performance," IEEE Trans.
on Aerosp. Electron. Syst. Vol. AES-15, pp. 803-814.

[3] D.H.Brandwood, "Cross Coupled Cancellation System for Improving Cross-Polarization Discrimination," Proc. Int. Conf. on Ant. and Prop. Part I, London, England. November 1978, pp41-5.

[4] Y.Bar-Ness and J. Rokah "Cross Coupled Bootstrapped Interference Canceler," Proc. Int.
Conf. on Ant. and Prop., Los Angeles, CA. June 1981, Vol. I, pp 292-295.

[5] Y.Bar-Ness, J.W. Carlin, and M.L. Steinberger, "Bootstrapping Adaptive Cross Pol Cancelers for Satellite Communications," Proc. Int. Conf. on Comm., Philadelphia, PA. June, 1982, paper no. 4F5.

[6] C.Jutten and J.Herault, "Blind separation of Source Part I" Signal Processing, Vol 245 No.1 July 199, pp 1-10.

[7] E.Weinstein, M.Fedor and A.V.Oppenhein. "Multi-Channel signal Separation Based on De-correlation", submitted to IEEE Trans. on ASSP 1992.

[8] M.L. Steinberger, "Adaptive Cross-Polarization Interference Cancellation Arrangements," U.S. parent 4,283,795, issued August 11, 1981.

[9] J.Carlin. Y.Bar-Ness and W.E.Studdiford "An IF Cross-Pol Canceler for Microwave Radio" Journal on Selected Area in Communication-Advances in Digital Communication by Radio, Vol. SAC-5, No.3. April 87,pp. 502-514.

[10] Y.Bar-Ness "Bootstrapped Algorithm for Cancellation" Proc. AFCEA-IEEE Technical Conf. on Tactical Communications. Fort Wayne, IN, May, 1988, pp.

[11] B.Widrow. J. McCool and M. Ball, "The Complex LMS Algorithm," IEEE Proc., April 1975. pp. 719-720.

IV. APPENDICES

A-1

For any function of a complex variable $f(\beta)$ we have,

$$E[f(\beta)f^*(\beta)] = E[f_R^2(\beta) + f_I^2(\beta)] = E[f_R^2(\beta)] + E[f_I^2(\beta)]$$

where R and I designate the real and imaginary parts, respectively. Since the two components of $f(\beta)$ are in quadrature relation to each other, they can not be minimized independently.

$$\frac{\partial}{\partial \beta_R} E[f(\beta)f^*(\beta)] = E[f(\beta)\frac{\partial f^*(\beta)}{\partial \beta_R}] + E[f^*(\beta)\frac{\partial f(\beta)}{\partial \beta_R}] \tag{1}$$

$$\frac{\partial}{\partial \beta_I} E[f(\beta)f^*(\beta)] = E[f(\beta)\frac{\partial f^*(\beta)}{\partial \beta_I}] + E[f^*(\beta)\frac{\partial f(\beta)}{\partial \beta_I}] \tag{2}$$

Accordingly, we define,

$$\frac{d}{d\beta} E[f(\beta)f^*(\beta)] \equiv \frac{\partial}{\partial \beta_R} E[f(\beta)f^*(\beta)] + j\frac{\partial}{\partial \beta_I} E[f(\beta)f^*(\beta)] \tag{3}$$

if $f(\beta)$ is linear in β

$$f(\beta) = a_1 + a_2\beta$$

then

$$\frac{\partial}{\partial \beta_R} E[f(\beta)f^*(\beta)] = E[f(\beta)a_2^*] + E[f^*(\beta)a_2]$$

and

$$\frac{\partial}{\partial \beta_I} E[f(\beta)f^*(\beta)] = E[-jf(\beta)a_2^*] + E[jf^*(\beta)a_2]$$

Using (P-3) we get

$$\frac{d}{d\beta} E[f(\beta)f^*(\beta)] = 2E[a_2^* f(\beta)]$$

This is the regular complex algorithm of Widrow [11]. This is also what we termed the general rule for a complex derivative.

Notice that we can write in an "informal form"

$$\frac{d}{d\beta}E[f(\beta)f^*(\beta)] = E[f(\beta)\frac{df^*(\beta)}{d\beta}] + E[f^*(\beta)\frac{df(\beta)}{d\beta}] \tag{4}$$

This is an informal since $df(\beta)/d\beta$ has not yet been defined. If $f(\beta)$ is a general function of β then we must use (P-1), (P-2) together with (P-3) to obtain.

$$\frac{dE[f(\beta)f^*(\beta)]}{d\beta} = E\left(f(\beta)[\frac{\partial f^*(\beta)}{\partial \beta_R} + j\frac{\partial f^*(\beta)}{\partial \beta_I}]\right) + E\left(f^*(\beta)[\frac{\partial f(\beta)}{\partial \beta_R} + j\frac{\partial f(\beta)}{\partial \beta_I}]\right)$$

from which (P-4) results if we define "informal" that for any function $f(\beta)$

$$\frac{df(\beta)}{d\beta} = \frac{\partial f(\beta)}{\partial \beta_R} + j\frac{\partial f(\beta)}{\partial \beta_I}$$

Notice that $\partial f(\beta)/\partial\beta$ is not a complex conjugate of the function $\partial f^*(\beta)/\partial\beta$. Also for $f(\beta)$ linear in β $\frac{df(\beta)}{d\beta} = 0$

A-2

To check the nature of the extremum points defined in (43), for β real, one can show from (48) and (48a) that

$$\partial^2 Q_2/\partial\alpha_R^2 = 2Real[A_2\partial^2 A_2^*/\partial\alpha_R^2] \tag{5}$$

and

$$\partial^2 A_2/\partial\alpha_R^2 = 2\beta(\overline{|s|^2} + |b|^2\overline{|n|^2})$$

Hence

$$\partial^2 Q_2/\partial\alpha_R^2 = 4Real[A\beta](\overline{|s|^2} + |b|^2\overline{|n|^2})$$

$$= 4(\overline{|s|^2} + |b|^2\overline{|n|^2})$$

$$[\beta_R^2(|\alpha + c|^2\overline{|s|^2} + |1 + \alpha b|^2\overline{|n|^2}) + \beta_R((\alpha + c)_R\overline{|s|^2} + (b_R + \alpha_R|b|^2)\overline{|n|^2})]$$

(P-6)

Using (43), (P-6) yields,

$$\partial^2 Q_2/\partial\alpha_R^2 = 4(\overline{|s|^2} + |b|^2\overline{|n|^2})$$

$$[\beta_R^2(|\alpha - c|^2\overline{|s|^2} + |1 + \alpha b|^2\overline{|n|^2}) - \frac{1}{2}(\overline{|s|^2} + |b|^2\overline{|n|^2})] \qquad (7)$$

With the smallness condition assumed for b and c and the constraint imposed on α and β, $\partial^2 Q_2/\partial\alpha_R^2 < 0$ and the point defined by (43) is a point of maximum. Similar steps can be used to show that $\partial^2 Q/\partial\beta_I^2 < 0$ and the point defined by (44) with $\beta = j\beta_I$ (pure imaginary) is a point of maximum.

A-3

Directly from (P-1) and (P-2) and similar to (36) and (37), we have $\partial P_1/\partial\beta_R = 2Real$

$[A_1^*\partial A_1/\partial\beta_R]$ and $\partial P_1/\partial\beta_I = 2Real[A_1^*\partial A_1/\partial\beta_I]$. From (47a) we get

$$\partial A_1/\partial\beta_R = (1 - b\alpha)^*(\delta_{n1})^2\overline{|n|^2} + c(c + \alpha)^*(\delta_{s1})^2\overline{|s|^2} \qquad (8)$$

$$\partial A_1/\partial\beta_I = j(1 + b\alpha)^*(\delta_{n1})^2\overline{|n|^2} + jc(c + \alpha)^*(\delta_{s1})^2\overline{|s|^2} = j\partial A_1/\partial\beta_R \qquad (9)$$

Using (P-9) we also have

$$Real[A_1^*\partial A_1/\partial\beta_I] = -I_m[A_1^*\partial A_1/\partial\beta_R] \qquad (10)$$

Therefore $\partial P_1/\partial\beta_R = 0$ and $\partial P_1/\partial\beta_I = 0$. simultaneously, if and only if

$$[A_1^*\partial A_1/\partial\beta_R] = 0 \qquad (11)$$

Notice that $\partial A_1/\partial\beta_R(\partial A_1/\partial\beta_I)$ is independent of $\beta_R(\beta_I)$, therefore A_1 is independent of β if $\partial A_1/\partial\beta_R = 0(\partial A_1/\partial\beta_I = 0)$. But for $\partial A_1/\partial\beta_R = 0$ we must have

$$\alpha = -\frac{|c|^2(\delta_{s1}^2\overline{|s|^2} + (\delta_{n1}^2\overline{|n|^2}}{c^*(\delta_{s1}^2\overline{|s|^2} + b(\delta_{n1})^2\overline{|n|^2}}$$

which will not occur in practice if we assume $|c|$ and $|b|$ are much smaller than unity and restrict $|\alpha|$ to be less than unity. Therefore, $\partial A_1/\partial\beta_R$ or $\partial A_1/\partial\beta_I = 0$ and $A_1 \neq 0$ do not constitute an extremum point. so that the only point of extremum for P_1 is that which is obtained by equating A_1 to zero.

Table. 1

	β_{op}	α_{op}	condition
The Power Power Scheme	$-\frac{(\alpha+c)^*\overline{\lvert s_1\rvert^2}+b(1+\alpha b)^*\overline{\lvert n_1\rvert^2}}{c(\alpha+c)^*\overline{\lvert s_1\rvert^2}+(1+\alpha b)^*\overline{\lvert n_1\rvert^2}}$	$-\frac{(b+\beta)^*\overline{\lvert n_2\rvert^2}+c(1+c\beta)^*\overline{\lvert s_2\rvert^2}}{b(b+\beta)^*\overline{\lvert n_2\rvert^2}+(1+c\beta)^*\overline{\lvert s_2\rvert^2}}$	$\lvert 1-\alpha\beta\rvert \neq 0$ $\lvert\alpha\rvert < 1$ $\lvert\beta\rvert < 1$ $\lvert b\rvert, \lvert c\rvert << 1$
The Power-Correlation Scheme	$-\frac{(\alpha+c)^*\overline{\lvert s_1\rvert^2}+b(1+\alpha b)^*\overline{\lvert n_1\rvert^2}}{c(\alpha+c)^*\overline{\lvert s_1\rvert^2}+(1+\alpha b)^*\overline{\lvert n_1\rvert^2}}$	$-c-(1+\alpha_{op}\beta)(b+\beta(1+\alpha_{op}b))^*\frac{\overline{\lvert n_2\rvert^2}}{\overline{\lvert s_2\rvert^2}}$	Eqn. (42) or $\lvert b\rvert = \lvert c\rvert \leq 0.295$ $\lvert\beta_R\rvert_{max} = \lvert\beta_I\rvert_{max} < 0.326$ $\lvert\alpha\rvert < 1$
The Correlation-Correlation	same as in the power-power scheme	same as in the power-power scheme	$\lvert\alpha\rvert < 1$ $\lvert\beta\rvert < 1$ $\lvert b\rvert, \lvert c\rvert << 1$

Table 1. Summary of equations representing the optimal weight of the three configuration and conditions under which these optimal weights are attainable. $s_1 = s\delta_1, s_2 = s\delta_{s2}, n_1 = n\delta_{n1}$, and $n_2 = n\delta_{n2}$; where δ_s, and δ_n represent the effect of the discriminitor networks.

		s-signal	n-signal
The Power-Power Scheme	P port	$\frac{1}{\lvert 1-\alpha\beta\rvert^2}\left\lvert\frac{(1+\alpha b)^*(1-cb)\lvert n_1\rvert^2/\lvert s_1\rvert^2}{c(\alpha+c)^*+(1+\alpha b)^*\lvert n_1\rvert^2/\lvert s_1\rvert^2}\right\rvert^2\lvert s\rvert^2$	$\frac{1}{\lvert 1-\alpha\beta\rvert^2}\left\lvert\frac{(\alpha+c)^*(1-cb)}{c(\alpha+c)^*+(1+\alpha b)^*\lvert n_1\rvert^2/\lvert s_1\rvert^2}\right\rvert^2\lvert n\rvert^2$
	Q port	$\frac{1}{\lvert 1-\alpha\beta\rvert^2}\left\lvert\frac{(\beta+b)^*(1-cb)}{b(\beta+b)^*+(1+\beta c)^*\lvert s_2\rvert^2/\lvert n_2\rvert^2}\right\rvert^2\lvert n\rvert^2$	$\frac{1}{\lvert 1-\alpha\beta\rvert^2}\left\lvert\frac{(1+\beta c)^*(1-bc)\lvert s_2\rvert^2/\lvert n_2\rvert^2}{b(\beta+b)^*+(1+\beta c)^*\lvert s_2\rvert^2/\lvert n_2\rvert^2}\right\rvert^2\lvert n\rvert^2$
The Power-Correlation Scheme	P port	$\left\lvert\frac{(1+\alpha b)^*(1-bc)\lvert n_1\rvert^2/\lvert s_1\rvert^2}{\lvert\alpha+c\rvert^2+\lvert 1+\alpha b\rvert^2\lvert n_1\rvert^2/\lvert s_1\rvert^2}\right\rvert^2\lvert s\rvert^2$	$\left\lvert\frac{(\alpha+c)^*(1-bc)}{\lvert\alpha+c\rvert^2+\lvert 1+\alpha b\rvert^2\lvert n_1\rvert^2/\lvert s_1\rvert^2}\right\rvert^2\lvert n\rvert^2$
	Q port	$\lvert(1+\alpha_{op}b)(b+\beta(1+\alpha_{op}b))\rvert^2\left[\lvert n_2\rvert^2/\lvert s_2\rvert^2\right]^2\lvert s\rvert^2$	$\left\lvert\frac{1-bc}{1+b(b+\beta(1+\alpha_{op}b))^*\lvert n_2\rvert^2/\lvert s_2\rvert^2}\right\rvert^2\lvert n\rvert^2$
The Correlation-Correlation Scheme		The same as in the Power-Power scheme except for the exclusion of the term $\frac{1}{\lvert 1-\alpha\beta\rvert^2}$	The same as in the Power-Power scheme except for the exclusion of the term $\frac{1}{\lvert 1-\alpha\beta\rvert^2}$

Table. 2

Table 2. Summary of equations representing the s-signal and n-signal optimal output powers at the P and Q ports of the three configurations.

		The Power-Power Scheme	The Power-Correlation Scheme	The Correlation-Correlation Scheme
P_{port}	p_s/p_n	$\left\|\frac{1+\alpha b}{\alpha+c}\right\|^2 (\delta_{n1}/\delta_{s1})^2 \overline{\|n\|^2}/\overline{\|s\|^2}$	$\left\|\frac{1+\alpha b}{\alpha+c}\right\|^2 (\delta_{n1}/\delta_{s1})^2 \overline{\|n\|^2}/\overline{\|s\|^2}$	
	$p_s/\overline{\|s\|^2}$	$\simeq 1$	$\left\|\frac{1-bc}{1+\alpha b}\right\|^2 \simeq 1$	
	$p_n/\overline{\|n\|^2}$	$\|\alpha+c\|^2/(\overline{\|n_1\|^2}/\overline{\|s_1\|^2})^2$	$= \left\|\frac{(\alpha+c)^*(1-cb)}{\|1+\alpha b\|^2}\right\|^2 /(\overline{\|n_1\|^2}/\overline{\|s_1\|^2})^2$ $\simeq \|\alpha+c\|^2/(\overline{\|n_1\|^2}/\overline{\|s_1\|^2})^2$	The same as in the power-power scheme
Q_{port}	Q_n/Q_s	$\left\|\frac{1+\beta c}{\beta+b}\right\|^2 (\delta_{s2}/\delta_{n2})^2 \overline{\|s\|^2}/\overline{\|n\|^2}$	$\left\|\frac{1}{b+\beta(1+\alpha_{op}b)}\right\|^2 (\delta_{s2}/\delta_{n2})^2 \overline{\|s\|^2}/\overline{\|n\|^2}$	
	$Q_n/\overline{\|n\|^2}$	$\simeq 1$	$\|1-bc\|^2 \simeq 1$	
	$Q_s/\overline{\|s\|^2}$	$\|\beta+b\|^2/(\overline{\|s_2\|^2}/\overline{\|n_2\|^2})^2$	$\|b+\beta(1+\alpha_{op}b)\|^2/(\overline{\|s_2\|^2}/\overline{\|n_2\|^2})^2$	

Table 3. Summary of equations representing the signals' power ratios ($P_s/P_n, Q_n/Q_s$), the input-output transmissions ($P_s/\overline{|s|^2}, Q_n/\overline{|n|^2}$) and the cancellation factors ($P_n/\overline{|n|^2}, Q_s/\overline{|s|^2}$) for the different configurations.

Table. 3

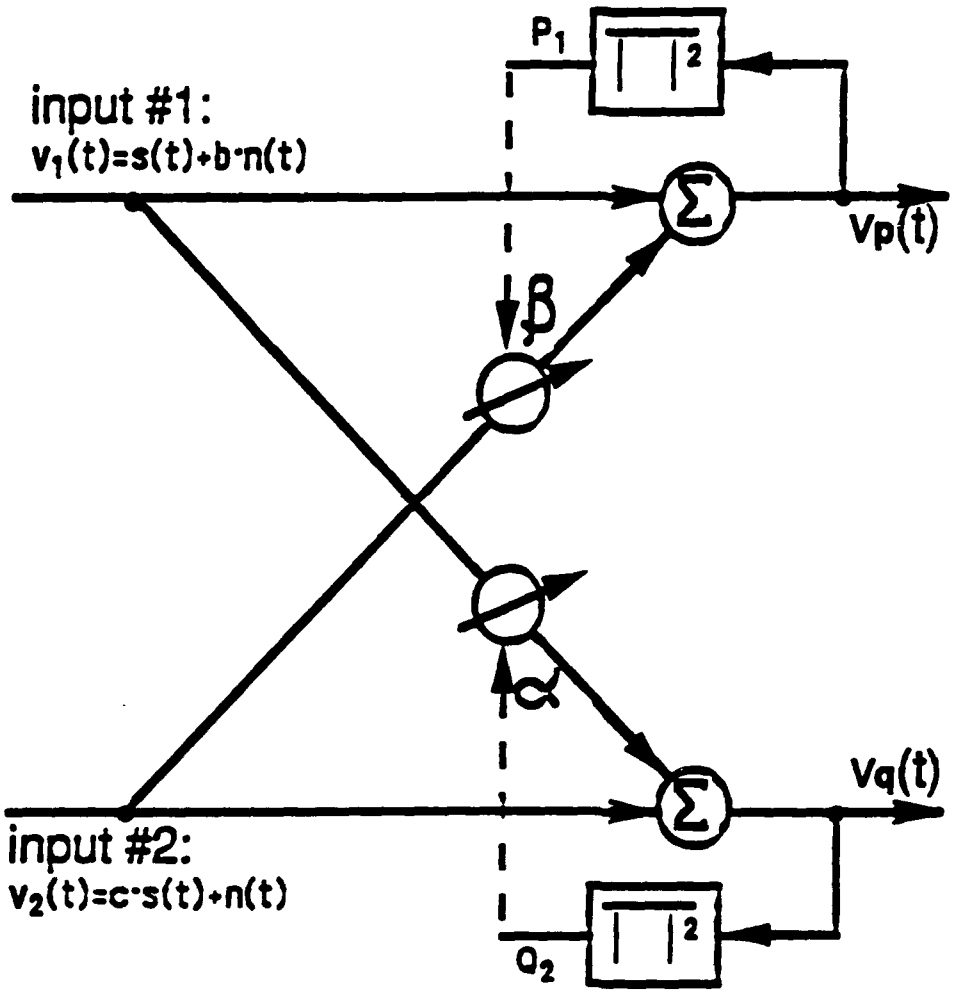


Fig. 1 Interference cancelling using two separate "noise cancelling" schemes

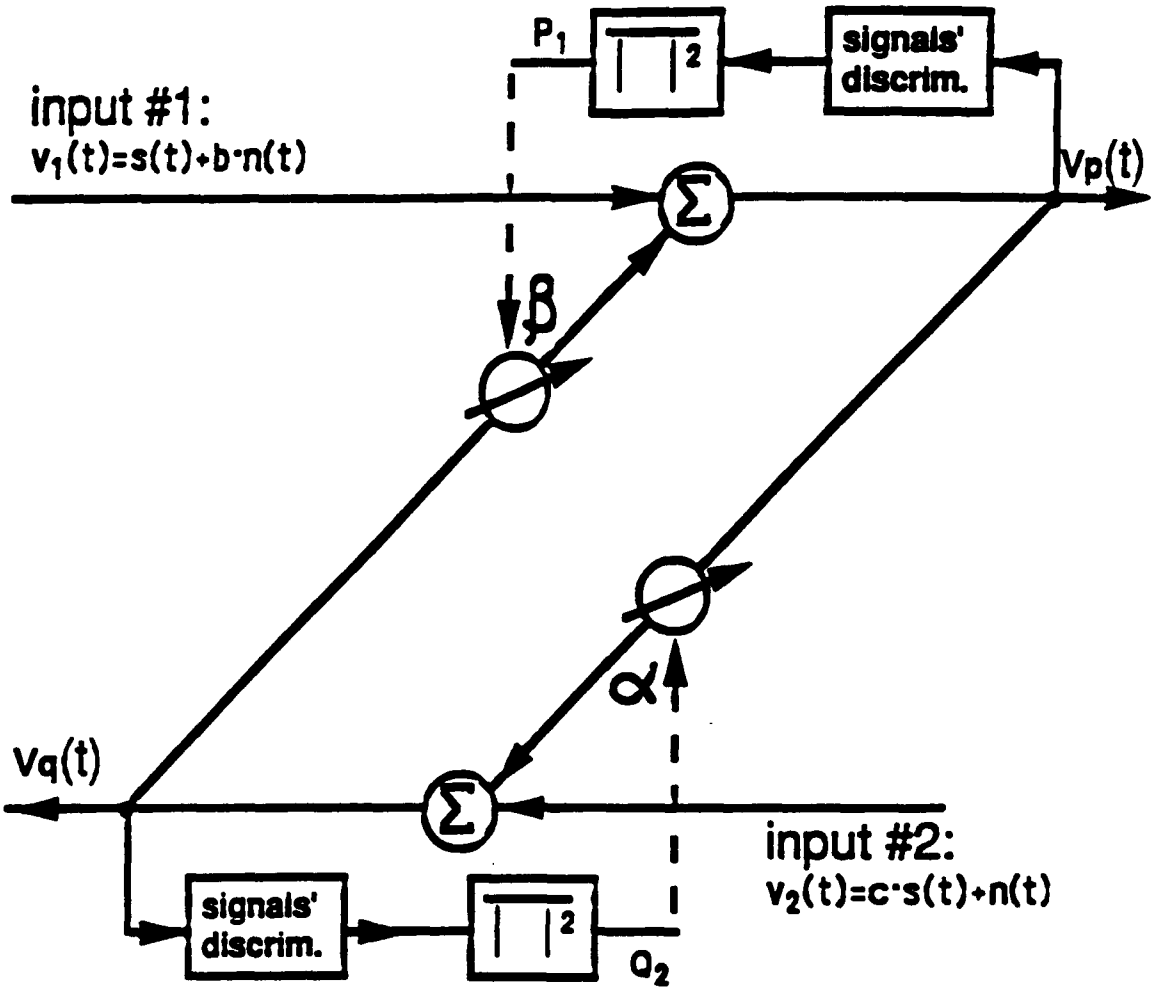


Fig. 2 Power-Power bootstrapped canceler

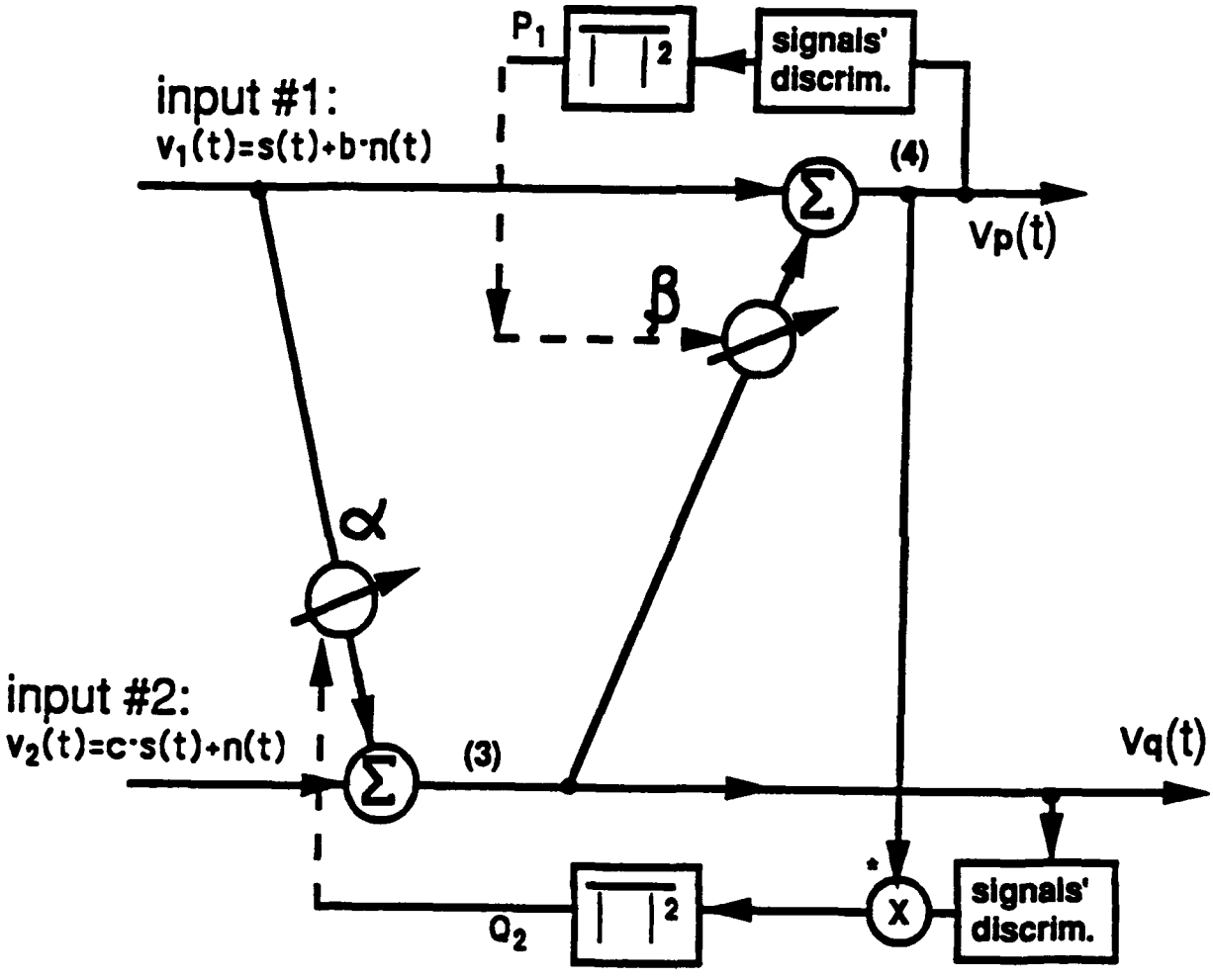


Fig. 3 Power-Correlation bootstrapped canceler

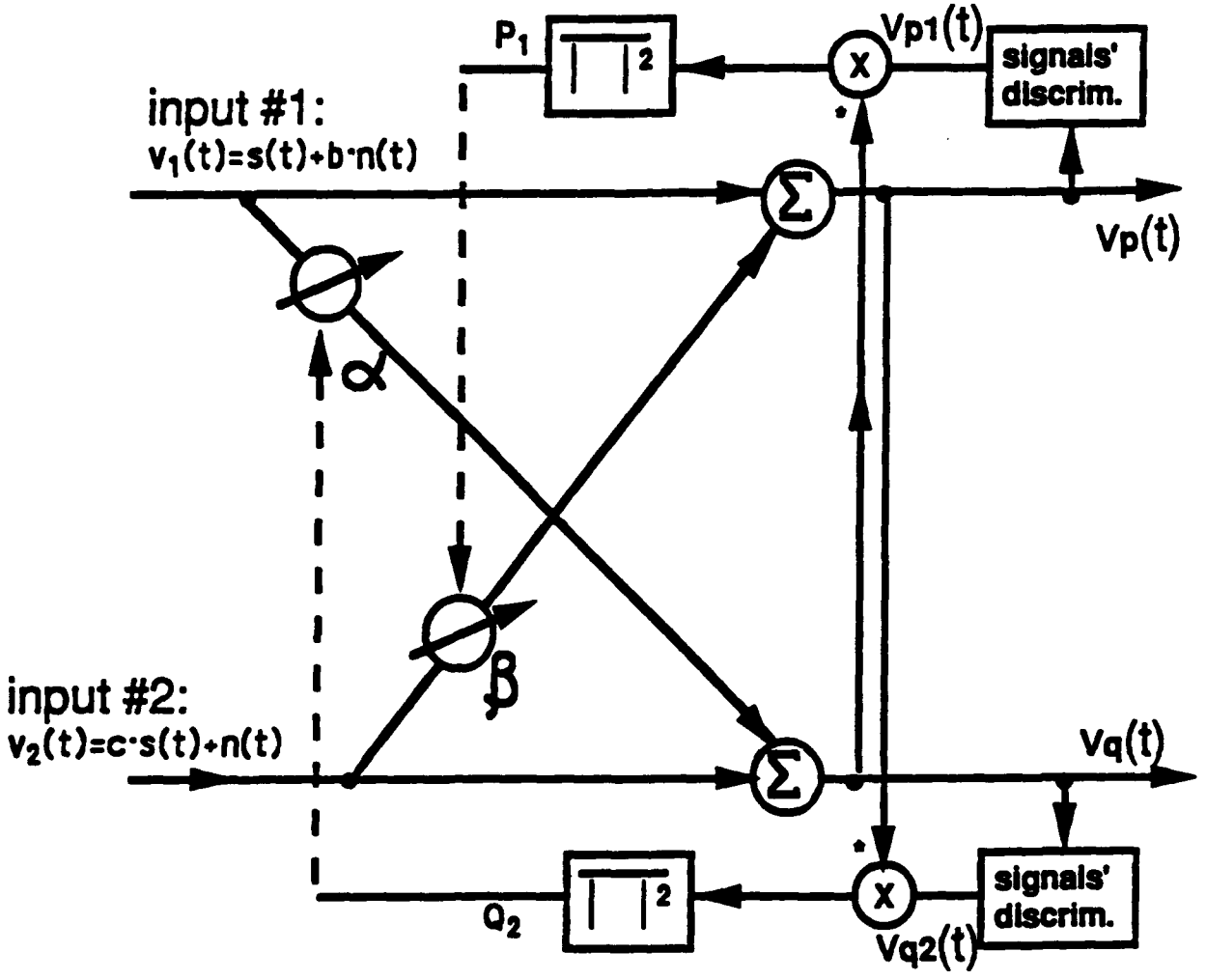


Fig. 4 Correlation-Cororrelation bootstrapped canceler

Appendix B

Bootstrapped Adaptive Separation of Superimposed Signals - Effect of Thermal Noise on Performance

Y. Bar-Ness, A. Dinç and H. Messer

Abstract

Bootstrapped adaptive signal separation techniques were shown to be powerful tools for separation of superimposed signals with an unknown parametric model. Estimation of all unknown model parameters are done simultaneously, using the estimated, separated signals under certain optimization criteria. When the signals to be separated are uncorrelated, random processes the algorithm converges to the unknown model parameters and the outputs are a clean version of the input signals, provided that no additive noise is present. In this appendix, we analyze the effect of additive thermal noise on the performance of the various structures of a bootstrapped signal separator. We show that this noise introduces *bias* to the model parameter estimates. We study this bias and we derive conditions under which its effect is limited. We provide a practical example where the separation of the input signals is hardly degraded due to the presence of additive thermal noise.

1 Introduction

Separation of superimposed signals is a major problem in many different fields. In fact, many problems are strongly related to the separation problem. For example, if a desired signal and an interference are considered as two signals to be separated, the well-known interference cancellation problem can be regarded as a signals' separation problem. In an interference cancellation application, however, one is usually not interested in the interference signal while in other applications all the signals to be separated are of interest. The general formulation of the problem is: **given $\mathbf{x}(t) = \mathbf{T}(\mathbf{s}(t), \theta) + \mathbf{n}(t)$, design a separation system G having the input $\mathbf{x}(t)$ and the output $\mathbf{y}(t) = \hat{\mathbf{s}}(t)$**; an estimate of $\mathbf{s}(t)$. $\mathbf{s}(t) = [s_1(t), ..., s_N(t)]^T$ is an N dimensional signal vector. $\mathbf{T}(\theta)$ is the mixing transformation which depends on the model parameter vector θ. $\mathbf{x}(t) = [x_1(t), ..., x_M(t)]^T$ is the M dimensional vector of the resulting superimposed signals, with

$$x_m(t) = \sum_{n=1}^{N} T_{mn}(\mathbf{s}(t), \theta) + n_m(t), \quad m = 1, 2, ..M \tag{1}$$

Also, **G** is the separation system and $\mathbf{y}(t)$ is the N dimensional vector of the separated signals. $\mathbf{n}(t) = [n_1(t), ..., n_M(t)]^T$ is an M dimensional vector of the additive noise (see Fig. 1). If the mixture model is linear in $\mathbf{s}(t)$ then

$$\mathbf{T}(\theta) = \begin{bmatrix} T_{11}(\theta) & \cdots & T_{1N}(\theta) \\ . & T_{ii}(\theta) & . \\ T_{M1}(\theta) & \cdots & T_{MN}(\theta) \end{bmatrix}, \tag{2}$$

In which case, it may take the form of a constant matrix, $\mathbf{T}(\theta) = T$, or a linear time invariant multi-input multi-output (MIMO) system; $T(t, \theta) = F^{-1}\{H(w.\theta)\}$ where $H(w, \theta)$ is an MxN transfer function matrix, or any non-linear operation which depends on a parameter vector θ.

If θ is known, then the problem is that of a multidimensional filtering. If the transformation is linear, its solution is well known. If $T(\theta)$ is a non-linear transformation,

the solution of the non-linear filtering problem depends on the type of non-linearity, but it can usually be found.

In practice, however, θ is rarely known. Thus, the separation problem is, inherently, a non-linear multidimensional filtering problem and is a very complicated one. Inspired by Widrow's adaptive LMS algorithm [1], Bar-Ness and Rockah [2] first suggested (for the two-inputs two-output case) to use each of the outputs of the separator as a reference input of an LMS algorithm which produces the other output. Since this arrangement successively improved the purity of the reference inputs and hence the corresponding outputs. Bar-Ness called his idea "a bootstrapped algorithm." The bootstrapped separator idea was extended to three different structures termed "power-power", "correlator-correlator" and "power-correlator" [3] and the results were reported in the open literature eg. [4-5]. The names of the different structures were chosen according to the performance indices which were used in each of these structures' optimal criteria. Application of the bootstrapped separators to satellite communication, microwave radio and tactical communication were reported in [4], [6] and [7], respectively. Independently, during the mid-eighties, a group of European researchers addressed the same separation problem and, under the name of "blind signal separation," developed a similar adaptive system [8]. Lately, a third group of signal processing researchers applied similar ideas to speech signal separation [9].

In this appendix, we review the three basic bootstrapped algorithm configurations of the separation schemes and relate them to the " blind separation " and " multichannel separation " schemes. We then analyze the steady-state performance of the different versions of the algorithms in the presence of an additive thermal noise. From the study of the very general case, we then move to a specific application - separation of cross polarized signals in satellite communication where we study the effect of the thermal noise on the separation performance in terms of the output desired-signal to

overall noise ratio.

For simplicity, our analysis is carried out for the case where M=N=2 (two complex signals to be separated, two complex measurements) and the mixture model is a constant, complex matrix. Notice, however, the same analysis *principles* are applicable to any ($M \geq N \geq 2$) and other mixture models.

That is, our model is

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{3}$$

where $\mathbf{s}(t) = [s_1(t)\, s_2(t)]^T$, $\mathbf{n}(t) = [n_1(t)\, n_2(t)]^T$, $\mathbf{A} = \begin{bmatrix} 1 & a_1 \\ a_2 & 1 \end{bmatrix}$, is the unknown complex matrix mixture model. (Notice that the assumption that $a_{11} = a_{22} = 1$ is not limiting in any way since the signals' power is assumed unknown). Thus $\mathbf{x}(t) = [x_1(t)\, x_2(t)]^T$ is complex, random vector. Given the vector $\mathbf{x}(t)$, separation is performed using a 2x2 complex matrix $\mathbf{W}$; that is

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) = \hat{s}(t) \tag{4}$$

In section II, we discuss three different structures for implementing $\mathbf{W}$: a backward/backward (BB) structure, a forward/forward (FF) structure and a forward/backward (FB) structure. For each structure, we propose different optimization criteria to be used for finding $\mathbf{W}$ that leads to separation, and we summarize the different algorithms which controls $\mathbf{W}$. Then, in section III, we analyze the steady state performance of the proposed algorithms in the presence of thermal additive noise. In section IV, we apply the results to the problem of cross polarization interference cancellation and show that in a practical situation, the effect of the noise on the separation performance is not dramatic.

2 Bootstrapped Signals Separator Configurations

2.1 The Backward-Backward Structure

In Fig. 2.a, cross-coupled feedback structure is depicted. w_1 and w_2 are complex weights. The input-output relation of the backward- backward configuration is summarized by

$$\mathbf{y}(t) = \mathbf{W}_{bb}\mathbf{x}(t) \tag{5}$$

where

$$\mathbf{W}_{bb} = \frac{1}{1 - w_1 w_2}\begin{bmatrix} 1 & -w_1 \\ -w_2 & 1 \end{bmatrix}, \tag{6}$$

By substituting (3) in (5) we get

$$\mathbf{y}(t) = \mathbf{W}_{bb}\mathbf{A}\mathbf{s}(t) + \mathbf{W}_{bb}\mathbf{n}(t) = \mathbf{H}_{bb}\mathbf{s}(t) + \mathbf{n}_{bb}(t) \tag{7}$$

where,

$$\mathbf{H}_{bb} = \frac{1}{1 - w_1 w_2}\begin{bmatrix} 1 - w_1 a_2 & a_1 - w_1 \\ a_2 - w_2 & 1 - w_2 a_1 \end{bmatrix}, \tag{8}$$

and $\mathbf{n}_{bb}(t) = \mathbf{W}_{bb}\mathbf{n}(t)$. Since signal separation is desired, w_1 and w_2 should be chosen such that $\mathbf{H}_{bb}$ is a diagonal matrix. Here, the desired solution is

$$w_1 = a_1, \quad w_2 = a_2, \tag{9}$$

for which $\mathbf{H}_{bb}$ =**I** (the identity matrix) and $\mathbf{y}(t) = \mathbf{s}(t) + \mathbf{A}^{-1}\mathbf{n}(t)$ is a noisy estimate of the signal vector (without noise $\mathbf{y}(t) = \mathbf{s}(t)$). Notice that (9) is a necessary and sufficient condition for $\mathbf{W}_{bb} = \mathbf{A}^{-1}$, so the weights can be regarded as estimates of a_1 and a_2. If $\mathbf{W}_{bb} \neq \mathbf{A}^{-1}$, the output correlation matrix is given by

$$\mathbf{R}_{bb} = E\{\mathbf{y}(t)\mathbf{y}^H(t)\} = \mathbf{H}_{bb}\mathbf{R}_s\mathbf{H}_{bb}^H + \mathbf{W}_{bb}\mathbf{R}_n\mathbf{W}_{bb}^H \tag{10}$$

where $\mathbf{R}_s = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$ and $\mathbf{R}_n = E\{\mathbf{n}(t)\mathbf{n}^H(t)\}$ are the signal and the noise covariance matrices, respectively, and the noise and signal vectors are assumed uncorrelated, H stands for Hermitian. We assume that

$$\mathbf{R}_n = \sigma^2\mathbf{I}, \quad \mathbf{R}_s = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \tag{11}$$

That is, the noise processes $n_1(t)$ and $n_2(t)$ are uncorrelated processes with power σ^2 and the signals to be separated: $s_1(t)$ and $s_2(t)$, are assumed uncorrelated, with power σ_1^2 and σ_2^2, respectively. By substituting (11), (8) and (6) in (10), we get

$$\begin{aligned} \mathbf{R}_{bb} &= E\{\mathbf{y}(t)\mathbf{y}^H(t)\} \\ &= \frac{1}{|1-w_1w_2|^2}\left(\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} + \sigma^2\begin{bmatrix} 1+|w_1|^2 & -(w_2^* + w_1) \\ -(w_2 + w_1^*) & 1+|w_2|^2 \end{bmatrix}\right), \end{aligned} \tag{12}$$

where

$$\begin{aligned} r_{11} &= \sigma_1^2|1-w_1a_2|^2 + \sigma_2^2|a_1 - w_1|^2 \\ r_{12} &= \sigma_1^2(a_2 - w_2)^*(1-w_1a_2) + \sigma_2^2(a_1-w_1)(1-w_2a_1)^* \\ r_{21} &= \sigma_1^2(a_2-w_2)(1-w_1a_2)^* + \sigma_2^2(a_1-w_1)^*(1-w_2a_1) = r_{12}^* \\ r_{22} &= \sigma_1^2|a_2-w_2|^2 + \sigma_2^2|1-w_2a_1|^2 \end{aligned} \tag{13}$$

The diagonal entries of R_{bb} are the output power, $E\{|y_1(t)|^2\}$ and $E\{|y_2(t)|^2\}$, respectively while the off-diagonal terms are the correlation between the two outputs, $E\{y_1(t)y_2^*(t)\}$ and $E\{y_1^*(t)y_2(t)\}$. In [3] it was shown that, for $\sigma^2 = 0$ (no thermal noise), the two outputs powers are minimal if and only if (9) is satisfied. Thus, simultaneous minimization of both output powers has been proposed as an optimization criterion, and the corresponding configuration was referred to as the "power-power configuration." Notice, however that, equating the derivative of the output power to zero is a necessary condition for minimization. That is

$$\begin{aligned} \frac{d}{dw_1}E\{|y_1(t)|^2\} &= E\{\frac{d}{dw_1}|x_1(t) - w_1y_2(t)|^2\} \\ &= E\{-2y_2^*(t)[x_1(t) - w_1y_2(t)]\} = -2E\{y_1(t)y_2^*(t)\} = 0 \end{aligned} \tag{14}$$

and similarly

$$\frac{d}{dw_2}E\{|y_2(t)|^2\} = -2E\{y_1^*(t)y_2(t)\} = 0 \quad (15)$$

where we used, from Fig. 2.a, the fact that, $y_1(t) = x_1(t) - w_1y_2(t)$; $y_2(t) = x_2(t) - w_2y_1(t)$. Although $E\{|y_1(t)|^2\}$ is not an analytical function, the derivatives in (14) and (15) are informally defined as follows [1]: if $A = |d + ex|^2$ and x is a complex variable then $\frac{dA}{dx} = 2e^*(d + ex)$.

From (14) and (15), we conclude that for minimum output powers. it is necessary to have minimum cross correlation between the outputs[6]. Thus, as optimization critera for the backward-backward (BB) configuration one can either use minimum output powers or de-correlation of the outputs. It can easily be shown that (14) leads to

$$\sigma_1^2(a_2 - w_{2o})^*(1 - w_{1o}a_2) + \sigma_2^2(a_1 - w_{1o})(1 - w_{2o}a_1)^* = 0 \quad (16)$$

while (15) leads to the complex conjugate of (16). Therefore, to find the optimal weights w_{1o} and w_{2o} we must either perform minimum power or de-correlation criterion by simultaneously solving the two dependent equations (14) and (15). To overcome the dependent equation problem, it was suggested in [2-3] and all the succeeding articles dealing with bootstrapped separators to add what was termed a "discriminator." This is a certain functional operation performed on the outputs of the separator and helps break the inherent dependency of the simultaneous equation to be solved. To choose an adequate discriminator. some designer ingenuity and prior knowledge of "some" differences between the signals to be separated is required. Examples of the particular discriminator, suitable for different application. were given in [4]. [6] and [7]. In [3], it was suggested to present the effect of the discriminator as follows: For the signal $s_1(t)$ and $s_2(t)$, the two discriminator outputs are given by

$$D_1(s_1(t) + s_2(t)) = s_1(t) + \delta_{12}s_2(t)$$

[6]it is also a sufficient condition since the power is quadratic function of the weight.

$$D_2(s_1(t) + s_2(t)) \;=\; \delta_{21} s_1(t) + s_2(t) \tag{17}$$

where $|\delta_{12}| < 1$, $|\delta_{21}| < 1$. That is, the discriminator D_i slightly emphasizes the signal $s_i(t)$ relative to the other signal, i=1,2.

Optimization Criteria and Respective Algorithms

Due to the equivalence cited above, we can use for optimization criteria either power-power minimization, correlation-correlation magnitude minimization or power-correlation magnitude minimization on the outputs. Notice that minimization of correlation magnitude leads to de-correlation.

Together with the effect of the discriminator. these criteria are given by [7]

$$\begin{aligned}
\textbf{power-power: } \mathbf{w}(t) &= \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} = \arg \min_{\mathbf{w}} \begin{bmatrix} E\{|D_1(y_1(t))|^2\} \\ E\{|D_2(y_2(t))|^2\} \end{bmatrix} \\
\textbf{correlation-correlation : } \mathbf{w}(t) &= \arg \min_{\mathbf{w}} \begin{bmatrix} |E\{D_1(y_1(t))y_2^*(t)\}|^2 \\ |E\{y_1(t)^* D_2(y_2(t))\}|^2 \end{bmatrix} \\
\textbf{power-correlation : } \mathbf{w}(t) &= \arg \min_{\mathbf{w}} \begin{bmatrix} E\{|D_1(y_1(t))|^2\} \\ |E\{y_1(t)^* D_2(y_2(t))\}|^2 \end{bmatrix} \\
\textbf{correlation-power : } \mathbf{w}(t) &= \arg \min_{\mathbf{w}} \begin{bmatrix} |E\{D_1(y_1(t))y_2^*(t)\}|^2 \\ E\{|D_2(y_2(t))|^2\} \end{bmatrix}
\end{aligned} \tag{18}$$

By using the steepest descent algorithm with the negative of the terms in (18) as the forcing increments, we get two linearly independent recursive control equations. For $\sigma^2 = 0$ (no noise) and for $|w_1| < 1$, and $|w_2| < 1$, these equations have a unique solution, $w_1 = a_1$ and $w_2 = a_2$ [3] and the bootstrapped separator converges to these values.

To summarize: with the backward-backward separation scheme, one can use four different adaptive separation algorithms. They are denoted by BBPP, BBCC, BBPC and BBCP where the first two letters (e.g. BB) stand for the configuration used (backward-backward) and the last two letters (say PC) stand for the optimization criteria used for the first and the second weight. respectively (say power-correlation.

[7] By minimization of a vector, we mean - minimization of all its components simultaneously.

which are minimum power and de-correlation, respectively). The separation algorithms related to the above configurations are depicted in Fig. 3.

2.2 The Forward-Forward Structure

In Fig. 2.b. the forward-forward structure (FF) is depicted. The input-output relation for this configuration is given by

$$\mathbf{y}(t) = \mathbf{W}_{ff}\mathbf{x}(t) \tag{19}$$

where

$$\mathbf{W}_{ff} = \begin{bmatrix} 1 & -w_1 \\ -w_2 & 1 \end{bmatrix} = (1 - w_1 w_2)\mathbf{W}_{bb} \tag{20}$$

where $\mathbf{W}_{bb}$ is given in (6). Since, $\mathbf{W}_{ff}$ and $\mathbf{W}_{bb}$ differ only by the scalar factor, $1 - w_1 w_2$, we easily adopt the results of section 2.1 to this case. In particular

$$\mathbf{y}(t) = \mathbf{H}_{ff}\mathbf{s}(t) + \mathbf{W}_{ff}\mathbf{n}(t) \tag{21}$$

where $\mathbf{H}_{ff} = (1 - w_1 w_2)\mathbf{H}_{bb}$ and

$$\mathbf{R}_{ff} = E\{\mathbf{y}(t)\mathbf{y}^H(t)\} = |1 - w_1 w_2|^2 \mathbf{R}_{bb} \tag{22}$$

where $\mathbf{R}_{bb}$ is given in (10). For this configuration it was proved [3] that, when $\sigma^2 = 0$ (no noise), $\mathbf{H}_{ff}$ is diagonal if and only if $\mathbf{R}_{ff}$ is. Therefore, de-correlation of the output signal is a useful criterion for signal separation, as in the BB configuration. Notice, however that for this configuration

$$\frac{d}{dw_1}E\{|y_1(t)|^2\} = \frac{d}{dw_1}E\{|x_1(t) - w_1 x_2|^2\} = -2E\{x_2^*(t)y_1(t)\} = 0 \tag{23}$$

Therefore, in contrast to the BB configuration wherein de-correlation of the outputs and minimization of their powers are equivalent, in FF such equivalence does not exist. Using de-correlation as an optimization criterion, i.e $E\{y_1(t)y_2^*(t)\} = 0$, $E\{y_1^*(t)y_2(t)\} = 0$, one might control both w_1 and w_2. It is very clear that, again

discrimination is necessary and the resultant separation algorithm, denoted by FFCC (forward-forward correlation-correlation), is depicted in Fig. 4.

As in the previous structure, we may look at each loop as an interference canceler which attempts to reduce the effect of the other signal. The upper loop in Fig. 4 attempts to cancel the interference caused by the input source $s_2(t)$ into $s_1(t)$, while the lower loop attempts to cancel interference caused by the $s_1(t)$ into $s_2(t)$. By virtue of discrimination which enhances the signal component due to $s_2(t)$ over that due to $s_1(t)$, the first loop is able to perform partial cancellation of the interference from $s_2(t)$ into $s_1(t)$. This results in a purer sample of $s_1(t)$ provided to the second loop allowing that loop to perform a partial cancellation in the other direction from $s_1(t)$ into $s_2(t)$. The purer sample of $s_2(t)$ is then used by the first loop to improve cancellation. The above cycle is repeated until essentially perfect signal separation has been achieved. This argument lead to the term bootstrap, although no feedback has been employed. In fact, unlike the BB structure, in the FF structure bootstrapping results from the control. In [3], it was shown that without noise ($\sigma^2 = 0$), the resultant algorithm converges to the unique solution $w_1 = a_1$ and $w_2 = a_2$ provided that $|a_1| < 1$ and $|a_2| < 1$. Notice, however, that the output signals are then given by $y_1(t) = (1 - a_1a_2)s_1(t)$ and $y_2(t) = (1 - a_1a_2)s_2(t)$ and not simply $y_i(t) = s_i(t)$ as in the BB structure. That is, the signals are indeed separated but they are scaled by a factor which is related to the model parameters.

2.3 The Forward-Backward Structure

In Fig. 2.c., a forward-backward (FB) bootstrapped structure is depicted. Its input-output relation is given by

$$\mathbf{y}(t) = \mathbf{W}_{fb}\mathbf{x}(t) \tag{24}$$

where

$$\mathbf{W}_{fb} = \begin{bmatrix} 1+w_1w_2 & -w_1 \\ -w_2 & 1 \end{bmatrix} = \mathbf{W}_{ff} + w_1w_2\mathbf{e}_1\mathbf{e}_1^T, \quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{25}$$

Therefore, this configuration can be alternatively viewed as a FF configuration in parallel with a direct weighted connection between $x_1(t)$ and $y_1(t)$ (see Fig. 5). The relation between the signals to be separated and the output vector is then given by $\mathbf{y}(t) = \mathbf{H}_{fb}\mathbf{s}(t) + \mathbf{W}_{fb}\mathbf{n}(t)$, where

$$\begin{aligned} \mathbf{H}_{fb} &= \mathbf{W}_{fb}\mathbf{A} = \mathbf{H}_{ff} + w_1w_2\mathbf{e}_1\mathbf{e}_1^T\mathbf{A} \\ &= \mathbf{H}_{ff} + w_1w_2 \begin{bmatrix} 1 & a_1 \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{26}$$

The output power matrix, $\mathbf{R}_{fb} = E\{\mathbf{y}(t)\mathbf{y}^H(t)\}$ is given by

$$\begin{aligned} \mathbf{R}_{fb} &= E\{\mathbf{y}(t)\mathbf{y}^H(t)\} \\ &= \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} + \sigma^2 \begin{bmatrix} 1+w_1w_2|^2+|w_1|^2 & -w_1|w_2|^2 - w_2^* - w_1 \\ -w_1^*|w_2|^2 - w_2 - w_1^* & 1+|w_2|^2 \end{bmatrix} \end{aligned} \tag{27}$$

where

$$\begin{aligned} r_{11} &= \sigma_1^2|1-w_1(a_2-w_2)|^2 + \sigma_2^2|a_1 - w_1(1-a_1w_2)|^2 \\ r_{12} &= \sigma_1^2((a_2-w_2)^* - w_1|a_2-w_2|^2) - \sigma_2^2(w_1|1-w_2a_1|^2 - a_1(1-w_2a_1)^*) \\ r_{21} &= \sigma_1^2((a_2-w_2) - w_1^*|a_2-w_2|^2) - \sigma_2^2(w_1^*|1-w_2a_1|^2 - a_1^*(1-w_2a_1)) = r_{12}^* \\ r_{22} &= |a_2-w_2|^2\sigma_1^2 + |1-w_2a_1|^2\sigma_2^2 \end{aligned} \tag{28}$$

Unsurprisingly, the output power of $y_2(t)$ is as in the FF configuration, while the output power of $y_1(t)$ and the correlation $E\{y_1(t)y_2^*(t)\}$ were changed. In [3], it was suggested to control w_1 by minimizing the power $E\{|y_1(t)|^2\}$ and w_2 by de-correlating the two outputs (or equivalently by minimizing $|E\{y_2(t)y_1^*(t)\}|^2$) and hence called power-correlation . With the terminology of the appendix, this is FBPC which is depicted in Fig. 6.a. However, using Fig. 2.c, we write

$$\frac{d}{dw_1}E\{|y_1(t)|^2\} = \frac{d}{dw_1}E\{|x_1(t) - w_1y_2(t)|^2\} = -2E\{y_1(t)y_2^*(t)\}. \tag{29}$$

Therefore, as in the FF structure, we can control w_1 by minimizing either the output power $E\{|y_1(t)|^2\}$ or the magnitude of the correlation $|E\{y_2(t)y_1^*(t)\}|^2$. Also, similar to the argument in (23), w_2, can only be controlled by de-correlating the two outputs. In conclusion, only two control algorithms are possible for the FB structure: the power-correlation (FBPC) and the correlation-correlation (FBCC). These two are depicted in Fig. 6.a and Fig. 6.b. respectively.

In [3] it was shown that if $\sigma^2 = 0$ and $|a_1| < 1$ and $|a_2| < 1$ then the FBPC algorithm converges to $w_2 = a_2$, and $w_1 = \frac{a_1}{1-w_2a_1} = \frac{a_1}{1-a_2a_1}$. Therefore, the desired outputs are $y_1(t) = s_1(t)$ and $y_2(t) = (1 - a_1a_2)s_2(t)$ so the first signal is accurately reconstructed while the second is multiplied by a scalar.

Although all seven possible configurations of the bootstrapped algorithms (Figs. 3.a, 3.b, 3.c, 3.d, 4, 6.a and 6.b) completely separate the signal (in a no-noise environment), they behave differently. While the BB structure resembles the AR model process generator, the FF and FB structures resemble the MA and ARMA process generators, respectively. These principle differences effect stability, bandwidth, convergence, etc. Some of these properties were studied and compared in [10]. In the rest of this appendix, we concentrate on the effect of noise on the average steady state behavior of the different structures.

3 Steady - State Optimal Weight Values

Since we are implementing the steepest descent algorithm, the steady-state values to which the weights converge (on the average) are the optimal values that satisfy the corresponding optimal criteria. We now study the steady-state behavior of the different configurations in the presence of noise. We do that by looking at the values of the weights w_1 and w_2 which satisfy the criteria of (18). We assume that the operation of the discriminators is described by (17) and that it does not effect the noise processes $n_1(t)$ and $n_2(t)$.

3.1 FF/BB Configurations

Using de-correlation in FF configuration, and due to the equivalence between power minimization and de-correlation, in case of BB configurations, the steady-state optimal weights of all FF and BB configurations can be obtained from the simultaneous solution of

$$E\{D_1(y_1)y_2^*\} = 0, \text{ and } E\{y_1^* D_1(y_2)\} = 0 \tag{30}$$

From (12) together with (13), we find that w_{1opt} and w_{2opt} must satisfy

$$\begin{aligned} \sigma_1^2(a_2 - w_2)^*(1 - w_1 a_2) + \delta_{12}^2\sigma_2^2(1 - w_2 a_1)^*(a_1 - w_1) - \sigma^2(w_2^* + w_1) &= 0 \\ \delta_{21}^2\sigma_1^2(a_2 - w_2)(1 - w_1 a_2)^* + \sigma_2^2(1 - w_2 a_1)(a_1 - w_1)^* - \sigma^2(w_2 + w_1^*) &= 0 \end{aligned} \tag{31}$$

These are two non-linear equations in w_1 and w_2. The solution to the equations is the mean of the steady-state value of the weights vector $\mathbf{w}_{opt} = [w_{1opt}\ w_{2opt}]^T$. For $\sigma^2 = 0$ the unique solution is $w_{1opt} = a_1$, $w_{2opt} = a_2$, provided that $|a_1| < 1$, $|a_2| < 1$ [3]. For $\sigma^2 \neq 0$, we write

$$\begin{aligned} w_{1opt} &= a_1 - \epsilon_1 \\ w_{2opt} &= a_2 - \epsilon_2 \end{aligned} \tag{32}$$

Then, using (32) in (31) we get the approximated matrix equation [8]

$$\begin{bmatrix} \Delta\delta_{12}^2 SNR_2 + 1 & \Delta^* SNR_1 + 1 \\ \Delta SNR_2 + 1 & \Delta^* \delta_{21}^2 SNR_1 + 1 \end{bmatrix} \begin{bmatrix} \epsilon_1^* \\ \epsilon_2 \end{bmatrix} = (a_1^* + a_2) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{33}$$

where $\Delta = 1 - a_1 a_2$, $SNR_1 = \frac{E\{|s_1(t)|^2\}}{\sigma^2}$ and $SNR_2 = \frac{E\{|s_2(t)|^2\}}{\sigma^2}$. Thus,

$$\epsilon = \begin{bmatrix} \epsilon_1^* \\ \epsilon_2 \end{bmatrix} = \frac{(a_1^* + a_2)}{det} \begin{bmatrix} \Delta^* SNR_1(\delta_{21}^2 - 1) \\ \Delta SNR_2(\delta_{12}^2 - 1) \end{bmatrix}. \tag{34}$$

and

$$det = |\Delta|^2 SNR_1 SNR_2(\delta_{12}^2 \delta_{21}^2 - 1) + \Delta SNR_2(\delta_{12}^2 - 1) + \Delta^* SNR_1(\delta_{21}^2 - 1) \tag{35}$$

If $\sigma^2 \rightarrow 0$ then $SNR_1 \rightarrow \infty$, $SNR_2 \rightarrow \infty$ and $\epsilon \rightarrow 0$ as expected. If $\sigma^2 \neq 0$ then ϵ is inversely proportional to $\Delta = 1 - a_1 a_2$. Notice that $|a_1| < 1$, $|a_2| < 1$ so $|\Delta| < 1$. The weights w_1 and w_2 can be regarded as estimates of the unknown model parameters a_1 and a_2. As such, they are random variables. Their expected value is given by the solution of (30). From (32), we have

$$\begin{aligned} E\{\hat{a}_1\} &= w_{1opt} = a_1 - \epsilon_1 \\ E\{\hat{a}_2\} &= w_{2opt} = a_2 - \epsilon_2 \end{aligned} \tag{36}$$

Thus, ϵ_1 and ϵ_2 of (34) represent the *bias* in the estimates of the unknown model parameters. To study this bias we first assume that, $\delta_{21}^2 = 1$, (that is, there is no discrimination in the corresponding control algorithms). In this case, $\epsilon_1^* = 0$

$$\epsilon_2 = \frac{(a_1^* + a_2)\Delta SNR_2(\delta_{12}^2 - 1)}{|\Delta|^2 SNR_1 SNR_2(\delta_{12}^2 - 1) + SNR_2\Delta(\delta_{12}^2 - 1)} \approx \frac{(a_1^* + a_2)}{SNR_1\Delta^*} = \frac{(a_1^* + a_2)}{SNR_1(1 - a_1 a_2)^*} \tag{37}$$

Thus, the bias is inversely proportional to the signal-to-noise ratio (SNR) and to the coupling effect in the model, Δ. Notice that for $|\Delta| = 0$, we must have $a_1 = a_2 = 1$ which means that at the inputs to the separator, the signals are superimposed equally.

[8] In the transition from (31) to (33) we assume that $1 - w_1 a_1 = \Delta + \epsilon_1 \approx \Delta$ and $1 - w_2 a_1 = \Delta + \epsilon_2 \approx \Delta$. This assumption will be discussed later.

$|\Delta| = 1$, we must have $a_1 = a_2 = 0$ which means that at the input the two signals are totally separated. In particular if

$$\frac{1 + \frac{a_1^*}{a_2}}{SNR_1 \Delta^*} > 1 \tag{38}$$

then, the bias is larger than the estimated parameter. Similar results can be obtained for ϵ_1 by letting $\delta_{12}^2 = 1$. Thus, as a rule of thumb, it can be said that $SNR_{min}\Delta$, where $SNR_{min} = min\{SNR_1, SNR_2\}$ is an important design factor. If $SNR_{min}\Delta >> 1$, the bias in the estimates of the model parameters might be negligible in comparison to the parameter to be estimated (particularly, if a_1 and a_2 are of the same order of magnitude) and our assumption in deriving Eq. (33) are satisfied. If, on the other hand, the SNR is not large enough or the coupling is too strong so $|\Delta| << 1$, then the bias may dominate the estimate and may cause severe degradation in separation performance.

3.2 FB Configuration

For the configurations depicted in Fig. 6.a and Fig. 6.b, $w_2 = \hat{a}_2$ as before, while w_1 is not an estimate of a_1, but rather related to it through $\hat{a}_1 = \frac{w_1}{1+w_1 w_2}$. As in section 3.1, and due to the equivalence between power minimization and de-correlation, we find the steady-state optimal values of w_1 and w_2 by imposing the de-correlation requirement of both discriminated output signals, that is from (27) together with (28), we find that w_{1opt} and w_{2opt} must satisfy

$$\begin{aligned}
\sigma_1^2[(a_2 - w_2)^* - w_1|a_2 - w_2|^2)] - \delta_{12}^2\sigma_2^2(w_1|1 - w_2a_1|^2 - a_1(1 - w_2a_1)^* & \\
-\sigma^2(w_2^* + w_1 + w_1|w_2|^2) &= 0 \\
\delta_{21}^2\sigma_1^2(a_2 - w_2) - w_1^*|a_2 - w_2|^2)] - \sigma_2^2[w_1^*|1 - w_2a_1|^2 - a_1^*(1 - w_2a_1)] & \\
-\sigma^2(w_2 + w_1^* + w_1^*|w_2|^2) &= 0
\end{aligned} \tag{39}$$

Define ϵ_1 and ϵ_2 as the bias in the estimate of a_1 and a_2 , respectively, so

$$\begin{aligned} w_{2opt} &= a_2 - \epsilon_2 \\ \frac{w_{1opt}}{1 + w_{1opt}w_{2opt}} &= a_1 - \epsilon_1 \\ w_{1opt} &= \frac{a_1 - \epsilon_1}{1 - (a_1 - \epsilon_1)(a_2 - \epsilon_2)} \end{aligned} \tag{40}$$

From (39) we get

$$(1 + w_1 w_2)^* \left[\sigma_1^2 \epsilon_2 (\Delta + \epsilon_1 a_2)^* + \delta_{12}^2 \sigma_2^2 (\Delta + a_1 \epsilon_2) \epsilon_1^* - \sigma^2 (a_1^* + a_2 - \epsilon_1^* - \epsilon_2) \right] = 0$$

$$(1 + w_1 w_2) \left[\delta_{21}^2 \sigma_1^2 \epsilon_2^* (\Delta + \epsilon_1 a_2) + \sigma_2^2 (\Delta^* + a_1 \epsilon_2^*) \epsilon_1 - \sigma^2 (a_1 + a_2^* - \epsilon_1 - \epsilon_2^*) \right] = 0 \tag{41}$$

Since, $|w_1 w_2| < 1$ (by restriction), we can write the approximation [9]

$$\begin{bmatrix} \Delta \delta_{12}^2 SNR_2 + 1 & \Delta^* SNR_1 + 1 \\ \Delta SNR_2 + 1 & \Delta^* \delta_{21}^2 SNR_1 + 1 \end{bmatrix} \begin{bmatrix} \epsilon_1^* \\ \epsilon_2 \end{bmatrix} = (a_1^* + a_2) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{42}$$

which is exactly the same as in (33). That is, the *bias* vector for the unknown model parameters $\epsilon = [\epsilon_1 \ \epsilon_2]^T$ is the same for all bootstrapped configurations, at least for small bias, where the approximation holds.

[9] the assumptions used are: $\epsilon_1 a_2 << \Delta$ and $\epsilon_2 a_1 << \Delta$, which is similar to the assumption in the BB/FF case, but ϵ_1 is different here.

4 Example: Separation of Cross Polarized Signals

In this section, we apply the bootstrapped signal separation algorithm to the separation of cross polarized signals [11]. Dual-polarization techniques have been applied to radio communication networks to increase the transmission capacity of limited bandwidth channels such as satellite communication channels and microwave radio channels. In a dual polarized transmission system, the available bandwidth is doubled by modulating the same carrier frequency with two independent information signals. The two modulated signals are then transmitted through the channel, with one having vertical polarization and the other having horizontal polarization. Because of antenna imperfections and/or non ideal transmission channel conditions (caused, for example, by fading), the received signals are not perfectly orthogonal. Therefore, cross polarization coupling of each information signal to other is created causing, in some cases, severe degradation in performance. The cross coupling parameters can be assumed to be slowly time varying and non-dispersive. Bootstrapped separation algorithms have been proposed [4,6,12,13] to mitigate degradation in performance due to cross polarization interference. It results in an increase in signals to cross polarization interference ratio and hence, an improved signal separation. In this application, the normalized coupling factor, $|a_1|$ and $|a_2|$ is sufficiently small, $(-10dB = 10log|a_i|^2)$ and signals-to-noise ratios are of the order of 20 dB or more. This leads to $SNR_{min}\Delta >> 1$ so that the effect of the additive thermal noise is small. In the following, we present results concerning the output signal-to-interference plus noise ratio (SINR) for three configurations.

The input SNIR at the channel output (at the separator input) say $x_1(t)$[10] , is given by

[10]Due to symmetry we get similar expression for the other input.

$$\left(\frac{P_{s1}}{P_{s_2}+P_n}\right)_{input} = \frac{\sigma_1^2}{|a_1|^2\sigma_2^2+\sigma^2}$$
$$\left(\frac{P_{s1}}{P_{s_2}+P_n}\right)_{input} = \frac{SNR_1}{|a_1|^2 SNR_2+1} \tag{43}$$

The output SNIR depends on the steady-state values of the weights, and therefore on ϵ_1 and ϵ_2. In our example taking $|a_1|,\ |a_2| < 0.3$ results in $|\Delta| = |1-a_1a_2| \approx 1$, (this is a fair assumption in the transmission of two orthogonally polarized M-ary QAM signals [6]). By taking SNR_1 and $SNR_2 >> 1$, one can show that under these assumptions the weight bias vector from (34) can be approximated by

$$\begin{bmatrix} \epsilon_1^* \\ \epsilon_2 \end{bmatrix} \approx \frac{(a_1^*+a_2)}{|\Delta|^2 SNR_1 SNR_2} \begin{bmatrix} \Delta^* SNR_1 \\ \Delta SNR_2 \end{bmatrix}, \tag{44}$$

This approximation is applicable to <u>all</u> configurations. Now, for both BB and FF[11] we get by using (12) with (13) and (22) the output SNIR at the separator output at $y_1 = \hat{s}_1$

$$\left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} = \frac{|1-w_{1opt}a_2|^2\sigma_1^2}{|a_1-w_{1opt}|^2\sigma_2^2+\sigma^2(1+|w_{1opt}|^2)} \tag{45}$$

Substituting for w_{1opt} from (32) and using the fact that $\Delta \approx 1 >> \epsilon_1 a_2$ and $SNR_2 >> 1$, we get

$$\left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} \approx \frac{|1-a_1a_2|^2 SNR_1}{|\epsilon_1|^2 SNR_2+1} \tag{46}$$

Using the approximated ϵ_1 from (44), we get

$$\left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} \approx \frac{|\Delta|^2 SNR_1}{|\frac{a_1+a_2^*}{\Delta^* SNR_2}|^2+1} \approx \frac{SNR_1}{\frac{|a_1+a_2^*|^2}{SNR_2}+1} \tag{47}$$

Notice that as $SNR_2 \to \infty$ the output SNIR approaches SNR_1, that is, as the signal-to-noise ratio of the second signal gets stronger, the interference cancellation

[11] From the relation between R_{ff} and R_{bb} in (22) the output power ratios are the same as in BB configuration and given by (45).

gets better. This result is applicable to both BB and FF configurations for any of the two outputs.

For the FB configuration, by using (27) with (28) we get the output SNIR at the separator output $y_1 = \hat{s}_1$,

$$\left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} = \frac{|1-w_{1opt}(a_2-w_{2opt})|^2\sigma_1^2}{|a_1-w_{1opt}(1-w_{2opt}a_1)|^2\sigma_2^2+\sigma^2[|1+w_{1opt}w_{2opt}|^2+|w_{1opt}|^2]} \tag{48}$$

at $y_2 = \hat{s}_2$, using (27) with (28), the SNIR is given by

$$\left(\frac{P_{s_2}}{P_{s_1}+P_n}\right)_{output} = \frac{|1-w_{2opt}a_1|^2 SNR_2}{|a_2-w_{2opt}|^2 SNR_1+[|1+|w_{2opt}|^2]} \tag{49}$$

Substituting for w_{1opt} and w_{2opt} from (40) and with high SNR assumption we get

$$\begin{aligned}\left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} &= \frac{|1-\frac{w_{1opt}}{1+w_{1opt}w_{2opt}}a_2|^2 SNR_1}{|a_1-\frac{w_{1opt}}{1+w_{1opt}w_{2opt}}|^2 SNR_2+(1+|\frac{w_{1opt}}{1+w_{1opt}w_{2opt}}|^2)} \\ &= \frac{|\Delta+\epsilon_1 a_2|^2 SNR_1}{|\epsilon_1|^2 SNR_2+(1+|a_1-\epsilon_1|^2)} \\ \left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} &\approx \frac{|\Delta+\epsilon_1 a_2|^2 SNR_1}{|\epsilon_1|^2 SNR_2+1}\end{aligned} \tag{50}$$

Now using ϵ_1 from (44) we get

$$\begin{aligned}\left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} &\approx \frac{|\Delta+\frac{a_1+a_2^*}{SNR_2}a_2|^2 SNR_1}{\frac{|a_1+a_2^*|^2}{|\Delta|^2 SNR_2}+1} \\ \left(\frac{P_{s_1}}{P_{s_2}+P_n}\right)_{output} &\approx \frac{|1+\frac{a_1+a_2^*}{SNR_2}a_2|^2 SNR_1}{\frac{|a_1+a_2^*|^2}{SNR_2}+1}\end{aligned} \tag{51}$$

For $SNR_2 >> 1$, this expression, as (47) can be approximated by simply SNR_1 (namely, perfect cancellation). However, for moderate levels of SNR_2, the output SNIR at $y_1 = \hat{s}_1$ for the FB configuration is slightly higher than that of the BB configuration or FF configurations. For the other output of the FB configuration, the output SNIR is exactly as for the BB/FF configurations.

To summarize the results of this section, we notice for cases of small coupling between the signals and high SNR, such that $|\Delta|^2 SNR_1 SNR_2 >> 1$, the output

SNIR of any of the bootstrapped separators depends on the SNR of the other signal. Practically, if the other signal SNR is high enough, the cancellation becomes perfect.

5 Conclusion

To summarize the results in this appendix, we emphasize the following

1. The basic 2-signals' separation cell contains two weighted and two unweighted paths. Each weighted path can be either a backward (feedback) or a forward directed.

2. The basic optimization criterion for signal separation is the de-correlation of the output signals. However, the backward path can equivalently be controlled by minimizing the corresponding output power.

3.Unless signals in the control loops are somewhat *discriminated*, the separation problem can *not* be solved. On the other hand, perfect separation is obtained for infinite signal to noise ratio (no noise) even if the discrimination is slight.

4. The bootstrapped separation algorithms provide separation of the signals as well as estimates of the model parameters. With no noise, these estimates are *unbiased*, in the steady-state. Additive thermal noise at the separator inputs causes *bias* in these estimates. This bias is inversely proportional to the input SNR and increases with the signal cross-coupling. The bias causes degradation in the separation performance, but this degradation is small if $|\Delta|SNR >> 1$, where $\Delta = 1 - a_1 a_2$ is the coupling factor.

5. The effect of the additive noise is an inherent factor. Since the input to the separator consists of the superimposed signals and uncorrelated noise processes, the output, if separation is done, consists of superimposed noise processes and uncorrelated signals. That is, the noise processes at the two outputs are inherently correlated.

6. It is interesting to study the case where only one signal is present. This case may be realistic if one of the signals to be separated is absent, temporarily. Formally, if one takes, say $\sigma_2^2 = 0$ in (12), we see that any criterian used will result in only *one*

control equation (assuming no noise).

$$\sigma_1^2(a_2 - w_2)(1 - w_1 w_2)^* = 0 \tag{52}$$

The operation of a discrimination will not produce a second linear independent equation whether noise is present or not. Thus, the problem is ill-stated and it has *no* solution for w_1 and w_2. This can also be explained by the fact that our model assumes $|a_2| < 1$ and $|a_1| < 1$ meaning that, at the inputs to the separator

$$\text{at 1 st input } x_1 \quad \frac{1}{SNR}\frac{\text{power of 1 st signal}}{\text{power of 2 nd signal}} = \frac{1}{|a_1|^2} > 1$$

$$\text{at 2 nd input } x_2 \quad \frac{1}{SNR}\frac{\text{power of 1 st signal}}{\text{power of 2 nd signal}} = |a_2|^2 < 1 \tag{53}$$

where $SNR = \frac{E\{|s_1(t)|^2\}}{E\{|s_2(t)|^2\}} = \frac{\sigma_1^2}{\sigma_2^2}$ If $\sigma_2^2 = 0$ then $\frac{1}{SNR} = 0$ so Eq. (30a) can never be satisfied. Thus, the bootstrapped separator is only applicable for separation of signals and not for noise reduction. etc. (as the LMS algorithm).

7. There exist many practical important separation problems for which $\Delta SNR >> 1$ (as the cross polarization example of sect. IV.). In this case, the output SINR is of the order of the input SNR. namely the separation is practically perfect.

8. All other papers dealing with the same separation problem [3,4] came to a separation configuration which is either a BB or a FF one. The papers which use a similar approach to ours, suggest the de-correlation of the outputs as the only optimization criteria. No other authors use our "discriminator." However, some suggested the use of nonlinearities in the control loop, which is a special case of statistical discrimination. All other papers in which the unknown model parameters are constants are special cases, mentioned in our appendix. However, there have been efforts to use similar principals for model parameters which are either digital filters coefficients or delay lines [14]. In general. similar principles are applicable and similar results are available. There are, however some differences. For example, in those cases, de-

correlation of the outputs is a necessary but not sufficient condition for minimization of the output powers.

Appendix

It is easy to show by using (40) that

$$|1 - w_{1opt}(a_2 - w_{2opt})|^2\sigma_1^2 = \left|\frac{1 - (a_1 - \epsilon_1)a_2}{1 - (a_1 - \epsilon_1)(a_2 - \epsilon_2)}\right|^2\sigma_1^2 \qquad (A.1)$$

$$|a_1 - w_{1opt}(1 - w_{2opt}a_1)|^2\sigma_2^2 = \left|\frac{\epsilon_1}{1 - (a_1 - \epsilon_1)(a_2 - \epsilon_2)}\right|^2\sigma_2^2 \qquad (A.2)$$

$$|1 + w_{1opt}w_{2opt}|^2\sigma^2 = \left|\frac{1}{1 - (a_1 - \epsilon_1)(a_2 - \epsilon_2)}\right|^2\sigma^2 \qquad (A.3)$$

$$|w_{1opt}|^2\sigma^2 = \left|\frac{|a_1 - \epsilon_1|^2}{1 - (a_1 - \epsilon_1)(a_2 - \epsilon_2)}\right|^2\sigma^2 \qquad (A.4)$$

References

[1] Widrow, B. et al.," The Complex LMS Algorithm." *IEEE Proceedings*, Vol. 63, No. 12, pp. 1692-1716, Dec. 1975.

[2] Bar-Ness, Y. and Rockach, S., " Cross-Coupled Bootstrapped Interference Canceler," *Proc. of Intern. Conf. on Ant. and Prop.*, pp. 292-5, Vol. I, LA, CA, June 1981.

[3] Bar-Ness, Y., " Bootstrapped Adaptive Cross-Pol Interference Cancelling Techniques- Steady State Analysis,"*Bell Labs. Report,* Jan. 26, 1982, also accepted for publication to the *Signal Processing*, Elsevier Science Publishers.

[4] Bar-Ness, Y., Carlin, J. W., Steinberger, M. L, " Bootstrapping Adaptive Cross-Pol Canceler for Satellite Communication," , *Proceedings of ICC-82*, paper 4F.S, June 1982.

[5] Bar-Ness, Y., Carlin, J. W., Steinberger, M. L , " Bootstrapping Adaptive Interference Cancellers: Some Practical Limitations," *Proceedings of The Globecom Conf.* paper F3.7, Nov. 1982.

[6] Carlin, J., Bar-Ness, Y. and Studdiford, W. E., " An IF Cross-Pol Canceler for Microwave Radio." *Journal on Selected Areas in Communication - Advances in Digital Communication by Radio* , Vol. SAC-5, No. 3 pp. 502-514, April 1987.

[7] Bar-Ness, Y., " Bootstrapped Algorithms for Interference Cancellation," , *AFCEA-IEEE Technical Conf. on Tactical Communications*, Fort Wayne, IN, May 1988.

[8] Jutten, C and Herault, J., " Blind Separation of Sources, Part I," *Signal Processing*, Vol. 245, No. 1, pp. 1-10, July 1990.

[9] Weinstein, E., Feder, M. and Oppenheim, A. V. , " Multi-Channel Signal Separation Based on De-correlation," *submitted to IEEE Trans. on ASSP* , 1992.

[10] Bar-Ness, Y., " Bootstrapped Adaptive Cross-Pol Interference Cancelling Techniques- Bandwidth Complexity Tradeoff," Bell Labs. Report, also *submitted to the Elsevier Science Publishers*

[11] Amitay N., " Signal-to-Noise Ratio Statistics for Non-dispersive Fading in Radio Channels with Cross Polarization Interference Cancellation," *IEEE Trans. on Commun.*, Vol. 27, pp. 498-502, Feb. 1979.

[12] Dinç, A. and Bar-Ness, Y., " Performance Comparison of LMS, Diagonalizer and Bootstrapped Adaptive Cross-Pol Cancellers For M-ary QAM," *Proceedings of MILCOM '90*, paper 3.7, Monterey, CA, Oct.1990.

[13] 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.

[14] Messer, H. and Bar-Ness, Y., " Bootstrapped Spatial Separation of Wide-band Superimposed Signals," EUSIPCO '92, Brussles, Belgium, pp. 819-822.

[15] Brandwood, D.H., " Cross Coupled Cancellation System for Improving Cross Polarization Discrimination." *Int. Conf. on Ant. Prop.*, 41-5, Part I., London, England, Nov. 1978.

[16] Blaquiere, A., " Nonlinear System Analysis," Academic Press, 1966.

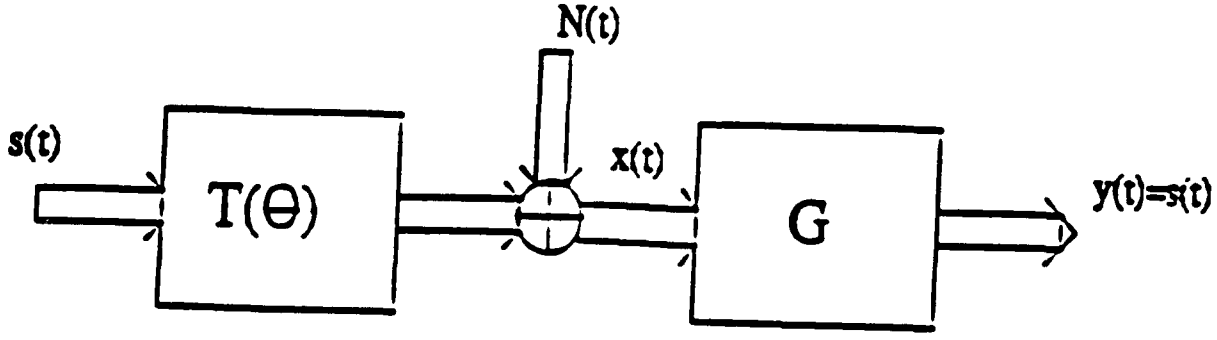


Fig. 1 Channel and Canceler Block Diagram

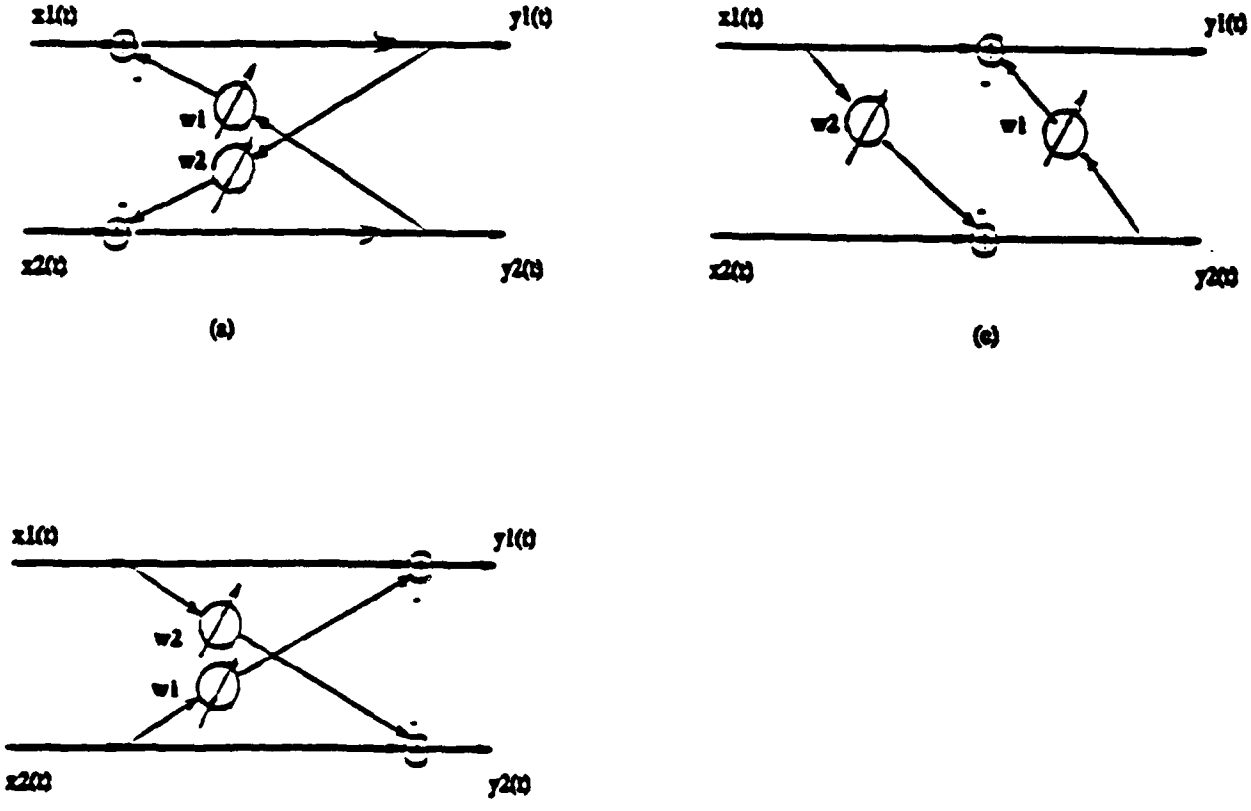


Fig. 2 The three configuration of the bootstrapped signal separator

a. Backward/Backward Structure
b. Forward/Forward Structure
c. Forward/Backward Structure

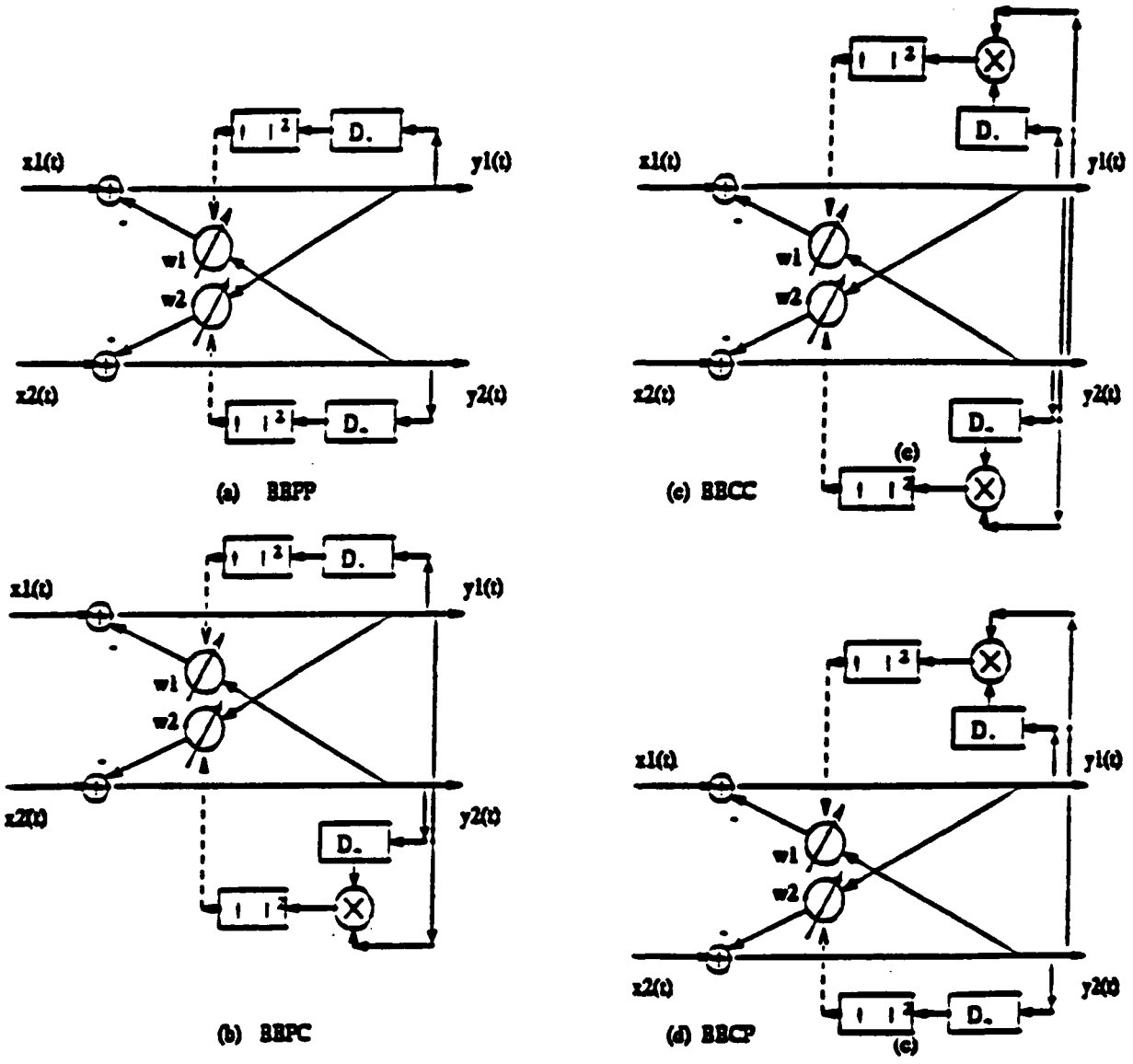


Fig. 3 Different Optimization Criteria for the Backward/Backward Structure

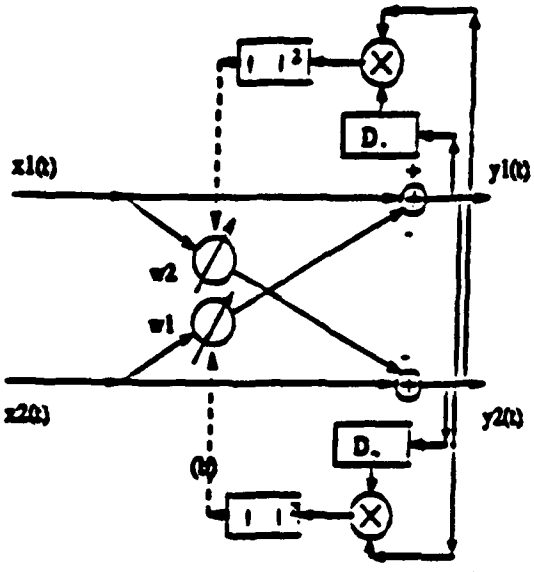


Fig. 4

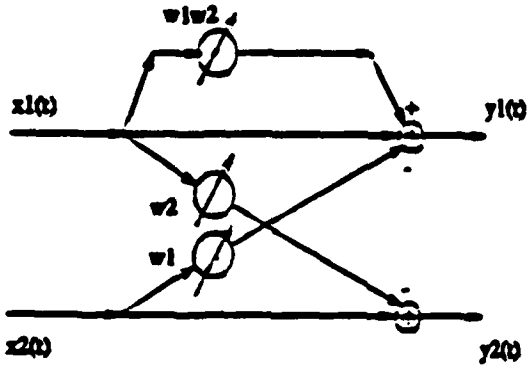


Fig. 5. An alternative view of the FB structure

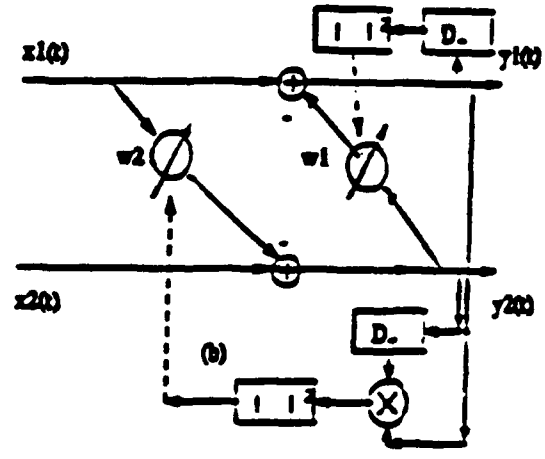


Fig. 6.a

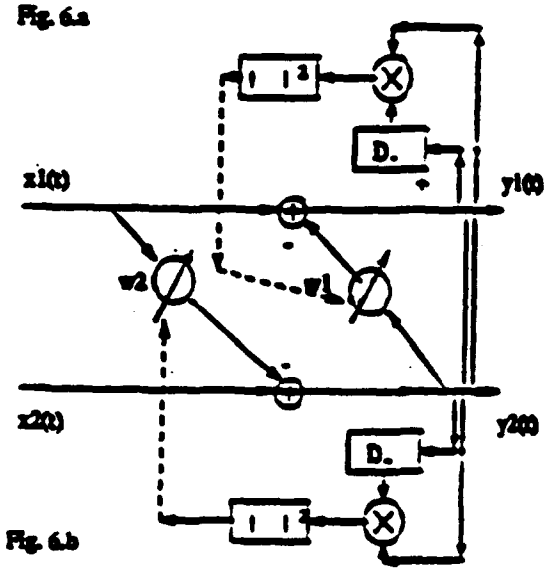


Fig. 6.b

Fig. 4 Correlation-Correlation Optimum Criterion for Forward/Forward Structure
Fig. 5 An Alternative View for Forward/Backward Structure
Fig. 6 Possible Optimization Criteria for Forward/Backward Structure

Appendix C

Power - Power Bootstrapped Cross-Pol Canceler for M-ary QAM Signals - Performance Evaluation and Comparison with LMS and Diagonalizer

Abdulkadir Dinç and Yeheskel Bar-Ness

1 Introduction

With many transmission systems, capacity can be doubled by using orthogonally polarized QAM carriers. The orthogonally polarized waves can suffer degradation due to carrier -to-cross polarization interference (C/XPI). Particularly, during multipath fading, such degradation could become intolerable.

Many methods have been suggested to cancel the XPI. Among these are the diagonalizer and the least mean square (LMS) cross-pol canceler proposed in [1]. In this paper, Kavehrad compares the error performance of these two cancelers and concludes that the LMS is substantially better than the diagonalizer. A similar diagonalizer was also proposed by Duvoison et. al. [2].

Still another method for cross-polarization interference XPI canceler is termed bootstrapped cancelers, of which three structures have been proposed and analyzed in [3,4,5]. It has been shown that, under the assumption of no additive noise or when the input signals to noise ratio is very high, all these structures converge to a state of total diagonalization. In other words, the XPI at each of the cancelers, two outputs are totally cancelled similar to Kavehrad's or the other diagonalizer considered. Nevertheless, due to the fact that these bootstrapped structures differ from the diagonalizer or the LMS cancelers, the noise terms at their outputs are different and so is the probability of error performance. Although the bootstrapped structures have been implemented in different applications, such as satellite dual-polarized communication [4], QAM microwave radio communication [5] and in tactical communication

[6], no attempt was made to quantify the bootstrapped cross-polarization cancelers (BXPC) probability of error performance. In this study, we derive the average probability of error of the BXPC as a function of its input signal-to-noise ratio and compare it with the performance of the other two methods, the diagonalizer and LMS which were presented in [1]. We will show that the BXPC outperforms both the LMS and the diagonalizer in cancelling XPI.

In this report, we first introduce a model for the dual-channel M-ary transmission in section (2) and then calculate the error probability caused by cross-polarization interference as well as the noise section (3). Two interference cancelers, namely the diagonalizer and the LMS cancelers are introduced and their error probabilities are estimated in section (4). These results are mostly based on previously published material and presented in this work for completeness and the convenience of the reader.

The power-power bootstrapped cross-pol cancelers will be discussed in section (5). There, after deriving the canceler's parameter such as optimal weights with and without noise effect, we find the canceler's optimal outputs. Assuming amplitude compensation alone, or amplitude and phase compensation of the canceler's outputs, we derive seperately decision parameters for obtaining the corresponding error terms. In section 5.2, we derive the Chernoff bound on the average probability of error on one hand and define an expression to be used in calculating an approximation to the probability of error by the method of moments, on the other. Finally, in section 5.3, we present results on the performance of the power-power canceler showing Chernoff bounds and actual approximations to error probability based on the moments, for different cases and with different parameters. These results are also compared, in order to draw the conclusions in section 5.4.

2 Dual-Channel M-ary QAM Transmission Model

The model for such channels have been well presented in the literature [7,8,9,10]. M-ary QAM bandpass signals with the same bandwidth and the same center frequency transmitted on two orthogonal channels can be presented as

$$s_i(t) = Re\{\tilde{s}_i(t)\exp(j2\pi f_c t)\} \tag{1}$$

where $Re\{\cdot\}$ stands for the real part, f_c denotes the carrier frequency and $\tilde{s}_i(t), i = 1, 2$ is the complex envelopes of each of the orthogonal signals, respectively. These complex envelope can be expressed as

$$\tilde{s}_i(t) = \sum_{k=0}^{\infty} I_{ki} h_i(t - kT) \tag{2}$$

where I_{ki} $i = 1, 2$ is a complex information symbol which takes on one of M different complex values, where $I_{ki} = I_{kR} + jI_{kI}$, and I_{kR} and I_{kI} (the in-phase and the quadrature component of the carrier), are independent M-ary symbols from the set $\{\pm c, \pm 3c, \ldots, \pm(\sqrt{M} - 1)c\}$. I_{kR} and I_{kI} can each take values equal to $2l - 1 - \sqrt{M}, l = 1, 2, \ldots, \sqrt{M}$. I_{kR} and I_{kI} are assumed to be statistically independent and equiprobable. Also, $h(t)$ is a complex low-pass equivalent of the overall system impulse response, M is the number of signal level of in-phase and quadrature component and c is a constant which determines the distance to the decision boundary from each signal location.

The channel is assumed to be slowly time varying and non-dispersive. It accepts two orthogonally independent random data streams $I_1(n)$ and $I_2(n)$. It causes distortion: a fraction of one stream of data is added to the other [10].

In matrix notation the received signal is given by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{3}$$

where $\mathbf{A}$ is the dual-channel cross coupling matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{4}$$

a_{12} and a_{21} are complex valued constants that denote the channels cross-polarization (inter-channel interference) responses. The factors a_{11} and a_{22} denote co-polarization (direct path attenuation) channel constants taken as real valued [16].

In (3) $\mathbf{s}(t) = [s_1(t), s_2(t)]^T$ and the noise $\mathbf{n}(t) = [n_1(t), n_2(t)]^T$ with $n_i(t) = n_{iR}(t) + jn_{iI}(t), i = 1, 2$

The received signals, sampled after matched filters, are denoted by;(see Fig 1)

$$\begin{aligned} x_1(n) &= a_{11}I_1(n) + a_{12}I_2(n) + n_1(n) \\ x_2(n) &= a_{21}I_1(n) + a_{22}I_2(n) + n_2(n) \end{aligned} \tag{5}$$

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. Also $n_1(n)$ and $n_2(n)$ are independent samples of Gaussian zero mean random process.

The channel coefficients a_{ij} $i = 1, 2$, $j = 1, 2$ are assumed to vary slowly with respect to the signal rate. These slow variations can be tracked by the adaptive algorithms.

We define the normalized cross-polarization coefficients,

$$\frac{a_{12}}{a_{22}} = r_1 e^{j\phi_1}, \quad \frac{a_{21}}{a_{11}} = r_2 e^{j\phi_2}, \tag{6}$$

where r_1, r_2 denote the magnitude of the normalized cross polarization constants and o_1, o_2 are the phases of these constants assumed to be independent and uniformly distributed over $[-\pi, \pi]$.

3 Performance of Dual Polarized M-ary QAM System

To estimate the performance of such a system and realize the effect of cross-polarization, we will find an estimate for the probability of errors that each output will suffer. As a standard procedure, and based on the kind of signal processing performed at the outputs of the receiver, we define the decision parameter. In this

chapter we will take as the decision parameter $\hat{I}_1(n) = \frac{y_1(n)}{a_{11}}$ and $y_1(n) \neq x_1(n)$[12]. In the next chapter this normalization will be termed "amplitude" only compensation. Define,

$$Z_1(n) \triangleq \hat{I}_1(n) - I_1(n) \tag{7}$$

From (5) and (7), we write

$$Z_1(n) = \frac{a_{12}}{a_{11}} I_2(n) + \frac{n_1(n)}{a_{11}} \tag{8}$$

Next, we write $Z_1(n)$ in terms of its real , Z_{1R}, and imaginary,Z_{1I}, parts. For this, assuming $a_{11} = a_{22}$ we use (6) in (8) and present $I_2(n)$ and $n_1(n)$ in terms of their real and imaginary parts;

$$\begin{aligned} Z_{1R} &= r_1 I_{2R} cos\phi_1 - r_1 I_{2I} sin\phi_1 + \frac{n_{1R}}{a_{11}} \\ Z_{1I} &= r_1 I_{2R} sin\phi_1 + r_1 I_{2I} cos\phi_1 + \frac{n_{1I}}{a_{11}} \end{aligned} \tag{9}$$

For a matter of convenience. we dropped in (9) the dependence on the sampling time n.

Based on the decision parameters in (9), Kavehrad [1] finds the Chernoff bound on the probability of error at the output. He also uses the Gauss quadrature rule [11] to obtain an approximate value of the probability of error.

For the convenience of the reader, we will summarize Kavehrad's GQR procedure:

Define,

$$\begin{aligned} X_I &= r_1(I_{2R} cos\phi_1 - I_{2I} sin\phi_1) \\ Y &= \frac{n_{1R}}{a_{11}} \end{aligned} \tag{10}$$

then

$$Z_{1R} = X_I + Y \tag{11}$$

[12]when we add canceler: $y_i(n)$ will be the output of the canceler.

where Y is zero mean Gaussian random variable with variance $\frac{\sigma_n^2}{a_{11}^2}$. Therefore, Z_{1R} is Gaussian with mean X_I and variance $\frac{\sigma_n^2}{a_{11}^2}$. It is possible to show that,

$$P_1(|Z_{1R}| > c|\phi_1, I_{2R}, I_{2I}) = 2Q\Big(\frac{c - X_I}{\sigma}\Big) \tag{12}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty exp(\frac{-t^2}{2})\, dt$$

and $\sigma^2 = \frac{\sigma_n^2}{a_{11}^2}$

Using the relation between the probability of error $P_1(e)$ and $P_1(|Z_{1R}| > c)$, we have

$$P_1(e|\phi_1, I_{2R}, I_{2I}) = 2(1 - \frac{1}{\sqrt{M}})Q\Big(a_{11}(c - X_I)\frac{1}{c}\sqrt{\frac{3\ SNR}{M-1}}\Big) \tag{13}$$

where we used the well known relation between SNR and M-QAM signal parameter;

$$SNR = \frac{M-1}{3}\frac{c^2}{\sigma_n^2} \tag{14}$$

Defining the random variable.

$$\mathbf{x} = \frac{X_I}{c} \tag{15}$$

then, (13) becomes.

$$P_1(e|x) = 2(1 - \frac{1}{\sqrt{M}})Q\Big(a_{11}(1 - \mathbf{x})\sqrt{\frac{3\ SNR}{M-1}}\Big) \tag{16}$$

The average probability can be approximated, using the GQR by;

$$P_1(e) \approx 2(1 - \frac{1}{\sqrt{M}})\sum_{i=1}^{N} w_i Q\Big(a_{11}\sqrt{\frac{3\ SNR}{M-1}}(1 - x_i)\Big) \tag{17}$$

where x_i nd w_i are the nodes and weights of the GQR which can be obtained from the moments of the random variable $\mathbf{x}$ defined in (15) together with (10).

Based on (17), we calculate and present in Fig. 2 the probability of error at the output of a channel for 16 QAM dually polarized signals as a function of the transmitter signal to noise ratio. The cross-coupling assumed to be -15 dB. 32 moments were used when applying the GQR in (17).

4 Cross-Polarization Cancelers

We notice from (5) that the output of the channel contains, besides the noise term, an interference (cross polarization) through the cross coupling a_{12} and a_{21}. This interference undoubtly causes degradation of performance. Several different canceler structures were proposed to mitigate the effect of cross-polarization. Among these are the diagonalizer [10] and the LMS [12] and the bootstrapped canceler [4].

For the convenience of the reader, we will summarize Kavehrad's approach to estimate the performances of the diagonalizer and the LMS cross-pol cancelers. We will introduce the decision parameters of these cancelers as they have been derived for M-ary QAM system by Kavehrad, follow his derivation for the Chernoff bound on the probability of error and compare the results of these two algorithms with each other for different cross-couplings. It should be noticed that the decision parameters for the outputs were derived in [1] under the assumption that only amplitude compensation is used at the output. Also, it should be emphasized that unlike the LMS canceler, Kavehrad completely neglects the effect of noise on optimal weight when he deals with the diagonalizer.

4.1 The Diagonalizer and Its Performance

The structure of this canceler is well presented in [10], [1], see Fig 3. The output of the diagonalizer is given by

$$\begin{aligned} y_1(n) &= w_{11}x_1(n) + w_{12}x_2(n) \\ y_2(n) &= w_{21}x_1(n) + w_{22}x_2(n) \end{aligned} \tag{18}$$

and substituting (5) in (18), we get

$$\begin{aligned} y_1(n) &= w_{11}[a_{11}I_1(n) + a_{12}I_2(n) + n_1(n)] + w_{12}[a_{21}I_1(n) + a_{22}I_2(n) + n_2(n)] \\ y_2(n) &= w_{21}[a_{11}I_1(n) + a_{12}I_2(n) + n_1(n)] + w_{22}[a_{21}I_1(n) + a_{22}I_2(n) + n_2(n)] \end{aligned}$$

(19)

The canceler weights are found by forcing the coefficients of the interference signal to zero on each channel. Therefore, from (19) we must choose the weights to satisfy,

$$\begin{aligned} w_{11}a_{12} + w_{12}a_{22} &= 0 \\ w_{21}a_{11} + w_{22}a_{21} &= 0 \end{aligned} \tag{20}$$

By substituting the constraint of (20) in (19), we get after using (5),

$$\begin{aligned} y_1(n) &= a_{11}[1 - r_1 r_2 e^{j(\phi_1+\phi_2)}][I_{1R}(n) + jI_{1I}(n)] + n_{1R}(n) + jn_{1I}(n) \\ &\quad -(n_{2R} + jn_{2I})r_1 e^{j\phi_1} \end{aligned}$$

$$\begin{aligned} y_2(n) &= a_{22}[1 - r_1 r_2 e^{j(\phi_1+\phi_2)}][I_{2R}(n) + jI_{2I}(n)] + n_{2R}(n) + jn_{2I}(n) \\ &\quad -(n_{1R} + jn_{1I})r_1 e^{j\phi_2} \end{aligned} \tag{21}$$

Following Kavehrad, we define, $\hat{I}_1(n) = \frac{y_1(n)}{a_{11}}$ as an estimate of the transmitted signal $I_1(n)$ and hence definition of decision parameter follows

$$Z_1(n) = \hat{I}_1(n) - I_1(n) \tag{22}$$

Using (22) in (21), we write the decision parameter for the output of channel 1, in terms of its real and imaginary parts;

$$\begin{aligned} Z_{1R} &= -I_{1R}(n)r_1 r_2 cos(\phi_1 + \phi_2) + I_{1I}(n)r_1 r_2 sin(\phi_1 + \phi_2) \\ &\quad + \frac{n_{1R}(n)}{a_{11}} - \frac{n_{2R}}{a_{11}} - \frac{n_{2R}}{a_{11}} r_1 cos\phi_1 + \frac{n_{2I}}{a_{11}} r_1 sin\phi_1 \end{aligned}$$

$$\begin{aligned} Z_{1I} &= -I_{1R}(n)r_1 r_2 sin(\phi_1 + \phi_2) - I_{1I}(n)r_1 r_2 cos(\phi_1 + \phi_2) \\ &\quad + \frac{n_{1I}(n)}{a_{11}} - \frac{n_{2R}}{a_{11}} - \frac{n_{2R}}{a_{11}} r_1 sin\phi_1 - \frac{n_{2I}}{a_{11}} r_1 cos\phi_1 \end{aligned} \tag{23}$$

The Chernoff bound on $P_1(|Z_{1R}| > c)$ is derived by [1] and some of the steps can be found detailed in section (5.2.1) From this analysis, we find that the probability of error is bounded as follows;

$$P_i(e) \leq (1 - \frac{1}{\sqrt{M}})\exp\left[\frac{-3}{2(M-1)}\frac{SNR\, a_{ii}}{1 + SNR\, r_1^2 r_2^2 a_{ii}^2 + r_i^2}\right] \; i = 1, 2 \qquad (24)$$

where we again use the relation (14);

Kavehrad in a different paper [11] uses another form of compensation;

From (21), he first finds the real and imaginary part of the canceler output,

$$\begin{aligned} y_{1R} &= a_{11}[1 - r_1 r_2 cos(\phi_1 + \phi_2)]I_{1R}(n) + a_{11} I_{1I}(n) r_1 r_2 sin(\phi_1 + \phi_2) + n_{1R}(n) \\ &\quad - n_{2R}(n) r_1 cos\phi_1 + n_{2I}(n) r_1 sin\phi_1, \\ y_{1I} &= a_{11}[1 - r_1 r_2 cos(\phi_1 + \phi_2)]I_{1I}(n) - a_{11} I_{1R}(n) r_1 r_2 sin(\phi_1 + \phi_2) + n_{1I}(n) \\ &\quad - n_{2R}(n) r_1 sin\phi_1 - n_{2I}(n) r_1 cos\phi_1, \end{aligned} \qquad (25)$$

then defines an estimate of the real part of $I_1(n)$,

$$\hat{I}_{1R}(n) = \frac{y_{1R}(n)}{a_{11}[1 - r_1 r_2 cos(\phi_1 + \phi_2)]}, \qquad (26)$$

and similar estimate for the imaginary part of $I_1(n)$. For the decision parameter Z_{1R}, then he uses

$$Z_{1R}(n) \triangleq \hat{I}_{1R}(n) - I_{1R}(n) \qquad (27)$$

Therefore,

$$Z_{1R}(n) = \frac{1}{\Delta}[a_{11} I_{1I}(n) r_1 r_2 sin(\phi_1 + \phi_2) + n_{1R}(n) - n_{2R}(n) r_1 cos\phi_1 + n_{2I}(n) sin\phi_1] \qquad (28)$$

where

$$\Delta \triangleq a_{11}[1 - r_1 r_2 cos(\phi_1 + \phi_2)]. \qquad (29)$$

This kind of compensation might be considered as "both amplitude and phase compensation" of the co-pol channel response. In the next chapter, we will use for the bootstrapped cancelers, a slightly different approach to Kavehrad's compensation; we will apply compensation on the complex output before making a decision. Obviously, there will be a difference in hardware needed to implement these approaches.

Both the Chernoff bound and the moment GQR can be used with (28). For the second approach one can find

$$P_1(e|\phi_1,\phi_2,I_{1I}) = 2(1-\frac{1}{\sqrt{M}})Q(\frac{c-X_I}{\sigma}) \tag{30}$$

where

$$\begin{aligned} X_I &= \frac{r_1 r_2 I_{1I} sin(\phi_1+\phi_2)}{\Delta(\phi_1,\phi_2)} \\ Y &= \frac{1}{\Delta(\phi_1,\varphi_2)}[n_{1R} - n_{2R} r_1 cos\phi_1 + n_{2I} r_1 sin\phi_1] \end{aligned} \tag{31}$$

and a variance

$$\sigma^2 = \frac{\sigma_n^2(1+r_1^2)}{[1-r_1 r_2 cos(\phi_1+\phi_2)]^2} \tag{32}$$

Let $\mathbf{x}$ be the random variable

$$\mathbf{x} = r_1 r_2[cos(\phi_1+\phi_2) - I_{1I} sin(\phi_1+\phi_2)] \tag{33}$$

then

$$P_1(e|\phi_1,o_2,I_{1I}) = 2(1-\frac{1}{\sqrt{M}})Q\left(\sqrt{\frac{3\,SNR}{(M-1)(1+r_1^2)}}(1-\mathbf{x})\right) \tag{34}$$

Equation (34) can be evaluated numerically. Some results on the performance of the diagonalizer of Kavehrad are given in the following figures. In Fig. 4, we use 16 QAM and compare the error probability obtained with the moment method to the Chernoff upper bound. The cross polarization used was r=-15 dB Fig. 5 depicts the same except for using 64 QAM instead. Comparing the performance when r=-10 dB to that when r=-15 dB is done in Fig. 6 using moment method.

4.2 LMS canceler

The structure of this canceler (Fig 7), for dually polarized non-dispersive channels is given in [12].

For the output of this canceler as in (18), Kavehrad performs an amplitude normalization and obtains an estimate of the transmitted signal $\hat{I}_i(n) = \frac{y_i(n)}{a_{ii}}$ $i = 1, 2$. The optimal LMS weights are found by minimizing the sum of the squares of the errors. $E\{|e_1(n)|^2 + |e_2(n)|^2\}$, where

$$e_i(n) \triangleq \hat{I}_1(n) - I_1(n) \quad i = 1, 2 \tag{35}$$

corresponding to the i-th output of the canceler.

These optimum weights are found by solving the matrix equation

$$\mathbf{R}\mathbf{w}_{\text{opt}} = \mathbf{S}^* \tag{36}$$

where

$$\mathbf{R} = \begin{bmatrix} |x_1(n)|^2 & x_1(n)x_2^*(n) \\ x_1(n)^* x_2(n) & |x_2(n)|^2 \end{bmatrix} \tag{37}$$

$$\mathbf{w}_{\text{opt}} = \begin{bmatrix} w_{11\text{opt}} & w_{12\text{opt}} \\ w_{21\text{opt}} & w_{22\text{opt}} \end{bmatrix}, \tag{38}$$

and

$$\mathbf{S} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \tag{39}$$

For $E\{I_1(n)\} = E\{I_2(n)\}$ and for a complex Gaussian noise, the optimum weights from (36) can be used in (18) to find the optimal output $y_1(n)$. (see appendix B of [1] for detail).

From the optimal output one can derive the decision parameter, and use it to find an approximation to, or upper bound on, the probability of error.

Some results of these calculations are given in Fig. 8 to 10.

5 The Power-Power Cross-Pol Canceler Structure

The system in Fig. 11 depicts the structure of the power-power cross-pol canceler. It consists of two distinct control loops: $Q - w_{21}$ loop and the $P - w_{12}$ loop. Let the power ratio of the two signals at point 3, be such that $I_1(n) > I_2(n)$. This being the input to the weight w_{12} results in power-inversion in $I_2(n) > I_1(n)$ at point 4. However, point 4, being the input to the weight w_{21}, results in $I_1(n)$ greater than $I_2(n)$ at point 3. This process continues resulting in a very high $\frac{I_1(n)}{I_2(n)}$ at one output and $\frac{I_2(n)}{I_1(n)}$ in the other. As a result, the power-power canceler acts as a high quality power separator. The adaptive control algorithm varies the cancellation the coefficient w_{12}, w_{21} so as to minimize the power P and Q at the output of the canceler. The blocks labeled "discrimination" perform functions which make the power detection more sensitive to the undesired ($a_{12}I_2$ at port P) signal than to the desired signal "$a_{11}I_1$". The effect of these blocks which will be discussed in later sections and prove to be essential for the bootstrapping operation.

5.1 Canceler Scheme and Parameters

As it was discussed in section 2 above, the received signals which are sampled after matched filters, are given by;

$$\begin{aligned} x_1(n) &= a_{11}I_1(n) + a_{12}I_2(n) + n_1(n) \\ x_2(n) &= a_{21}I_1(n) + a_{22}I_2(n) + n_2(n) \end{aligned} \tag{40}$$

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.

5.1.1 Canceler Outputs

From Fig. 11 the outputs $y_1(n)$ and $y_2(n)$ of the canceler are as follows

$$y_1(n) = x_1(n) + y_2(n)w_{12}$$

$$y_2(n) = x_2(n) + y_1(n)w_{21} \tag{41}$$

Solving the system of equation (41) leads to

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

$$y_2(n) = \frac{x_2(n) + x_1(n)w_{21}}{1 - w_{12}w_{21}} \tag{42}$$

Substituting for $x_1(n)$ and $x_2(n)$ from (40) we get

$$y_1(n) = \frac{I_1(n)(a_{11} + w_{12}a_{21}) + I_2(n)(a_{12} + w_{12}a_{22}) + n_1(n) + n_2(n)w_{12}}{1 - w_{12}w_{21}} \tag{43}$$

$$y_2(n) = \frac{I_1(n)(a_{21} + w_{21}a_{11}) + I_2(n)(a_{22} + w_{21}a_{12}) + n_1(n)w_{21} + n_2(n)}{1 - w_{12}w_{21}} \tag{44}$$

5.1.2 Optimal Weights

The control algorithm simultaneously minimizes the output powers

$$P(w_{12}^i, w_{21}^i) = E\{|y_{1d}^i(n)|^2\} \tag{45}$$

$$Q(w_{12}^i, w_{21}^i) = E\{|y_{2d}^i(n)|^2\} \tag{46}$$

where $y_{1d}(n)$ and $y_{2d}(n)$ are the samples of the corresponding output after the discriminations. 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 $E\{\cdot\}$ and $|\cdot|$ denote the expected and magnitude respectively. The search for optimum weights can be performed by successive use of the following recursive equations, provided that $1 - w_{12}w_{21} \neq 0$.

$$w_{12}^{i+1} = w_{12}^i - \mu_1 \frac{\partial}{\partial w_{12}^i} P(w_{12}^i, w_{21}^i) \tag{47}$$

$$w_{21}^{i+1} = w_{21}^{i} - \mu_2 \frac{\partial}{\partial w_{21}^{i}} Q(w_{12}^{i}, w_{21}^{i}) \tag{48}$$

where μ_1 and μ_2 are the constants which determine the stability of convergence.

The optimum weights that minimize the powers are the steady state weights obtained from

$$\frac{\partial P(w_{12}^{i}, w_{21}^{i})}{\partial w_{12}^{i}} = 0 \tag{49}$$

$$\frac{\partial Q(w_{12}^{i}, w_{21}^{i})}{\partial w_{21}^{i}} = 0 \tag{50}$$

From (43) and (44), we first find the powers at the output of the discriminators, as

$$\begin{aligned} P(w_{12}, w_{21}) &= \frac{1}{|1 - w_{12}w_{21}|^2}\Big[\delta_{11}E\{|I_1(n)|^2\}|a_{11} + w_{12}a_{21}|^2 \\ &\quad +\delta_{12}E\{|I_2(n)|^2\}|a_{12} + w_{12}a_{22}|^2 + E\{|n_1(n)|^2\} \\ &\quad +E\{|n_2(n)|^2\}|w_{12}|^2\Big], \end{aligned} \tag{51}$$

$$\begin{aligned} Q(w_{12}, w_{21}) &= \frac{1}{|1 - w_{12}w_{21}|^2}\Big[\delta_{21}E\{|I_1(n)|^2\}|a_{21} + w_{21}a_{11}|^2 \\ &\quad +\delta_{22}E\{|I_2(n)|^2\}|a_{22} + w_{21}a_{12}|^2 + E\{|n_1(n)|^2\}|w_{21}|^2 \\ &\quad +E\{|n_2(n)|^2\}\Big] \end{aligned} \tag{52}$$

where $\delta_{i,j}$ i.j=1.2 denotes the effect of the i th discriminator on the different signal ($I_1(n)$ or $I_2(n)$) powers.

Notice that in calculating the power, we assumed $I_1(n)$ and $I_2(n)$ are uncorrelated and zero mean. We will take for the derivative of any real function with respect to a complex variable [13].

$$\frac{\partial}{\partial w} = \frac{\partial}{\partial w_r} + j\frac{\partial}{\partial w_i} \tag{53}$$

where $w = w_r + jw_i$. Hence, for the functions P in (51), we get, after some algebraic manipulation (see Chapter 3 of [14]).

$$\begin{aligned}\frac{\partial P}{\partial w_{12}} &= \frac{2(1-w_{12}w_{21})}{|1-w_{12}w_{21}|^4}\Big[(a_{11}+w_{12}a_{21})(a_{21}+w_{21}a_{11})^*\delta_{11}E\{|I_1(n)|^2\} \\ &\quad +(a_{12}+w_{12}a_{22})(a_{22}+w_{21}a_{12})^*\delta_{12}E\{|I_2(n)|^2\} \\ &\quad +E\{|n_1(n)|^2\}w_{21}^* + E\{|n_2(n)|^2\}w_{12}\Big].\end{aligned} \tag{54}$$

Similarly the derivative of the power Q in (52) can be finally written as.

$$\begin{aligned}\frac{\partial Q}{\partial w_{21}} &= \frac{2(1-w_{12}w_{21})}{|1-w_{12}w_{21}|^4}\Big[(a_{21}+w_{21}a_{11})(a_{11}+w_{12}a_{21})^*\delta_{21}E\{|I_1(n)|^2\} \\ &\quad +(a_{12}+w_{12}a_{22})^*(a_{22}+w_{21}a_{12})\delta_{22}E\{|I_2(n)|^2\} \\ &\quad +E\{|n_1(n)|^2\}w_{21} + E\{|n_2(n)|^2\}w_{12}^*\Big]\end{aligned} \tag{55}$$

Provided $|1-w_{12}w_{21}| \neq 0$. equating (54) and (55) simultaneously to zero will result in $w_{12\text{opt}}$ and $w_{21\text{opt}}$.

$$w_{12\text{opt}} = \frac{-1}{Dw_{12\text{opt}}}\Big[a_{11}(a_{21}+w_{21\text{opt}}a_{11})^*E\{|I_1(n)|^2\}\delta_{11}$$

$$+a_{12}(a_{22}+w_{21\mathrm{opt}}a_{12})^{*}E\{|I_2(n)|^2\}\delta_{12}+w_{21\mathrm{opt}}^{*}E\{|n_1(n)|^2\}\Big] \quad (56)$$

where,

$$\begin{aligned} Dw_{12\mathrm{opt}} &= a_{21}(a_{21}+w_{21\mathrm{opt}}a_{11})^{*}E\{|I_1(n)|^2\}\delta_{11} \\ &\quad +a_{22}(a_{22}+w_{21\mathrm{opt}}a_{12})^{*}E\{|I_2(n)|^2\}\delta_{12}+E\{|n_2(n)|^2\}, \end{aligned} \quad (57)$$

and,

$$w_{21\mathrm{opt}} = \frac{-1}{Dw_{21\mathrm{opt}}}\Big[a_{21}(a_{11}+w_{12\mathrm{opt}}a_{21})^{*}E\{|I_1(n)|^2\}\delta_{21}$$

$$+a_{22}(a_{12}+w_{12\mathrm{opt}}a_{22})^{*}E\{|I_2(n)|^2\}\delta_{22}+w_{12\mathrm{opt}}^{*}E\{|n_2(n)|^2\}\Big] \quad (58)$$

where

$$\begin{aligned} Dw_{21\mathrm{opt}} &= a_{11}(a_{11}+w_{12\mathrm{opt}}a_{21})^{*}E\{|I_1(n)|^2\}\delta_{21} \\ &\quad +a_{12}(a_{12}+w_{12\mathrm{opt}}a_{22})^{*}E\{|I_2(n)|^2\}\delta_{22}+E\{|n_1(n)|^2\} \end{aligned} \quad (59)$$

The effect of the discriminators are presented by $\delta_{i,j}\quad i,j=1,2$ which are real valued satisfying $\delta_{11}\delta_{22}<\delta_{12}\delta_{21}$. Note that, the first and second terms in (55) are complex conjugates of the terms in (54). Therefore, to find a unique solution for w_{12} and w_{21} using these equations, discriminators which enforce the constant $\delta_{i,j}\quad i,j=1,2$ satisfying the above condition, are essential. The simultaneous solution of these non-linear equations give two equilibrium points; $[w_{12\mathrm{opt}1}, w_{21\mathrm{opt}1}]$ and $[w_{12\mathrm{opt}2}, w_{21\mathrm{opt}2}]$. One is a stable equilibrium which provides a solution to our problem.

5.1.3 Effect of Noise On Optimal Weights

In the absence of noise, that is when $E\{|n_1(n)|^2\} = E\{|n_2(n)|^2\} = 0$ the stable equilibrium points can easily found to be

$$w_{12\text{opt}} = -\frac{a_{12}}{a_{22}}, \qquad w_{21\text{opt}} = -\frac{a_{21}}{a_{11}}. \tag{60}$$

With noise, we will write

$$w_{12\text{opt}} = -\frac{a_{12}}{a_{22}} - \epsilon_1, \qquad w_{21\text{opt}} = -\frac{a_{21}}{a_{11}} + \epsilon_2. \tag{61}$$

where ϵ_1 and ϵ_2 are perturbations. due to noise on the optimal weights that we intend to find.

Using (56) and (60) in (61), we can finally write, after some algebraic manipulation (see chapter 3 of [14])

$$\begin{aligned}\epsilon_1 &= \frac{a_{22}}{\Delta_{\epsilon_1}}\Big[\Big(a_{11}^2(\frac{a_{12}}{a_{22}}\frac{a_{21}}{a_{11}} - 1)\delta_{11}E\{|I_1(n)|^2\} - E\{|n_1(n)|^2\}\Big)\epsilon_2^* \\ &\quad +(\frac{a_{21}}{a_{11}})^*E\{|n_1(n)|^2\} + \frac{a_{12}}{a_{22}}E\{|n_2(n)|^2\}\Big].\end{aligned} \tag{62}$$

where.

$$\begin{aligned}\Delta_{\epsilon_1} &= a_{22}\Big[\Big(a_{21}a_{11}\delta_{11}E\{|I_1(n)|^2\} + a_{22}a_{12}^*\delta_{12}E\{|I_2(n)|^2\}\Big)\epsilon_2^* \\ &\quad +a_{22}(a_{22} - \frac{a_{21}}{a_{11}}a_{12})^*\delta_{12}E\{|I_2(n)|^2\} + E\{|n_2(n)|^2\}\Big].\end{aligned} \tag{63}$$

Similar steps can be followed to determine ϵ_2 from (61) together with (58):

$$\begin{aligned}\epsilon_2 &= \frac{1}{\Delta_{\epsilon_2}}a_{11}\Big[\Big(a_{22}^2(\frac{a_{12}}{a_{22}}\frac{a_{21}}{a_{11}} - 1)\delta_{22}E\{|I_2(n)|^2\} - E\{|n_2(n)|^2\}\Big)\epsilon_1^* \\ &\quad +(\frac{a_{12}}{a_{22}})^*E\{|n_2(n)|^2\} + \frac{a_{21}}{a_{11}}E\{|n_1(n)|^2\}\Big].\end{aligned} \tag{64}$$

where,

$$\Delta_{\epsilon_2} = a_{11}\Big[\Big(a_{21}^{*}a_{11}\delta_{21}E\{|I_1(n)|^2\} + a_{12}a_{22}\delta_{22}E\{|I_2(n)|^2\}\Big)\epsilon_1^{*}$$

$$+a_{11}^2(1-\frac{a_{21}}{a_{11}}\frac{a_{11}}{a_{12}})^{*}\delta_{21}E\{|I_1(n)|^2\} + E\{|n_1(n)|^2\}\Big]. \quad (65)$$

Equation (62) and (64) can be solved simultaneously to obtain

$$\epsilon_1 = \frac{1}{\Delta}\Big(\Big[a_{11}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})E\{|I_1(n)|^2\}\delta_{21} + E\{|n_1(n)|^2\}\Big]\Big[(\frac{a_{21}}{a_{11}})^{*}E\{|n_1(n)|^2\}$$

$$+\frac{a_{12}}{a_{22}}E\{|n_2(n)|^2\}\Big] - \Big[a_{11}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})E\{|I_1(n)|^2\}\delta_{11} + E\{|n_1(n)|^2\}\Big].$$

$$\Big[(\frac{a_{21}}{a_{11}})^{*}E\{|n_1(n)|^2\} + \frac{a_{12}}{a_{22}}E\{|n_2(n)|^2\}\Big]\Big) \quad (66)$$

$$\epsilon_2 = \frac{-1}{\Delta^{*}}\Big[\Big[a_{22}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})E\{|I_2(n)|^2\}\delta_{22} + E\{|n_n(n)|^2\}\Big].\Big[\frac{a_{21}}{a_{11}}E\{|n_n(n)|^2\}$$

$$+(\frac{a_{12}}{a_{22}})^{*}E\{|n_n(n)|^2\}\Big] + \Big[a_{22}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})E\{|I_2(n)|^2\}\delta_{12} + E\{|n_n(n)|^2\}\Big]$$

$$\Big[\frac{a_{21}}{a_{11}}E\{|n_n(n)|^2\} + (\frac{a_{12}}{a_{22}})^{*}E\{|n_n(n)|^2\}\Big]\Big], \quad (67)$$

with

$$\Delta = \Big[a_{22}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})^{*}E\{|I_2(n)|^2\}\delta_{12} + E\{|n_n(n)|^2\}\Big]\Big[a_{11}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}}).$$

$$E\{|I_1(n)|^2\}\delta_{21} + E\{|n_n(n)|^2\}\Big] - \Big[a_{11}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})E\{|I_1(n)|^2\}\delta_{11}$$

$$+E\{|n_n(n)|^2\}\Big].\Big[a_{22}^2(1-\frac{a_{12}}{a_{11}}\frac{a_{21}}{a_{22}})^{*}E\{|I_2(n)|^2\}\delta_{22} + E\{|n_n(n)|^2\}\Big]. \quad (68)$$

Without loss of generality, we will assume that the noise variance in the V and H polarized channels are equal: $E\{|n_1(n)|^2\} = E\{|n_2(n)|^2\} = E\{|n_n(n)|^2\}$. This assumption is particularly useful when calculating the probability of error. In fact, this is the most difficult case that puts the worst requirement on the discriminators. Using this assumption and the definition of the channel parameter, we can write the final expression of the perturbation ϵ_1 and ϵ_2 as follows:

First Δ_R and Δ_I be the real and imaginary part of Δ, that is; $\Delta = \Delta_R + j\Delta_I$, then, some algebraic manipulation leads to,

$$\Delta_R = a_{11}^2 a_{22}^2 (1 - 2r_1 r_2 cos(\phi_1 - \phi_2) + r_1^2 r_2^2).E\{|I_1(n)|^2\}E\{|I_2(n)|^2\}$$

$$(\delta_{12}\delta_{21} - \delta_{11}\delta_{22}) + \Big[|a_{22}|^2 E\{|I_2(n)|^2\}E\{|n_n(n)|^2\}(\delta_{12} - \delta_{22}) + a_{11}^2 E\{|I_1(n)|^2\}.$$

$$E\{|n_n(n)|^2\}(\delta_{21} - \delta_{11})\Big].(1 - r_1 r_2 cos(\phi_1 + \phi_2)), \tag{69}$$

$$\Delta_I = \Big[a_{22}^2 E\{|I_2(n)|^2\}(\delta_{12} - \delta_{22}) - a_{11}^2 E\{|I_1(n)|^2\}(\delta_{21} - \delta_{11})\Big].$$

$$r_1 r_2 sin(\phi_1 + \phi_2) E\{|n_n(n)|^2\}. \tag{70}$$

Defining ;

$$\epsilon_1 = \frac{\epsilon_{1AR} + j\epsilon_{1AI}}{\Delta_R + j\Delta_I}, \tag{71}$$

then,

$$\epsilon_1 = \epsilon_{1R} + j\epsilon_{1I}$$

$$= \frac{1}{\Delta_R^2 + \Delta_I^2}\Big[(\epsilon_{1AR}\Delta_R + \epsilon_{1AI}\Delta_I) + j(\epsilon_{1AI}\Delta_R - \epsilon_{1AR}\Delta_I)\Big] \tag{72}$$

with

$$\epsilon_{1R} = \frac{\epsilon_{1AR}\Delta_R + \epsilon_{1AI}\Delta_I}{|\Delta|^2}, \tag{73}$$

$$\epsilon_{1I} = \frac{\epsilon_{1AI}\Delta_R - \epsilon_{1AR}\Delta_I}{|\Delta|^2}, \tag{74}$$

and

$$|\Delta|^2 = \Delta_R^2 + \Delta_I^2. \tag{75}$$

From (66) with the definition of the channel parameter one can easily conclude that,

$$\Big[r_1 cos\phi_1 + r_2 cos\phi_2 - r_1^2 r_2 cos(2\phi_1 + \phi_2) - r_1 r_2^2 cos\phi_1\Big].$$

$$E\{|n_n(n)|^2\}E\{|I_1(n)|^2\}(\delta_{21} - \delta_{11}) \tag{76}$$

$$\epsilon_{1AI} = a_{11}^2\Big[r_1 sin\phi_1 - r_2 sin\phi_2 - r_1^2 r_2 sin(2\phi_1 + \phi_2) - r_1 r_2^2 sin\phi_1\Big].$$

$$E\{|n_n(n)|^2\}E\{|I_1(n)|^2\}(\delta_{21} - \delta_{11}). \tag{77}$$

$$\epsilon_{2AI} = a_{22}^2\Big[-r_1 sin\phi_1 + r_2 sin\phi_2 - r_1^2 r_2 sin\phi_2 - r_1 r_2^2 sin(\phi_1 + 2\phi_2)\Big].$$

$$E\{|n_n(n)|^2\}E\{|I_2(n)|^2\}(\delta_{12} - \delta_{22}). \tag{78}$$

$$\epsilon_{2AR} = a_{22}^2\Big[r_1 cos\phi_1 + r_2 cos\phi_2 - r_1^2 r_2 cos\phi_2 - r_1 r_2^2 cos(\phi_1 + 2\phi_2)\Big].$$

$$E\{|n_n(n)|^2\}E\{|I_2(n)|^2\}(\delta_{12} - \delta_{22}), \tag{79}$$

Clearly,

$$\epsilon_2 = \epsilon_{2R} + j\epsilon_{2I} \tag{80}$$

$$= \frac{1}{\Delta_R^2 + \Delta_I^2}[(\epsilon_{2AR}\Delta_R - \epsilon_{2AI}\Delta_I) + j(\epsilon_{2AR}\Delta_I + \epsilon_{2AI}\Delta_R)], \tag{81}$$

then,

$$\epsilon_{2R} = \frac{\epsilon_{2AR}\Delta_R - \epsilon_{2AI}\Delta_I}{|\Delta|^2} \tag{82}$$

$$\epsilon_{2I} = \frac{\epsilon_{2AR}\Delta_I + \epsilon_{2AI}\Delta_R}{|\Delta|^2} \tag{83}$$

5.1.4 Canceler Optimal Outputs

Next using (61) in (43) with $w_{12} = w_{12\text{opt}}$ and $w_{21} = w_{21\text{opt}}$, we obtain $y_1(n)$

$$\begin{aligned} y_1(n) &= \frac{1}{\left(1 - (\frac{-a_{12}}{a_{22}} + \epsilon_1)(\frac{-a_{21}}{a_{11}} + \epsilon_2)\right)}\Big[I_1(n)[a_{11} + (\frac{-a_{12}}{a_{22}} + \epsilon_1)a_{21}] \\ &\quad + I_2(n)[a_{12} + (-\frac{a_{12}}{a_{22}} + \epsilon_1)a_{22}] + n_1(n) + n_2(n)[\frac{-a_{12}}{a_{22}} + \epsilon_1]\Big], \end{aligned} \tag{84}$$

and after combining terms,

$$y_1(n) = \frac{a_{11}I_1(n)[1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \epsilon_1\frac{a_{21}}{a_{11}}] + I_2(n)\epsilon_1 a_{22} + n_1(n) - n_2(n)\frac{a_{12}}{a_{22}} + n_2(n)\epsilon_1}{1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \frac{a_{12}}{a_{22}}\epsilon_2 + \frac{a_{21}}{a_{11}}\epsilon_1 - \epsilon_1\epsilon_2} \tag{85}$$

5.1.5 Decision Parameters with Amplitude Compensation at the Canceler Output

The co-pole horizontally polarized signal at the output of the channel is $a_{11}I_1(n)$ and hence, it is reasonable to take $y_1(n)a_{11}$ as an estimate of this signal by compensating for the attenuation in the co-pol by a_{11}. Therefore, we will take $\hat{I}_1(n) = \frac{y_1(n)}{a_{11}}$ as an estimate of the transmitted signal $I_1(n)$. We will also assume the $\epsilon_1\epsilon_2$ in (85) is negligible with respect to the other terms in the denominator of this equation. Hence from (85) we can write,

$$\hat{I}_1(n)[1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \frac{a_{12}}{a_{22}}\epsilon_2 + \frac{a_{21}}{a_{11}}\epsilon_1] = I_1(n)[1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \frac{a_{21}}{a_{11}}\epsilon_1 + \frac{a_{12}}{a_{22}}\epsilon_2 - \frac{a_{12}}{a_{22}}\epsilon_2]$$

$$+I_2(n)\epsilon_1\frac{a_{22}}{a_{11}} + \frac{n_1(n)}{a_{11}} + n_2(n)(\frac{-a_{12}}{a_{11}a_{22}} + \frac{\epsilon_1}{a_{11}}) \tag{86}$$

Define,

$$Z_1(n) \triangleq \hat{I}_1(n) - I_1(n), \tag{87}$$

with $Z_1(n)$ is taken as the decision parameter. That is, the probability of error is given by $P_1(\epsilon) = P\{|Z_1(n)| > c\}$ where c is the half of the distance between two signals in the corresponding signal space. From (86) together with (87), we can drive, after lengthy algebraic manipulations (see chapter 3 of [14]), the real and imaginary part of the decision variable $Z_1(n)$;

$$\begin{aligned} Z_{1R} = \frac{1}{|\Delta_Z|^2}\Big[& I_{1R}(K_R Z_{DR} + K_I Z_{DI}) + I_{1I}(K_I Z_{DR} - K_R Z_{DI}) \\ & + I_{2R}(\epsilon_{1R}Z_{DR} + \epsilon_{1I}Z_{DI})\frac{a_{22}}{a_{11}} + I_{2I}(\epsilon_{1R}Z_{DI} - \epsilon_{1I}Z_{DR})\frac{a_{22}}{a_{11}} \\ & + n_{1R}\frac{Z_{DR}}{a_{11}} + n_{1I}\frac{Z_{DI}}{a_{11}} \\ & + n_{2R}[(\frac{\epsilon_{1R}}{a_{11}} - \frac{r_1 cos\phi_1}{a_{11}})Z_{DR} + (\frac{\epsilon_{1I}}{a_{11}} - \frac{r_1 sin\phi_1}{a_{11}})Z_{DI}] \\ & + n_{2I}[(\frac{-\epsilon_{1I}}{a_{11}} + \frac{r_1 sin\phi_1}{a_{11}})Z_{DR} + (\frac{\epsilon_{1R}}{a_{11}} - \frac{r_1 cos\phi_1}{a_{11}})Z_{DI}]\Big]. \end{aligned} \tag{88}$$

and

$$
\begin{aligned}
Z_{1I} &= \frac{1}{|\Delta_Z|^2}\Big[-I_{1R}(K_I Z_{DR} - K_R Z_{DI}) - I_{1I}(K_R Z_{DR} + K_I Z_{DI}) \\
&\quad -I_{2R}(\epsilon_{1R} Z_{DI} - \epsilon_{1I} Z_{DR})\frac{a_{22}}{a_{11}} + I_{2I}(\epsilon_{1R} Z_{DR} + \epsilon_{1I} Z_{DI})\frac{a_{22}}{a_{11}} \\
&\quad -n_{1R}\frac{Z_{DI}}{a_{11}} + n_{1I}\frac{Z_{DR}}{a_{11}} \\
&\quad -n_{2R}[(\frac{-\epsilon_{1I}}{a_{11}} + \frac{r_1 sin\phi_1}{a_{11}})Z_{DR} + (\frac{\epsilon_{1R}}{a_{11}} - \frac{r_1 cos\phi_1}{a_{11}})Z_{DI}] \\
&\quad +n_{2I}[(\frac{\epsilon_{1R}}{a_{11}} - \frac{r_1 cos\phi_1}{a_{11}})Z_{DR} + (\frac{\epsilon_{1I}}{a_{11}} - \frac{r_1 sin\phi_1}{a_{11}})Z_{DI}]\Big] \qquad (89)
\end{aligned}
$$

where K_R and K_I are the real and imaginary part of K given by

$$K_R = \frac{K_{AR}\Delta_R - K_{AI}\Delta_I}{\Delta_R^2 + \Delta_I^2} \qquad (90)$$

$$K_I = \frac{K_{AR}\Delta_I + K_{AI}\Delta_R}{\Delta_R^2 + \Delta_I^2} \qquad (91)$$

K_{AR} and K_{AI} are given by

$$
\begin{aligned}
K_{AR} &= a_{22}^2 E\{|n_n(n)|^2\}E\{|I_2(n)|^2\}(\delta_{12} - \delta_{22}). \\
&\quad \Big[r_1^2 + r_1 r_2 cos(\phi_1 + o_2) - r_1^3 r_2 cos(\phi_1 + \phi_2) - r_1^2 r_2^2 cos(2(\phi_1 + \phi_2))\Big] \qquad (92)
\end{aligned}
$$

$$
\begin{aligned}
K_{AI} &= a_{22}^2 E\{|n_n(n)|^2\}E\{|I_2(n)|^2\}(\delta_{12} - \delta_{22}). \\
&\quad \Big[r_1 r_2 sin(\phi_1 + \phi_2) - r_1^3 r_2 sin(\phi_1 + \phi_2) - r_1^2 r_2^2 sin(2(\phi_1 + \phi_2))\Big] \qquad (93)
\end{aligned}
$$

while Δ_R and Δ_I are given in (69,70). Note that K_R and K_I are functions of the random variables ϕ_1 and ϕ_2. Also, the real and imaginary part of the denominator $Z_D = Z_{DR} + jZ_{DI}$ is given by

$$Z_{DR} = 1 - r_1 r_2 cos(\phi_1 + \phi_2) + K_R + V_R \tag{94}$$

$$Z_{DI} = -r_1 r_2 sin(\phi_1 + \phi_2) + K_I + V_I. \tag{95}$$

$|\Delta_Z|^2 = Z_{DR}^2 + Z_{DI}^2$. V_{AR} and V_{AI} are the real and imaginary part of V defined as follows

$$V_R = \frac{V_{AR}\Delta_R + V_{AI}\Delta_I}{\Delta_R^2 + \Delta_I^2} \tag{96}$$

$$V_I = \frac{V_{AI}\Delta_R - V_{AR}\Delta_I}{\Delta_R^2 + \Delta_I^2} \tag{97}$$

with V_{AR} and V_{AI} are given by

$$V_{AR} = a_{11}^2 E\{|n_n(n)|^2\} E\{|I_1(n)|^2\}(\delta_{21} - \delta_{11})\Big[r_1 r_2 cos(\phi_1 + \phi_2) + r_2^2$$

$$-r_1 r_2^3 cos(\phi_1 + \phi_2) - r_1^2 r_2^2 cos(2(\phi_1 + \phi_2))\Big] \tag{98}$$

$$V_{AI} = a_{11}^2 E\{|n_n(n)|^2\} E\{|I_1(n)|^2\}(\delta_{21} - \delta_{11})\Big[r_1 r_2 sin(\phi_1 + \phi_2)$$

$$-r_1 r_2^3 sin(\phi_1 + \phi_2) - r_1^2 r_2^2 sin(2(\phi_1 + \phi_2))\Big] \tag{99}$$

and Δ_R and Δ_I are given in (69). Similar to the real and imaginary part of K, V_R and V_I are functions of the random variables ϕ_1 and ϕ_2.

Finally we write the real and imaginary parts of $Z_1(n)$ in terms of the random variable representing the real and imaginary part of signal and noises of channel 1

$$Z_{1R} = I_{1R}Y_1 + I_{1I}Y_2 + I_{2R}Y_3 + I_{2I}Y_4 + n_{1R}Y_5 + n_{1I}Y_6 + n_{2R}Y_7 + n_{2I}Y_8 \quad (100)$$

$$Z_{1I} = -I_{1R}Y_2 + I_{1I}Y_1 - I_{2R}Y_4 + I_{2I}Y_3 - n_{1R}Y_6 + n_{1I}Y_5 - n_{2R}Y_8 + n_{2I}Y_7 \quad (101)$$

where

$$Y_1 = -\frac{K_R Z_{DR} + K_I Z_{DI}}{|\Delta_Z|^2} \quad (102)$$

$$Y_2 = \frac{K_I Z_{DR} - K_R Z_{DI}}{|\Delta_Z|^2} \quad (103)$$

$$Y_3 = \frac{a_{22}}{a_{11}} \frac{\epsilon_{1R} Z_{DR} + \epsilon_{1I} Z_{DI}}{|\Delta_Z|^2} \quad (104)$$

$$Y_4 = \frac{a_{22}}{a_{11}} \frac{\epsilon_{1R} Z_{DI} - \epsilon_{1I} Z_{DR}}{|\Delta_Z|^2} \quad (105)$$

$$Y_5 = \frac{Z_{DR}}{a_{11}|\Delta_Z|^2} \quad (106)$$

$$Y_6 = \frac{Z_{DI}}{a_{11}|\Delta_Z|^2} \quad (107)$$

$$Y_7 = \frac{(\epsilon_{1R} - r_1 cos\phi_1) Z_{DR} + (\epsilon_{1I} - r_1 sin\phi_1) Z_{DI}}{a_{11}|\Delta_Z|^2} \quad (108)$$

$$Y_8 = \frac{(r_1 sin\phi_1 - \epsilon_{1I}) Z_{DR} + (\epsilon_{1R} - r_1 cos\phi_1) Z_{DI}}{a_{11}|\Delta_Z|^2} \quad (109)$$

$$|\Delta_Z|^2 = Z_{DR}^2 + Z_{DI}^2 \quad (110)$$

Notice that Y_i, $i = 1.2..8$ depend only on the random variables ϕ_1 and ϕ_2.

5.1.6 Decision Parameter with Both Amplitude and Phase Compensation at the Canceler Output

Instead of the amplitude compensation used in section 5.1.5, in this section we use compensation on both amplitude and phase of the co-pol signal. That is at the output of the canceler, the co-pol signal is the same as that sent by the transmitter and the error will be caused only by the cross coupling and the noise processes.

From (85) we can write

$$y_1(n) = \frac{1}{\Delta_I}\Big[I_1(n)\Delta_y + I_2(n)\epsilon_1 a_{22} + n_1(n) + n_2(n)[-\frac{a_{12}}{a_{22}} + \epsilon_1]\Big] \qquad (111)$$

where

$$\Delta_I = 1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \epsilon_2\frac{a_{12}}{a_{22}} + \epsilon_1\frac{a_{21}}{a_{11}} - \epsilon_1\epsilon_2$$

$$\Delta_y = a_{11}[1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \epsilon_1\frac{a_{21}}{a_{11}}], \qquad (112)$$

and with amplitude and phase compensation, we take $\hat{I}_1(n) = y_1(n)\frac{\Delta_I}{\Delta_y}$ as estimate of the transmitted signal $I_1(n)$.

Define.

$$Z_1(n) \triangleq \hat{I}_1(n) - I_1(n) \qquad (113)$$

with Z_1 denotes the amplitude and phase compensated output decision variable at channel 1.

We will perform an analysis similar to that in the previous section to find the real and imaginary parts of the decision variable $Z_1(n)$.

From (111) and (113) we get.

$$Z_1(n) = \frac{1}{\Delta_y}\Big[I_2(n)\epsilon_1 a_{22} + n_1(n) + n_2(n)[-\frac{a_{12}}{a_{22}} + \epsilon_1]\Big] \qquad (114)$$

Using (6) in (114), we get the decision variable.

$$Z_1 = \frac{1}{\Delta_y}\Big[(I_{2R}+jI_{2I})(\epsilon_{1R}+j\epsilon_{1I})a_{22}$$

$$+n_{1R}+jn_{1I}-r_1(n_{2R}+jn_{2I})(cos\phi_1+jsin\phi_1)$$

$$+(n_{2R}+jn_{2I})(\epsilon_{1R}+j\epsilon_{1I})\Big], \tag{115}$$

where we drop the dependence of terms on the sampling time n.

Also from (112) the real and imaginary part of the denominator Δ_y are respectively given by,

$$\Delta_{yR} = [1-r_1r_2cos(\phi_1+o_2)+V_R]a_{11} \tag{116}$$

$$\Delta_{yI} = [-r_1r_2sin(\phi_1+\phi_2)+V_I]a_{11} \tag{117}$$

where V_R and V_I are defined in the previous section. Clearly Δ_{yR} and Δ_{yI} are functions of the random variables ϕ_1 and ϕ_2.

The real and imaginary part of numerator of (115), Z_N are given by

$$Z_{NR} = (I_{2R}\epsilon_{1R}-I_{2I}\epsilon_{1I})a_{22}$$

$$+n_{1R}+n_{2R}(\epsilon_{1R}-r_1cos\phi_1)-n_{2I}(\epsilon_{1I}-r_1sin\phi_1) \tag{118}$$

$$Z_{NI} = (I_{2R}\epsilon_{1I}+I_{2I}\epsilon_{1R})a_{22}$$

$$+n_{1I}+n_{2R}(\epsilon_{1I}-r_1sin\phi_1)+n_{2I}(\epsilon_{1R}-r_1cos\phi_1) \tag{119}$$

Finally, we can write (115) as.

$$Z_1 = Z_{1R} + jZ_{1I} = \frac{Z_{NR} + jZ_{NI}}{\Delta_{yR} + j\Delta_{yI}} \tag{120}$$

with

$$Z_{1R} = \frac{Z_{NR}\Delta_{yR} + Z_{NI}\Delta_{yI}}{\Delta_{yR}^2 + \Delta_{yI}^2} \tag{121}$$

$$Z_{1I} = \frac{Z_{NI}\Delta_{yR} - Z_{NR}\Delta_{yI}}{\Delta_{yR}^2 + \Delta_{yI}^2} \tag{122}$$

Using (116), (117), (118) and (119) in (121) and (122) we can get, after some simplification which emphasizes the dependency of the different terms on the different random variables, the real part of the decision variable:

$$\begin{aligned} Z_{1R} = \frac{1}{|\Delta_y|^2}\Big[& I_{2R}(\epsilon_{1R}\Delta_{yR} + \epsilon_{1I}\Delta_{yI})a_{22} + I_{2I}(\epsilon_{1R}\Delta_{yI} - \epsilon_{1I}\Delta_{yR})a_{22} \\ & + n_{1R}\Delta_{yR} + n_{1I}\Delta_{yI} \\ & + n_{2R}[(\epsilon_{1R} - r_1 cos\phi_1)\Delta_{yR} + (\epsilon_{1I} - r_1 sin\phi_1)\Delta_{yI}] \\ & + n_{2I}[(-\epsilon_{1I} + r_1 sin\phi_1)\Delta_{yR} + (\epsilon_{1R} - r_1 cos\phi_1)\Delta_{yI}]\Big] \end{aligned} \tag{123}$$

Similar expression can be found for Z_{1I};

$$\begin{aligned} Z_{1I} = \frac{1}{|\Delta_y|^2}\Big[& -I_{2R}(\epsilon_{1R}\Delta_{yI} - \epsilon_{1I}\Delta_{yR})a_{22} + I_{2I}(\epsilon_{1R}\Delta_{yR} + \epsilon_{1I}\Delta_{yI})a_{22} \\ & - n_{1R}\Delta_{yI} + n_{1I}\Delta_{yR} \end{aligned}$$

$$-n_{2R}[(-\epsilon_{1I}+r_1 sin\phi_1)\Delta_{yR}+(\epsilon_{1R}-r_1 cos\phi_1)\Delta_{yI}$$
$$+n_{2I}[(\epsilon_{1R}-r_1 cos\phi_1)\Delta_{yR}+(\epsilon_{1I}-r_1 sin\phi_1)\Delta_{yI}]\Big] \quad (124)$$

with $|\Delta_y|^2 = \Delta_{yR}^2 + \Delta_{yI}^2$

Finally we write the real and imaginary parts of $Z_1(n)$ in terms of the random variable representing the real and imaginary part of signal and noises of channel 1

$$Z_{1R} = I_{2R}Y_{1AP} + I_{2I}Y_{2AP} + n_{1R}Y_{3AP} + n_{1I}Y_{4AP} + n_{2R}Y_{5AP} + n_{2I}Y_{6AP} \quad (125)$$

$$Z_{1I} = -I_{2R}Y_{2AP} + I_{2I}Y_{1AP} - n_{1R}Y_{4AP} + n_{1I}Y_{3AP} - n_{2R}Y_{6AP} + n_{2I}Y_{5AP} \quad (126)$$

where

$$Y_{1AP} = a_{22}\frac{\epsilon_{1R}\Delta_{yR}+\epsilon_{1I}\Delta_{yI}}{|\Delta_y|^2} \quad (127)$$

$$Y_{2AP} = a_{22}\frac{\epsilon_{1R}\Delta_{yI}-\epsilon_{1I}\Delta_{yR}}{|\Delta_y|^2} \quad (128)$$

$$Y_{3AP} = \frac{\Delta_{yR}}{|\Delta_y|^2} \quad (129)$$

$$Y_{4AP} = \frac{\Delta_{yI}}{|\Delta_y|^2} \quad (130)$$

$$Y_{5AP} = \frac{(\epsilon_{1R}-r_1 cos\phi_1)\Delta_{yR}+(\epsilon_{1I}-r_1 sin\phi_1)\Delta_{yI}}{|\Delta_y|^2} \quad (131)$$

$$Y_{6AP} = \frac{(r_1 sin\phi_1-\epsilon_{1I})\Delta_{yR}+(\epsilon_{1R}-r_1 cos\phi_1)\Delta_{yI}}{|\Delta_y|^2} \quad (132)$$

$$|\Delta_y|^2 = \Delta_{yR}^2 + \Delta_{yI}^2 \quad (133)$$

5.2 The Performance Analysis

5.2.1 Chernoff Bound

With Amplitude Compensation

Using the real and imaginary parts of the decision variable for channel 1 (100) and (101) obtained under the assumption of amplitude compensation, we calculate an upper bound for the average symbol error probability for the power-power scheme of BXPC with dual-polarized M-ary QAM system.

An error is made on this channel if the decision variable $|Z_{1R}| > c$ or $|\quad| > c$. The probability of error on channel 1 can be written [1] as,

$$P_1(e) = \frac{1}{2}\left[1 - \frac{1}{\sqrt{M}}\right]\{P(|Z_{1R}| > c) + P(|Z_{1I}| > c)\} \tag{134}$$

For a bound on the probability, $P(|Z_{1R}| > c)$ or $P(|Z_{1I}| > c)$ we will use the Chernoff bound [1]. Such a bound is defined as follows: for any random variable Z and a constant c, one can find a $\lambda \geq 0$ such that

$$P(Z > c) \leq E\{e^{\lambda(Z-c)}\} \quad \lambda \geq 0 \tag{135}$$

Obviously λ that minimizes the right hand side of (135) establishes the least upper bound on $P(Z > c)$. Using (100) in (135) we find

$$P(|Z_{1R}| > c) \leq e^{-\lambda c} E_{\phi_1,\phi_2}\Big\{ E_{I_{1R}}[\exp(\lambda I_{1R} Y_1)].E_{I_{1I}}[\exp(\lambda I_{1I} Y_2)].$$

$$E_{I_{2R}}[\exp(\lambda I_{2R} Y_3)].E_{I_{2I}}[\exp(\lambda I_{2I} Y_4)].$$

$$E_{n_{1R}}[\exp(\lambda n_{1R} Y_5)].E_{n_{1I}}[\exp(\lambda n_{1I} Y_6)].$$

$$E_{n_{2R}}[\exp(\lambda n_{2R}Y_7)].E_{n_{2I}}[\exp(\lambda n_{2I}Y_8)]\Big\}, \qquad (136)$$

where we used the fact that I_{iR}, I_{iI}, n_{iR} and n_{iI}, $i = 1, 2$ are independent of each other. Also note that all the expected value operations inside the large parenthesis, are conditional on ϕ_1 and ϕ_2, and hence the random function Y_i $i = 1, ..8$ conditioned on ϕ_1 and ϕ_2 are constant with respect to these operations.

Following Kavehrad [1], we derive these expected values: The random variable I_{1R} is a discrete M-ary random variable which takes the values $\{\pm 1c. \pm 3c. ..\pm(\sqrt{(M)}-1)c\}$ with equal probability. For such a random variable, we derive in appendix A an upper bound on $E\{\exp(aI_{1R})\}$, with a given constant a.(see detail in [14]).

$$\begin{aligned} E_{I_{1R}}[\exp(\lambda I_{1R}Y_1)] &= \frac{2}{\sqrt{M}} \sum_{i=1}^{\sqrt{M}/2} cosh[(\lambda Y_1)(2i-1)c] \\ &\leq \exp(\frac{\lambda^2}{2}c^2\frac{M-1}{3}Y_1^2) \end{aligned} \qquad (137)$$

and terms are in effect for the other terms in (136). The additive noise $n_1(n)$ and $n_2(n)$ are assumed to be independent samples of zero mean complex Gaussian random variables with $E\{|n_i(n)|^2\} = 2\sigma_n^2$ $i = 1, 2$. Therefore, n_{iR} and n_{iI} $i = 1, 2$ are real, zero mean Gaussian random variable with variance$= \sigma_n^2$. For such a random variable $\mathbf{n}$, we can derive the value of $E\{\exp(a\mathbf{n})\}$ with a as given constant. Again noting that conditioned on ϕ_1 and ϕ_2, λY_5 is a constant, and hence

$$E_{n_{1R}}[\exp(\lambda n_{1R}Y_5)] = \exp(\frac{\lambda^2}{2}c^2\sigma_n^2Y_5^2). \qquad (138)$$

and terms are in effect for the other terms in (136). Finally, by collecting term, we get

$$P(|Z_{1R}| > c) \leq E_{\phi_1,\phi_2}\{\exp(-\lambda c + \lambda^2[U(\phi_1, \phi_2) + W(\phi_1, \phi_2)])\} \qquad (139)$$

where

$$U(\phi_1,\phi_2) = \frac{c^2}{2}\frac{M-1}{3}(Y_1^2+Y_2^2+Y_3^2+Y_4^2)$$
$$W(\phi_1,\phi_2) = \frac{\sigma_n^2 c^2}{2}(Y_5^2+Y_6^2+Y_7^2+Y_8^2) \tag{140}$$

Minimizing the exponent of (139) with respect to λ, we obtain by using (140) the least upper bound

$$P(|Z_{1R}|>c) \le E_{\phi_1,\phi_2}\left\{\exp\left(\frac{-c^2}{4[\frac{c^2}{2}\frac{M-1}{3}U_1(\phi_1,\phi_2)+\frac{\sigma_n^2}{2}W_1(\phi_1,\phi_2)]}\right)\right\} \tag{141}$$

where.

$$U_1(\phi_1,\phi_2) = Y_1^2+Y_2^2+Y_3^2+Y_4^2$$
$$W_1(\phi_1,\phi_2) = Y_5^2+Y_6^2+Y_7^2+Y_8^2 \tag{142}$$

Rearranging terms, we get

$$\begin{aligned} P(|Z_{1R}|>c) &\le E_{\phi_1,\phi_2}\left\{\exp\left[\frac{\frac{-c^2}{\sigma_n^2}}{2[c^2\frac{M-1}{3\sigma_n^2}U_1(\phi_1,\phi_2)+W_1(\phi_1,\phi_2)]}\right]\right\} \\ &= E_{\phi_1,\phi_2}\left\{\exp\left[\frac{-3(SNR)}{2(M-1)[(SNR)U_1(\phi_1,\phi_2)+W_1(\phi_1,\phi_2)]}\right]\right\}, \end{aligned} \tag{143}$$

where in the last step, we used (see [17])

$$SNR = \frac{S}{N} = \frac{M-1}{3}\frac{c^2}{\sigma_n^2} \tag{144}$$

Due to symmetry $P(|Z_{1R}|>c) = P(|Z_{1I}|>c)$. By using the minimum upper bound on $P(|Z_{1R}|>c)$ from (143) in (139), we can write the resulting least upper bound on the probability error

$$P_1(e) \leq (1-\frac{1}{\sqrt{M}})\frac{1}{4\pi^2}\cdot \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\exp\left[\frac{-3\,(SNR)}{2(M-1)}\frac{1}{(SNR)U_1(o_1.\phi_2)+W_1(\phi_1,\phi_2)}\right]d\phi_1 d\phi_2 \tag{145}$$

where we used the fact that probability density function of ϕ_i $i=1,2$ are, $p(\phi_1)=p(\phi_2)=\frac{1}{2\pi}$.

With Amplitude and Phase Compensation

To calculate the Chernoff bound in the case where the decision variable is obtained with both amplitude and phase compensation, we follow the same steps as in the previous section except with a different value of Y_i $i=1,2..,6$.

As in (139) we have here

$$P(|Z_{1R}|>c) \leq E_{\phi_1,\phi_2}\Big\{\exp[-\lambda c+\frac{\lambda^2c^2}{2}\frac{M-1}{3}(Y_{1AP}^2+Y_{2AP}^2) + \frac{\lambda^2c^2}{2}\sigma_n^2(Y_{3AP}^2+Y_{4AP}^2+Y_{5AP}^2+Y_{6AP}^2)]\Big\} \tag{146}$$

where Y_{iAP} $i=1.2...6$ are defined by (127) to (132).

Minimizing the right hand side of (146) and taking the expected value over ϕ_1 and ϕ_2, we write error the bound for the amplitude and phase compensated channel 1 output,

$$P_{1AP}(e) \leq (1-\frac{1}{\sqrt{M}})\frac{1}{4\pi^2}\cdot \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\exp\left[\frac{-3\,(SNR)}{2(M-1)}\frac{1}{(SNR)U_{1AP}(o_1.o_2)+W_{1AP}(o_1,o_2)}\right]d\phi_1 d\phi_2 \tag{147}$$

where

$$U_{1AP}(\phi_1,\phi_2) = Y_{1AP}^2 + Y_{2AP}^2$$

$$W_{1AP}(\phi_1,\phi_2) = Y_{3AP}^2 + Y_{4AP}^2 + Y_{5AP}^2 + Y_{6AP}^2 \tag{148}$$

and

$$SNR = \frac{S}{N} = \frac{M-1}{3}\frac{c^2}{\sigma_n^2} \tag{149}$$

5.2.2 Method of Moments for Probability of Error Calculation

In some cases, the Chernoff bound might not be sufficiently tight [17]. Therefore to use it as a measure of performance in comparing different systems might not be adequate. Hence, in this section. we will present another method which actually computes an approximation rather than a bound for the average probability of error for the power-power scheme of BNPC. The method is based upon Gauss quadrature rules (GQR) which were shown to assure accurate and satisfactory results. We will first give a brief description of GQR and apply it to calculate the average probability of error of the power-power scheme. These calculations will be performed for both amplitude compensated, and amplitude and phase compensated received signals. respectively.

With Amplitude Compensation

Using (100) in (134), we find the conditional probability,

$P(|Z_{1R}| > c|\phi_1,\phi_2,I_{1R},I_{1I},I_{2R},I_{2I})$. For this we define $Z_{1R} = X_I + Y$ with the random variable $Y = n_{1R}Y_5 + n_{1I}Y_6 + n_{2R}Y_7 + n_{2I}Y_8$.

The random variable Y is zero mean Gaussian and have variance

$$\sigma^2(\phi_1,\phi_2) = (Y_5^2 + Y_6^2 + Y_7^2 + Y_8^2)\sigma_n^2 \tag{150}$$

Conditioned on $\phi_1,\phi_2,I_{1R},I_{1I}$, I_{2R} and I_{2I} , the random variable Z_{1R} is Gaussian with mean equals X_I and variance $\sigma^2(\phi_1,\phi_2)$. Therefore

$$P(|Z_{1R}| > c|\phi_1,\phi_2,I_{1R},I_{1I},I_{2R},I_{2I}) = \frac{2}{\sqrt{2\pi}}\int_c^\infty \exp\left(\frac{-1}{2}\left[\frac{z_{1R}-X_I}{\sigma}\right]^2\right)dz_{1R}$$

Hence.

$$P(|Z_{1R}| > c|\phi_1, \phi_2, I_{1R}, I_{1I}, I_{2R}, I_{2I}) \;=\; 2Q(\frac{c - X_I}{\sigma}) \tag{151}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(\frac{-t^2}{2})\, dt \tag{152}$$

Again due to symmetry $P(|Z_{1R}| > c)$= $P(|Z_{1I}| > c)$, so that together with (134), we can write

$$P_1(e|\mathbf{x}) = 2(1 - \frac{1}{\sqrt{M}})Q(\frac{c}{\sigma_n}\mathbf{x}), \tag{153}$$

where the random variable $\mathbf{x}$;

$$\mathbf{x} = \frac{c - X_I}{\sigma_o} \tag{154}$$

with

$$\sigma_o^2 = (Y_5^2 + Y_6^2 + Y_7^2 + Y_8^2)c^2, \tag{155}$$

is a function of the random variables $(\phi_1, \phi_2, I_{1R}, I_{1I}, I_{2R}, I_{2I})$.

Clearly, the average error probability on channel 1 can be evaluated from

$$P_1(e) = \int_x P_1(e|x) f_{\mathbf{x}}(x)\, dx \tag{156}$$

with $f_{\mathbf{x}}(x)$ as the pdf of the random variable $\mathbf{x}$.

Using the method of moments in calculating (156), we get by using (149)

$$P_1(e) = 2(1 - \frac{1}{\sqrt{M}}) \sum_{i=1}^{N} w_i Q\left(\sqrt{\frac{3(SNR)}{M-1}} x_i\right) \tag{157}$$

The GQR nodes x_i and the weights w_i are determined from the moments of random variable $\mathbf{x}$.

Using the Gauss Quadrature integration, the average probability of error in (157) can be calculated numerically by evaluating the $2N + 1$ moments of random variable $\mathbf{x}$ in (154). One can derive a general equation for the moment of $\mathbf{x}$.

$$E\{\mathbf{x^n}\} = \sum_{k=0}^{n} \binom{n}{k} c^k (-1)^{n-k} \sum_{m=0}^{n-k} \binom{n-k}{m} E\{A_x B_x\} \tag{158}$$

where

$$A_x = \sum_{l=0}^{m} \binom{m}{l} I_{1R}^{l} \frac{Y_1^l}{\sigma_o^n} I_{1I}^{(m-l)} \frac{Y_2^{(m-l)}}{\sigma_o^n}$$

$$B_x = \sum_{u=0}^{n-k-m} \binom{n-k-m}{u} I_{2R}^{u} \frac{Y_3^u}{\sigma_o^n} I_{2I}^{(n-k-m-u)} \frac{Y_4^{(n-k-m-u)}}{\sigma_o^n} \tag{159}$$

Recall that I_{ij}, $i = 1,2$, $j = R,I$ are all independent, equally likely M-ary symbols from the set $\{\pm 1c, \pm 3c, ... \pm(\sqrt{M}-1)c\}$, and Y_k, $k = 1,2..8$ are functions of ϕ_1 and ϕ_2 which are independent and uniformly distributed over $[-\pi, \pi]$. In the processes of evaluating (159), we note that the n th moment of equally likely M-ary symbol [17,18] is given by

$$E_{I_{ij}}\{I_{ij}^n\} = \frac{1}{\sqrt{M}} \sum_{m=0}^{\sqrt{M}-1} (2m + 1 - \sqrt{M})^n c^n, \tag{160}$$

and for the case of independent and zero mean M-ary symbols I_{ij}, we have

$$\begin{aligned} E_{\mathbf{I_{ij},I_{kl}}}\{I_{ij} I_{kl}\} &= 0 \quad i \neq k, \ or \ \ j \neq l \\ E_{\mathbf{I_{ij}}}\{I_{ij}\} &= 0 \quad i = 1,2 \ \ j = I,R \end{aligned} \tag{161}$$

Furthermore for the n th moment of Y_i $i = 1,..8$ which are function of ϕ_1 and ϕ_2, we use

$$\begin{aligned} E_{\phi_1,\phi_2}\{Y_i(\phi_1,\phi_2)^n\} &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} Y_i(\phi_1,\phi_2)^n f_{\phi_1,\phi_2}(\phi_1,\phi_2) \ d\phi_1, d\phi_2 \\ &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} Y_i(\phi_1,\phi_2)^n f_{\phi_1}(\phi_1) f_{\phi_2}(\phi_2) \ d\phi_1 d\phi_2 \quad i = 1.2..8 \end{aligned} \tag{162}$$

With Amplitude and Phase compensation

Similar analysis is used to find the average probability of error for the case when the decision variable is obtained with both amplitude and phase compensation.

In this case, using (125), we first calculate the conditional probability, $P(|Z_{1R}| > c|\phi_1, \phi_2, I_{2R}, I_{2I})$. That is, we integrate on the joint probability of the random variable Y_{AP},

Define

$$Y_{AP} = n_{1R}Y_{3AP} + n_{1I}Y_{4AP} + n_{2R}Y_{5AP} + n_{2I}Y_{6AP} \tag{163}$$

$$X_{IAP} = I_{2R}Y_{1AP} + I_{2I}Y_{2AP}, \tag{164}$$

then from (125),

$$Z_{1R} = X_{IAP} + Y_{AP} \tag{165}$$

The random variable Y_{AP} is zero mean Gaussian and has variance

$$\sigma^2_{oAP}(\phi_1, \phi_2) = (Y^2_{3AP} + Y^2_{4AP} + Y^2_{5AP} + Y^2_{6AP})\sigma^2_n \tag{166}$$

Also conditioned on ϕ_1,ϕ_2, I_{2R} and I_{2I} , the random variable Z_{1R} is Gaussian with mean equals X_{IAP} and variance $\sigma^2_{oAP}(\phi_1, \phi_2)$. Similar to (153),

$$P_{1AP}(e|\phi_1, \phi_2, I_{2R}, I_{2I}) = 2(1 - \frac{1}{\sqrt{M}})Q\Big(\frac{c - X_{IAP}}{\sigma_{AP}}\Big). \tag{167}$$

or

$$P_{1AP}(e|\mathbf{x_{AP}}) = 2(1 - \frac{1}{\sqrt{M}})Q(\frac{c}{\sigma_n}\mathbf{x_{AP}}), \tag{168}$$

where the random variable $\mathbf{x_{AP}}$,

$$\mathbf{x_{AP}} = \frac{c - X_{IAP}}{\sigma_{oAP}}, \tag{169}$$

with

$$\sigma^2_{oAP} = (Y^2_{3AP} + Y^2_{4AP} + Y^2_{5AP} + Y^2_{6AP})c^2, \tag{170}$$

is a function of the random variables $(\phi_1, \phi_2, I_{2R}, I_{2I})$. Because of independence assumption,

$$f_{\mathbf{x_{AP}}}(x) = f_{\mathbf{I_{2R}}}(I_{2R}) f_{\mathbf{I_{2I}}}(I_{2I}) f_{\phi_1}(\phi_1) f_{\phi_2}(\phi_2) \tag{171}$$

Similar to (157), we can use GQR to calculate the probability of error from the moments of the random variable $\mathbf{x_{AP}}$; viz,

$$P_{1AP}(e) = 2(1 - \frac{1}{\sqrt{M}}) \sum_{i}^{N} w_i Q\left(\sqrt{\frac{3(SNR)}{M-1}} x_i\right) \tag{172}$$

where again x_i and w_i are the nodes and the weights of the GQR.

The moments of $\mathbf{x_{AP}}$ can be calculated using the simple binomial rule

$$E\{\mathbf{x_{AP}^n}\} = E\left\{ \frac{\sum_{k=0}^{n} \binom{n}{k} c^k (-1)^{n-k} (I_{2R} Y_{1AP} + I_{2I} Y_{2AP})^{n-k}}{\sigma_{oAP}^n} \right\} \tag{173}$$

We take the expected values of the inner terms in (173), then,

$$E\{\mathbf{x_{AP}^n}\} = \sum_{k=0}^{n} \binom{n}{k} c^k (-1)^{n-k} E\{A_{xAP}\} \tag{174}$$

where

$$A_{xAP} = \sum_{l=0}^{n-k} \binom{n-k}{l} I_{2R}^{l} \frac{Y_{1AP}^{l}}{\sigma_{oAP}^{n}} I_{2I}^{(n-k-l)} \frac{Y_{2AP}^{(n-k-l)}}{\sigma_{oAP}^{n}} \tag{175}$$

5.3 Results

The Chernoff upper bound on the average probability of error as a function of signal-to-noise (SNR) ratio is evaluated for various cross coupling constants and for 16 QAM and 64 QAM signals. The Gauss quadrature rule is also used to find approximations to the probability of errors.

In Fig. 12, the bounds on error probability with 16 QAM and with cross polarization coupling r= -15 dB, -10 dB and -5 dB are calculated and compared. Equation

(145) is used in calculating these bounds when only amplitude compensation is employed, while (147) is used when both amplitude and phase compensation is employed. Notice that adding phase compensation improves the bound when the cross coupling is high (i.e., r=-5 dB). The effect of adding phase compensation is hardly noticeable with low cross coupling (r=-15 dB). Fig. 13 depicts the same for 64 QAM. The effect of compensation is similar. Nevertheless, as expected the bounds are higher for 64 QAM displaying possibility of higher error rates with the same SNR. Comparison of these bounds for 16 QAM and 64 QAM are shown in Fig. 14.

In Fig. 15 and Fig. 16 we depict the probability of error as it is calculated using the Gauss quadrature rule. for 16 QAM and 64 QAM, respectively. These calculations were done with cross coupling of -15 dB. -10 dB and -5 dB, and in each a total of 9 moments were used. Only the case with amplitude compensation was shown since adding phase compensation did not change these results very much. In order to show how tight are the Chernoff bounds shown in in Fig. 12, we depict in Fig. 17 a comparison of the results obtained with GQR moments calculations to their corresponding Chernoff bounds for 16 QAM and cross coupling of -15 dB, -10 dB and -5 dB. Fig. 18 shows the same for the 64 QAM case. To show the effect of increasing the number of moments used in obtaining the GQR results, we show in the next two figures these results with 7, 9 and 11 moments. Chernoff bound was added to these curves for comparison. In Fig. 19, we present these comparisons for the 16 QAM case, while Fig. 20 presents the same for the 64 QAM case.

It is important to emphasize that the error probability, although an important factor, is certainly not the only advantage of the bootstrapped canceler. We mention the following other points in favour of these cancelers:

1. Under the same system condition, the bootstrapped canceler steady state interference residue is smaller.

2. To implement the bootstrapped algorithm one needs less complex hardware than for the diagonalizer, which needs a zero forcing algorithm, and for the LMS canceler, which needs decision feedback information. In fact, it is clear that adding a decision feedback to the bootstrap schemes will result in faster convergence and still better performance than that which we obtained in the current analysis.

3. The fact that the bootstrapped cancelers do not need decision feedback makes them ideal for acquisition and, hence, suitable for channels with fast and deep fading which causes occasional system outage.

In Fig. 21, we compare the performance of the power-power canceler to that of the LMS canceler, for 16 QAM and $r = -10$ dB. Fig. 22 depicts the same comparison with $r = -15$ dB. To emphasize the need for cancelers in dual polarized systems, we add, to these two curves, the error performance without cancelers.

Figs. 23 and 24 are the same as Figs. 21 and, 22 except for the use of 64 QAM instead of 16 QAM. In the last four figures, the moment method was used. For each curve, the number of the moments is marked in parentheses.

The three cancelers, power-power canceler, LMS canceler and the diagonalizer, are compared in Figs. 25 and 26. A 16 QAM signal is assumed in these figures, with $r = -10$ dB and $r = -15$ dB, respectively. Although GQR calculation has been done for amplitude and the phase compensated diagonalizer (see chapter 2), the GQR calculation has not been done for the amplitude compensated diagonalizer. Therefore, the comparisons are based on the Chernoff bound.

5.4 Conclusion

The power-power bootstrapped canceler was analyzed and its performance was studied in this chapter. In particular the average probability of error was estimated using the moment generating method or by finding the Chernoff bounds. Results of

the analysis as well as computer calculations show, as expected, that 16 QAM performance is much better than 64 QAM. It is also shown that adding phase compensation to the canceler output adds very little to the performance, when only amplitude compensation is included.

From comparing the results obtained with the moment generating method to the corresponding Chernoff bound, we concluded that these bounds are sufficiently tight. Comparing the results when different numbers of moments are used, and the concluded tightness of the Chernoff bound, we infer that approximately 10 moments are sufficient for deriving a good approximation for the average probability of error using the Gauss quadrature rule.

References

[1] Kavehrad, M. "Performance of Cross-Polarized M-ary QAM Signals Over Nondispersive Fading Channels", *AT&T Bell Lab. Tech. J.*, Vol. 63, pp. 499–521, March 1984.

[2] Duvoisin P., et. al.," Performance of the Adaptive Baseband Diagonalizer, A Crosstalk Canceller, in Dual-Polarized 16-QAM System" *Record, IEEE Southeastern '88*, Knoxville, TN, April 10-13, 1988.

[3] Bar-Ness Y., and Rokach, J.," Cross-Coupled Bootstrapped Interference Canceller", *Int. Conf. on Ant. and Pro.*, Vol. I. pp. 292-295, June 1981, Los Angles, CA.

[4] Bar-Ness Y., et al.," Bootstrapping Adaptive Cross-Pol Canceller for Satellite Communications", *Int. Conf. on Comm.*, June 1982, Philadelphia, PA.

[5] Carlin, J., et. al.," An IF Cross-Pol Canceler for Microwave Radio," *IEEE J. Select. Areas Commun.*, Vol. SAC-5, No. 3 pp. 502–514, April 1987.

[6] Bar-Ness Y.," Bootstrapped Algorithm for Interference Cancellation," *AFCEA-IEEE Technical Conference Tactical Comm.*, Fort Wayne, Indiana, May 1988.

[7] Chu, T. S.,"Restoring the orthogonality of two polarizations in radio communication systems," *Bell Sys. Tech. J.*, Vol. 50, pp. 3063-3069. Nov. 1971.

[8] Lin, S. H.,"Impact of Microwave Depolarization During Multipath fading on Digital Radio Performance," *Bell Sys. Tech. J.*, Vol. 56, No. 5, pp. 645-674, May 1977.

[9] Steinberger, M. L., "Design of a Terrestial Cross-Pol Canceller," *Int. Conf. on Commun.*, Philedelphia, PA, June 1982.

[10] Amitay N.,"Signal-to-Noise Ratio Statistics for Nondispersive Fading in Radio Channels with Cross Polarization Interference Cancellation ," *IEEE Trans. Commun.*, Vol. 27, pp. 498–502, Feb. 1979.

[11] Kavehrad, M. et. al.,"On the Performance of Combined Quadrature Amplitude Modulation and Convolutional Codes for Cross-Coupled Multidimensional Channels," *IEEE Trans. Commun.*, Vol. 34, pp. 1190–1201, Dec. 1986.

[12] Nichols H. E., et. al.,"MLD and mse Algorithms for Adaptive Detection of Digital Signals in the Presence of Interchannel Interference," *IEEE Trans. Inform. Theory*, Vol. IT-23. No. 5, pp. 563–574, Sept. 1977.

[13] Widrow, B. et. al. ," The Complex LMS Algorithm ," *IEEE Proceedings* , Vol. 63 , No. 12, pp. 1692–1716, Dec. 1975.

[14] Dinc, A. ," The Theory of Bootstrapped Algorithms and Their Applications to Cross Polarization Interference Cancelation ," *Ph.D Dissertation* , N.J. Ins. of Technology, May 1991.

[15] Gautschi, W.," On the Construction of Gaussian Quadrature Rules from Modified Moments," *Math Comput.*, Vol. 24, pp. 245–260, Apr. 1970.

[16] Golub, G. H. and Welsch, J. H.," Calculation of Gauss Quadrature Rules," *Math Comput.*, Vol. 26, pp. 221–230, Apr. 1969.

[17] Benedotto. S. et. al.,"Error Probability in the Presence of Intersymbol Interference and Additive Noise for Multilevel Digital Signals," *IEEE Trans. Commun.*, Vol. 21. pp. 181–190, March 1973.

[18] Korn, I.."*Digital Communications*," Van Nostrand Reinhold Company 1985.

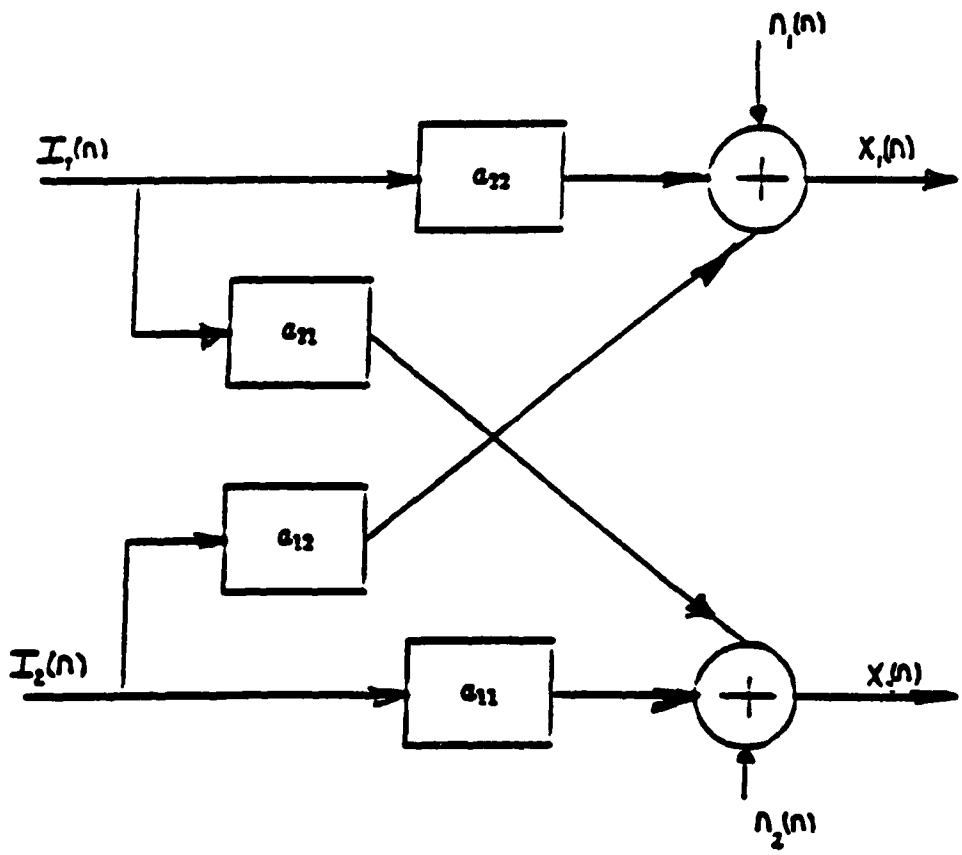


Fig. 1 Dually Polarized Channel Model

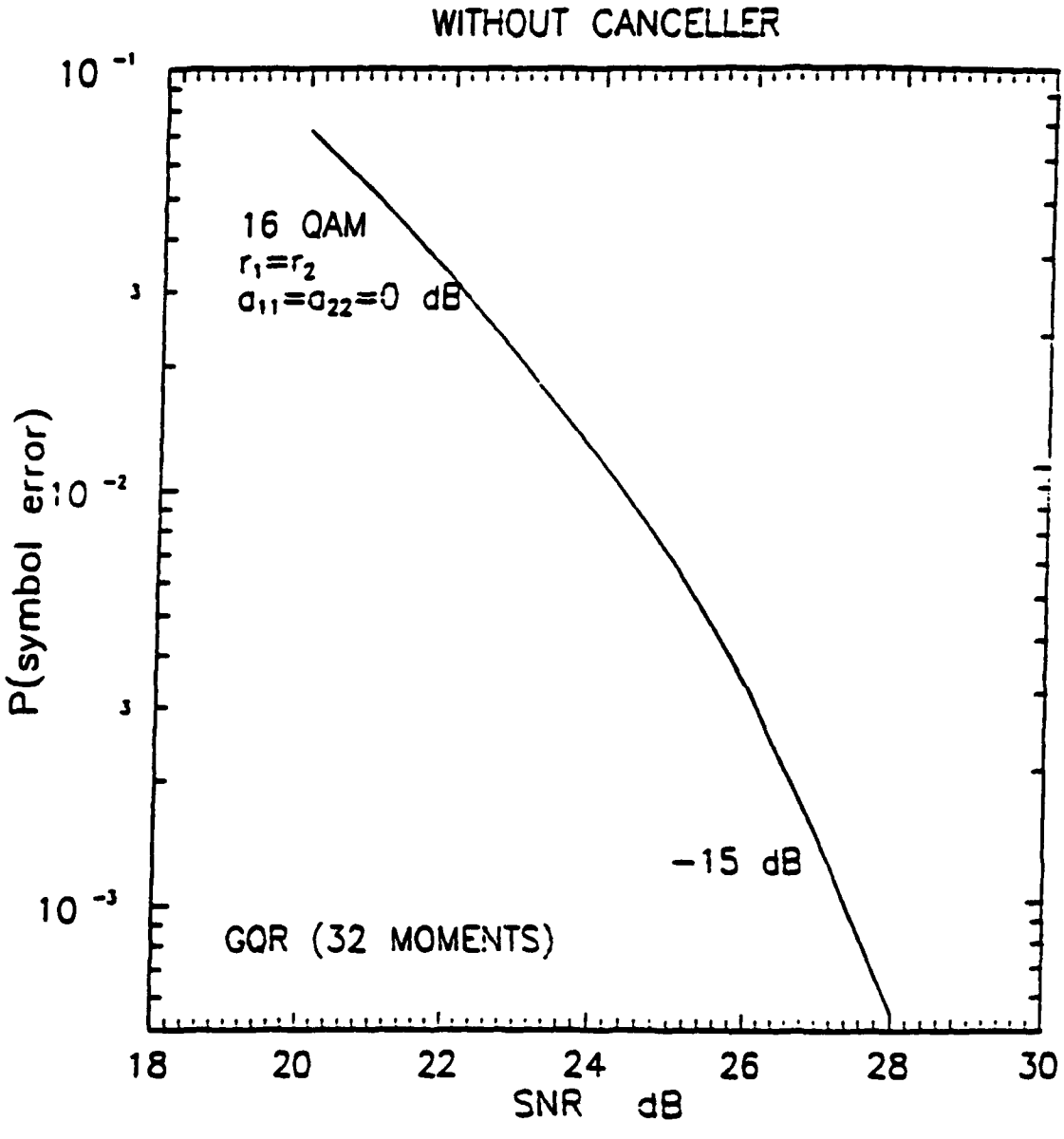


Fig. 2 Performance of Dually Polarized 16 QAM System, without Cross-Pol Interference Canceler

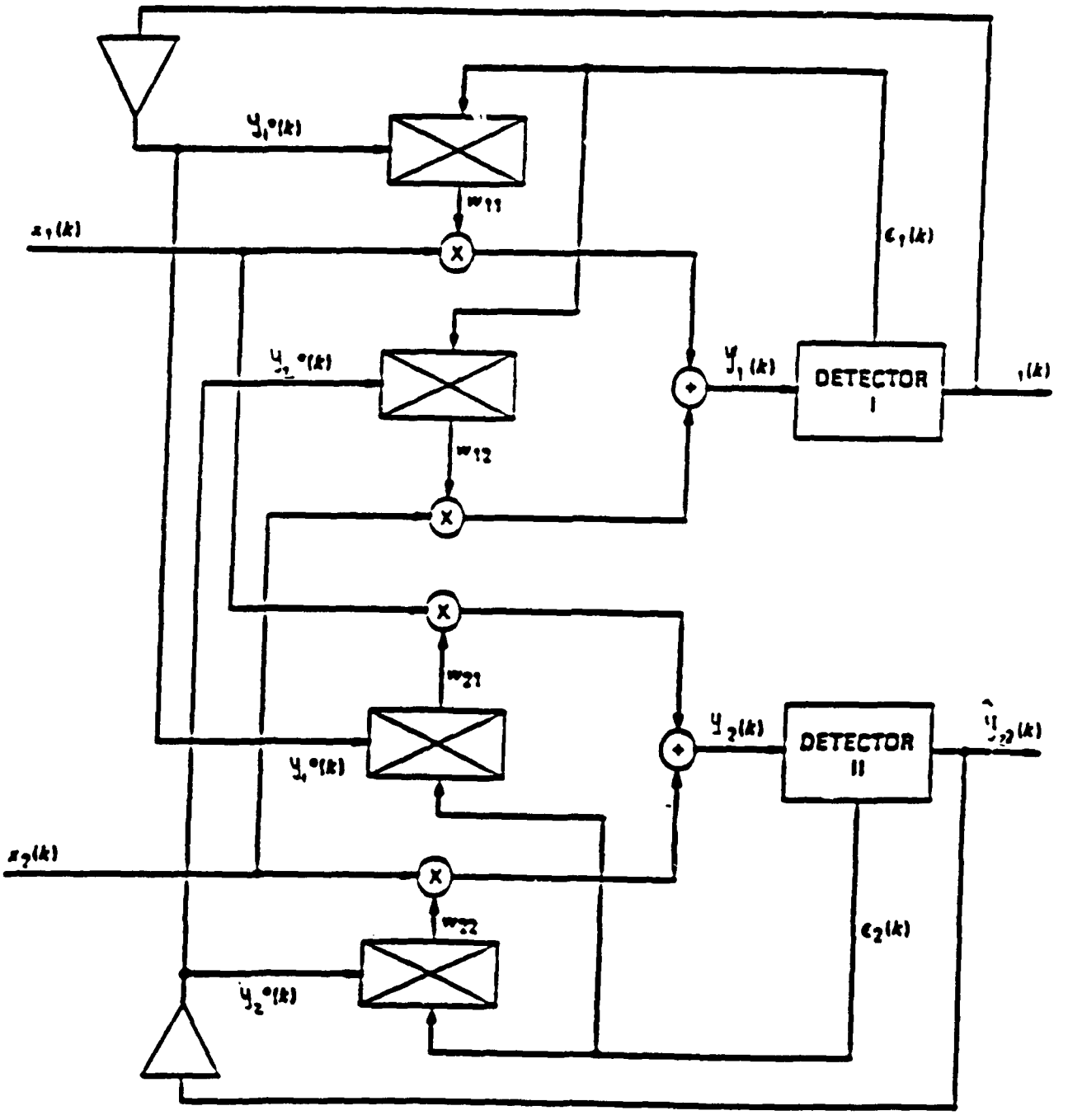


Fig. 3 Diagonalizer Cross-Pol Interference Canceler

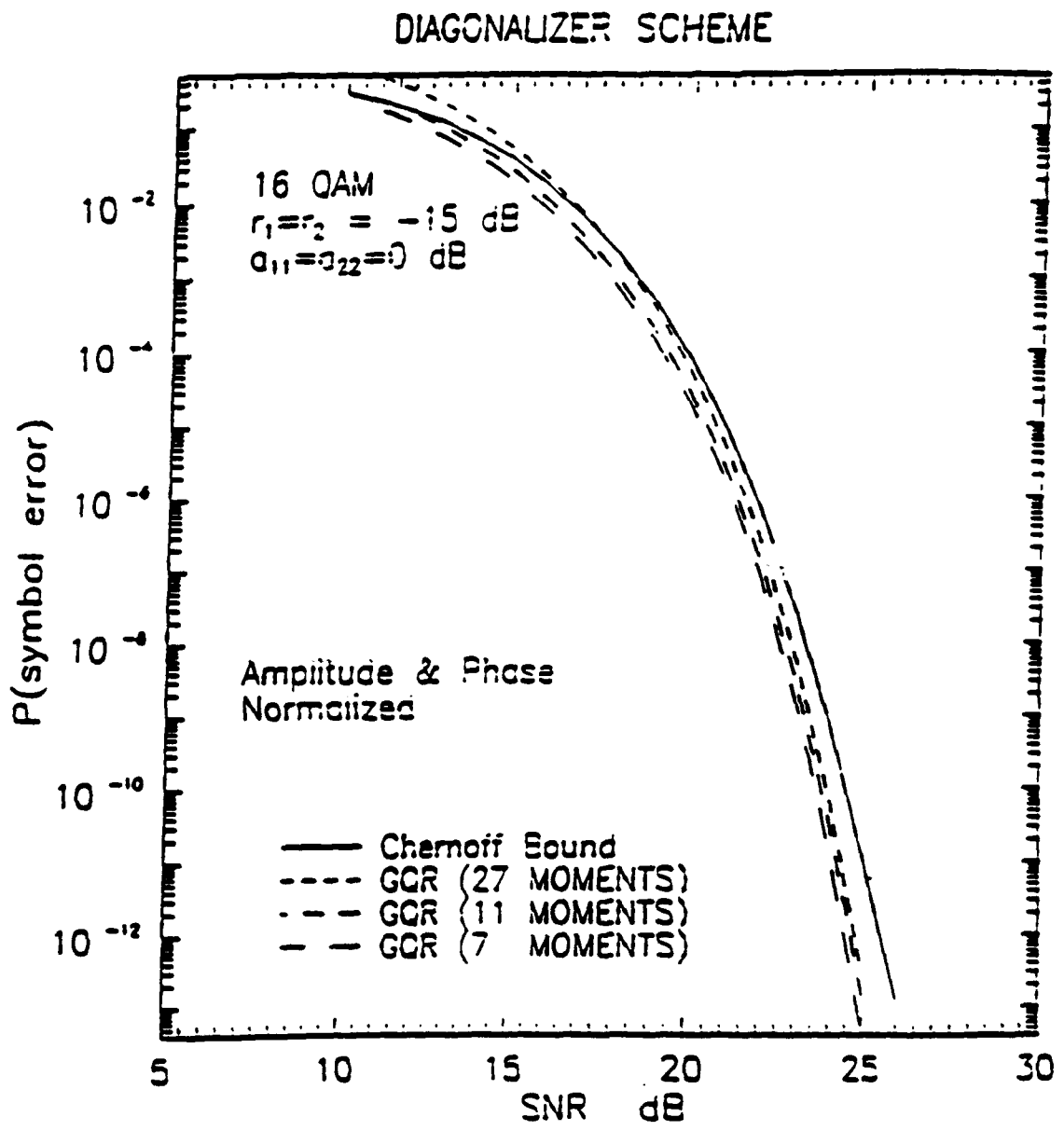


Fig. 4 Diagonalizer Cross-Pol Interference Canceler. Chernoff Bound and GQR calculation with amplitude and phase compensation. cross coupling -15 dB.

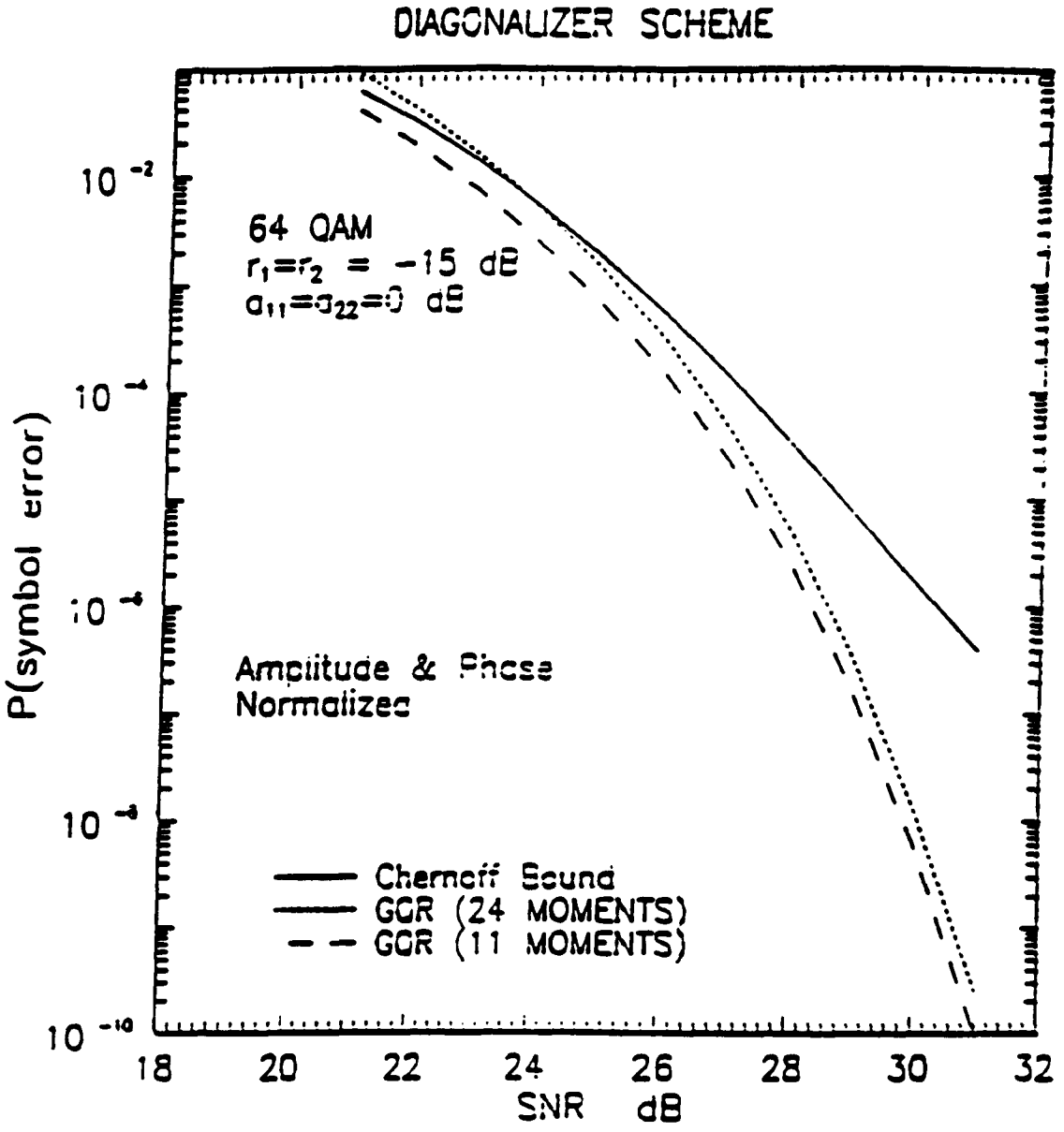


Fig. 5 Diagonalizer Cross-Pol Interference Canceler. Chernoff Bound and GQR calculation. 64 QAM with amplitude and phase compensation, cross coupling -15 dB.

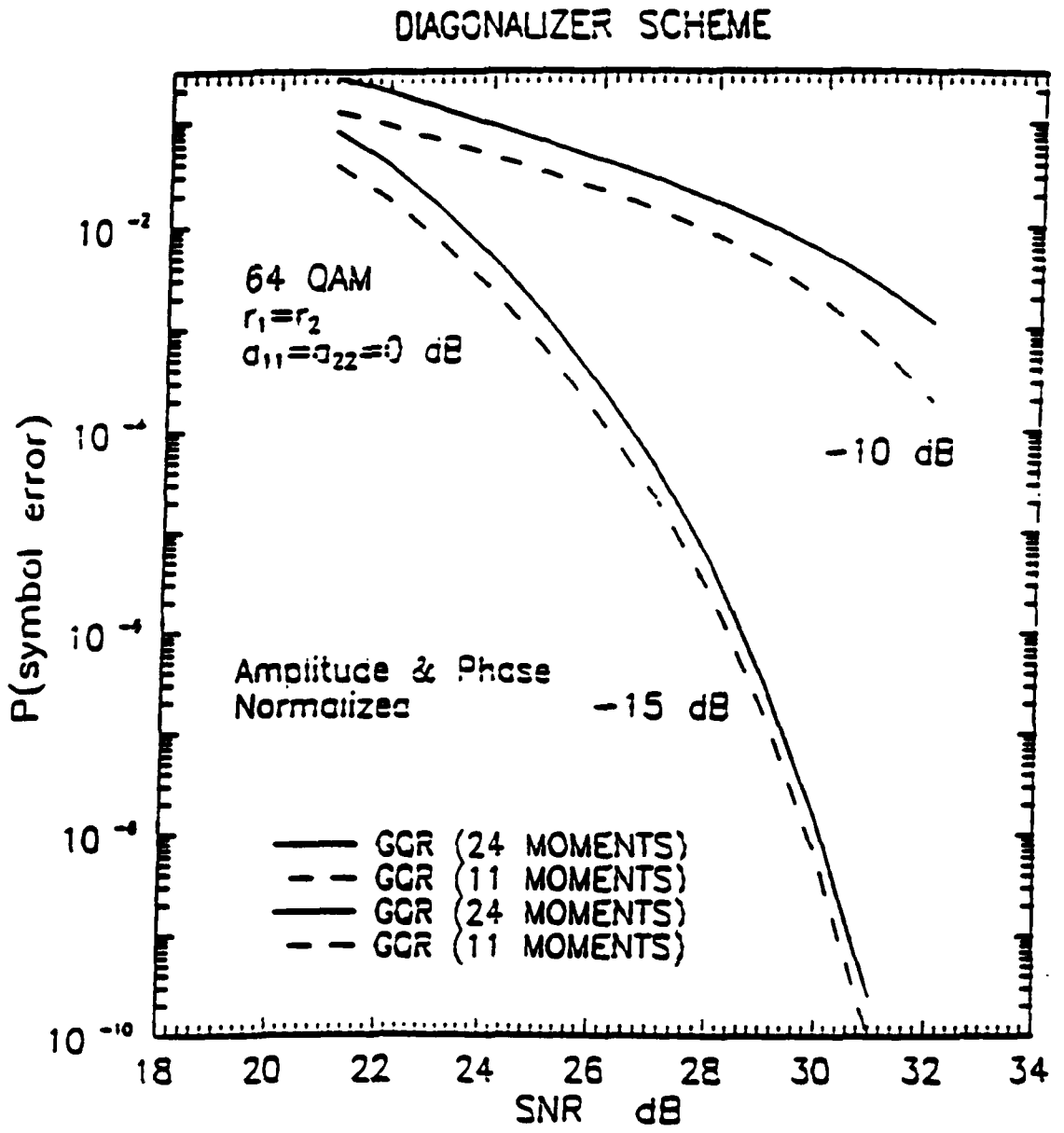


Fig. 6 Diagonalizer Cross-Pol Interference Canceler. GQR calculation, 16 vs. 64 QAM with amplitude and phase compensation, cross coupling -15 dB, -10 dB.

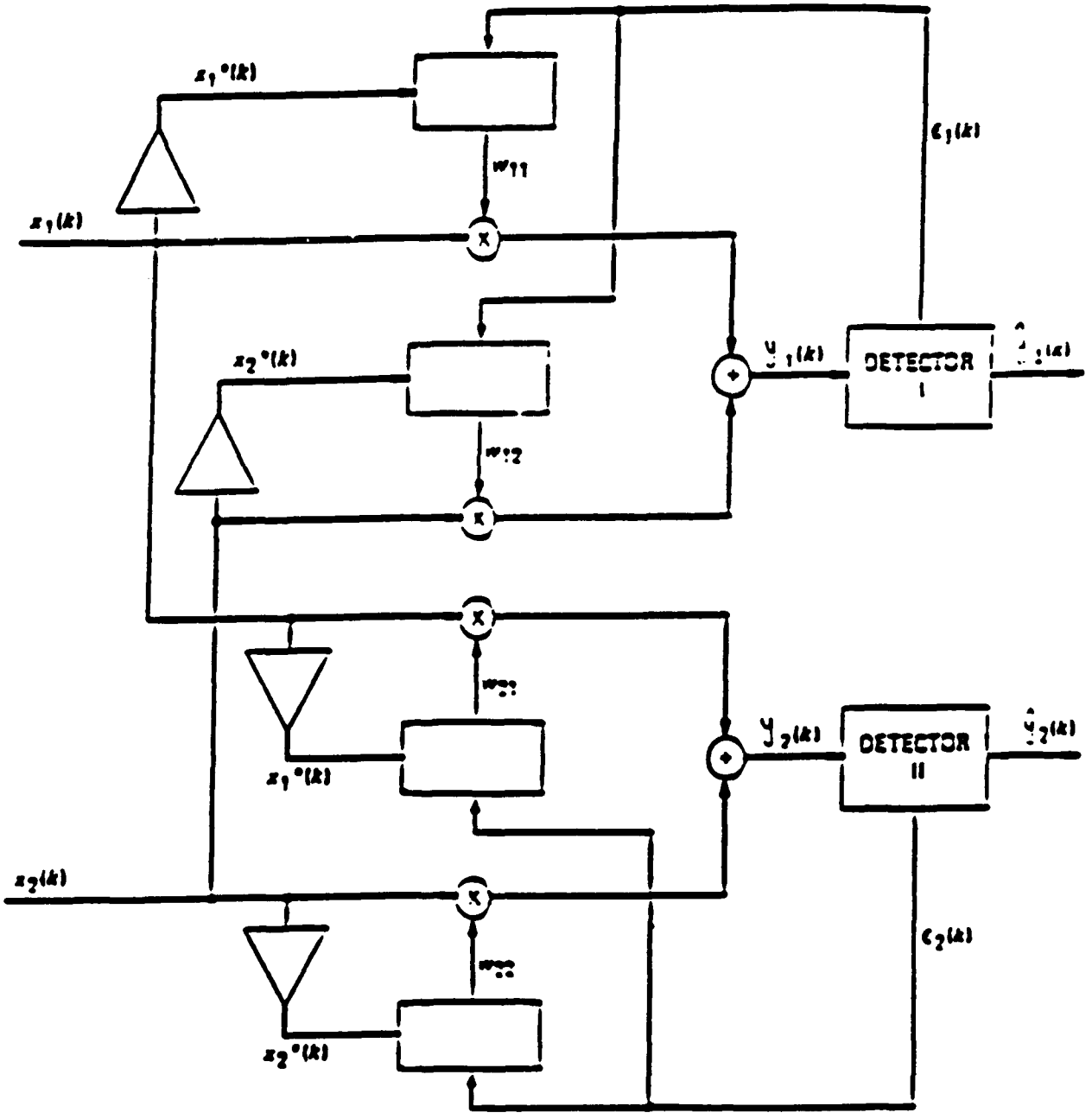


Fig. 7 LMS Cross-Pol Interference Canceler

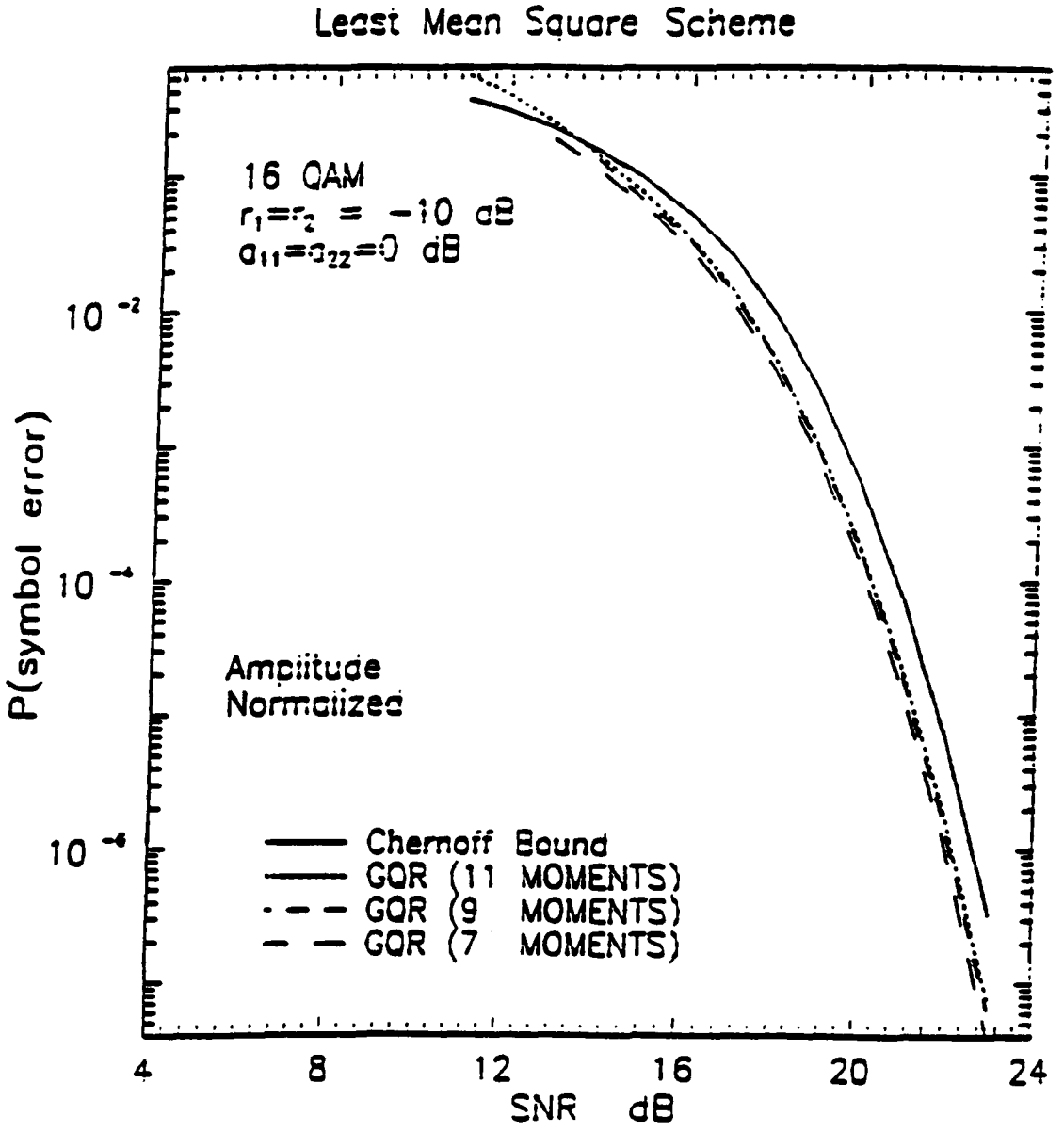


Fig. 8 LMS Cross-Pol Interference Canceler. Chernoff Bound and GQR calculation. 16 QAM with amplitude and phase compensation. cross coupling -10 dB.

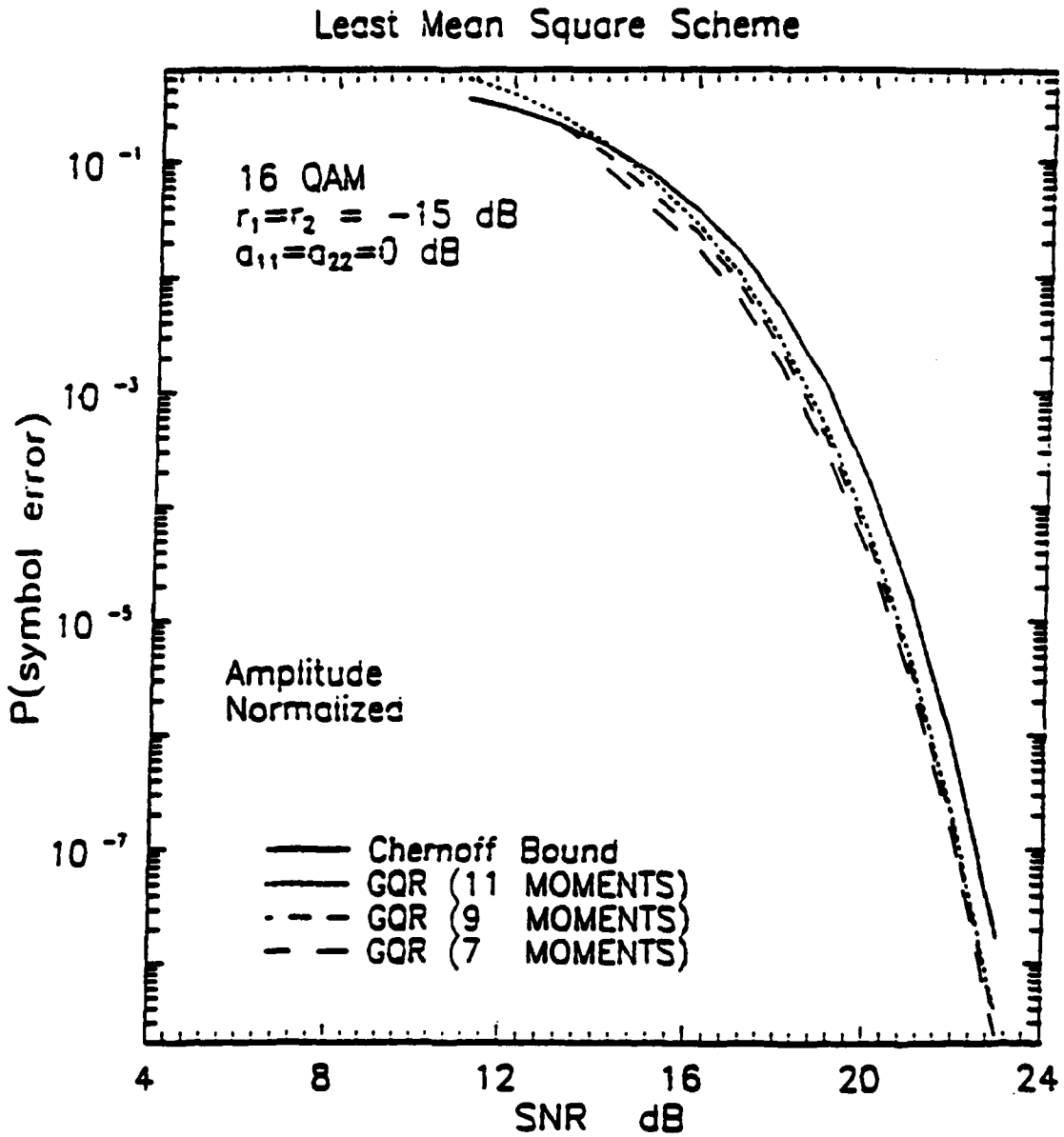


Fig. 9 LMS Cross-Pol Interference Canceler, Chernoff Bound and GQR calculation. 16 QAM with amplitude and phase compensation, cross coupling -15 dB.

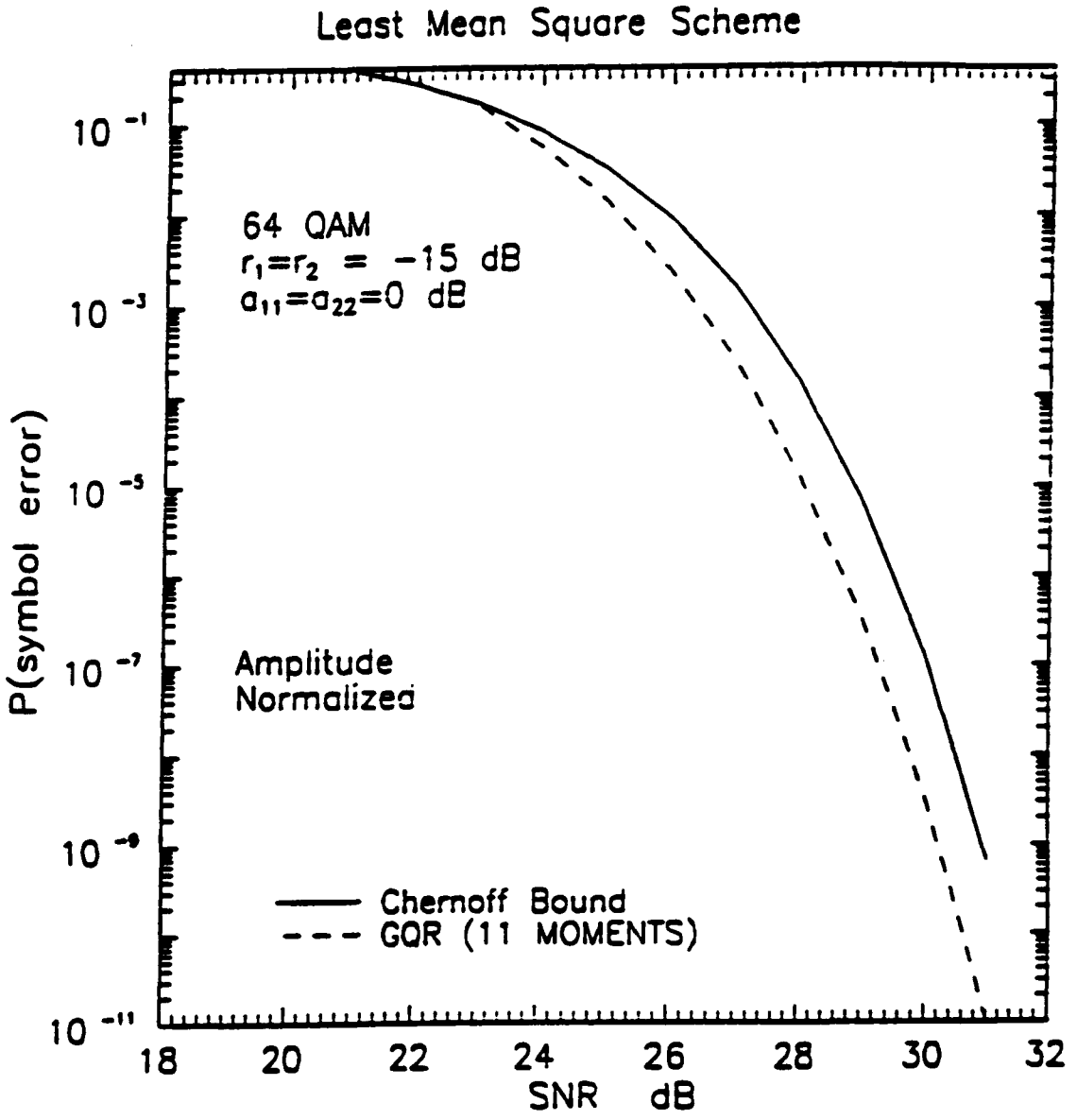


Fig. 10 LMS Cross-Pol Interference Canceler, Chernoff Bound and GQR calculation, 64 QAM with amplitude and phase compensation, cross coupling -15 dB.

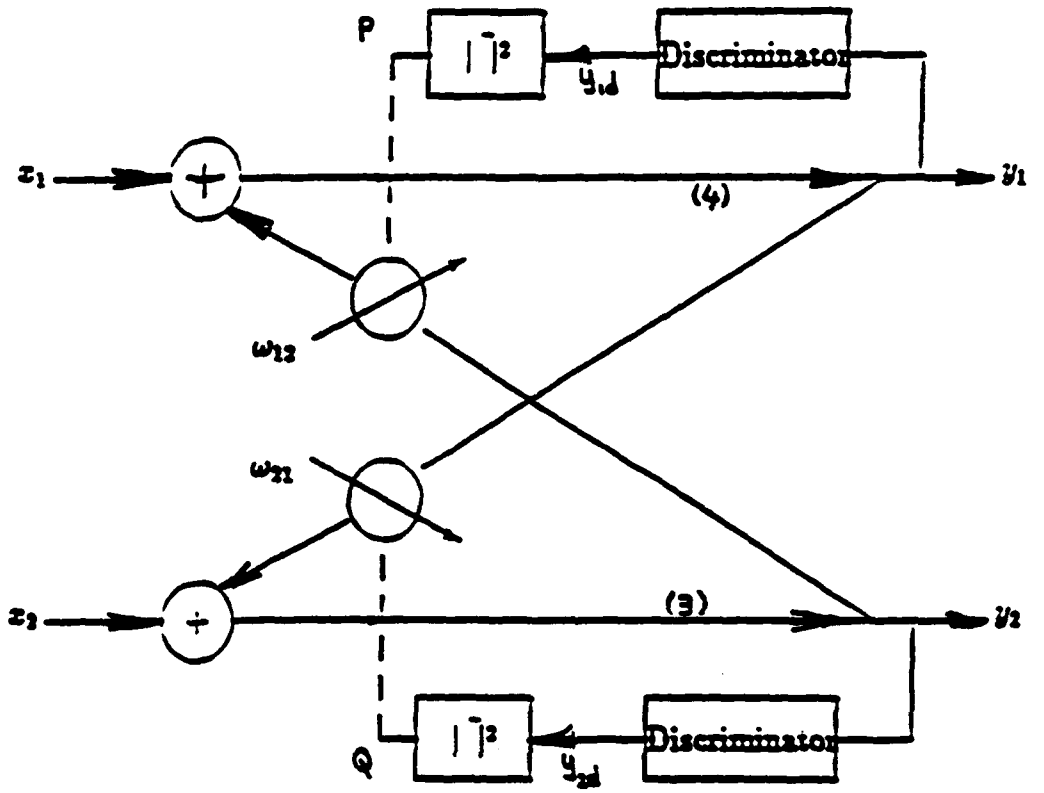


Fig. 11 Power-Power Cross-Pol Interference Canceler

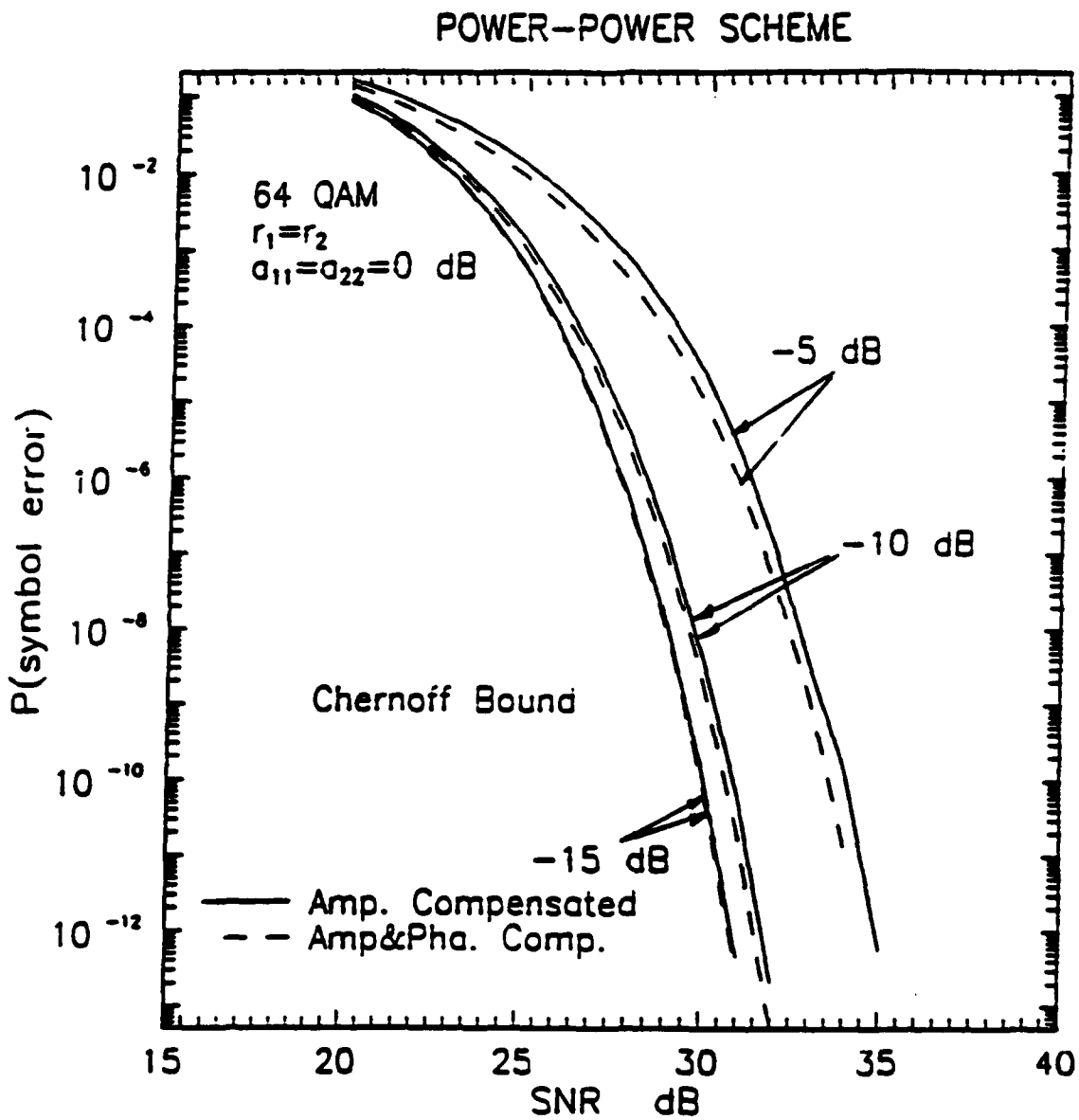


Fig. 12 Power-Power Cross-Pol Interference Canceler, Chernoff Bound 16 QAM

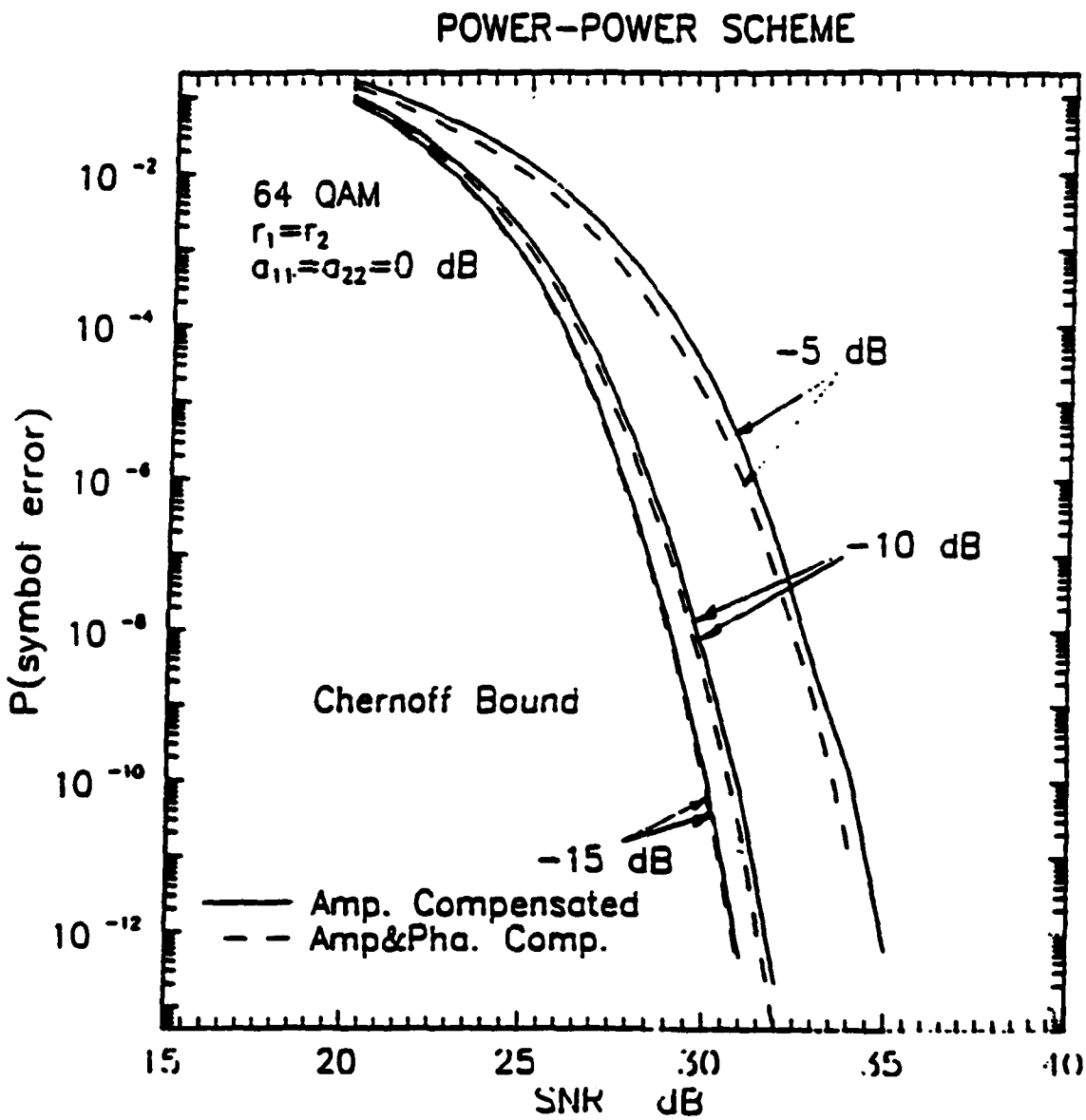


Fig. 13 Power-Power Cross-Pol Interference Canceler, Chernoff Bound 64 QAM

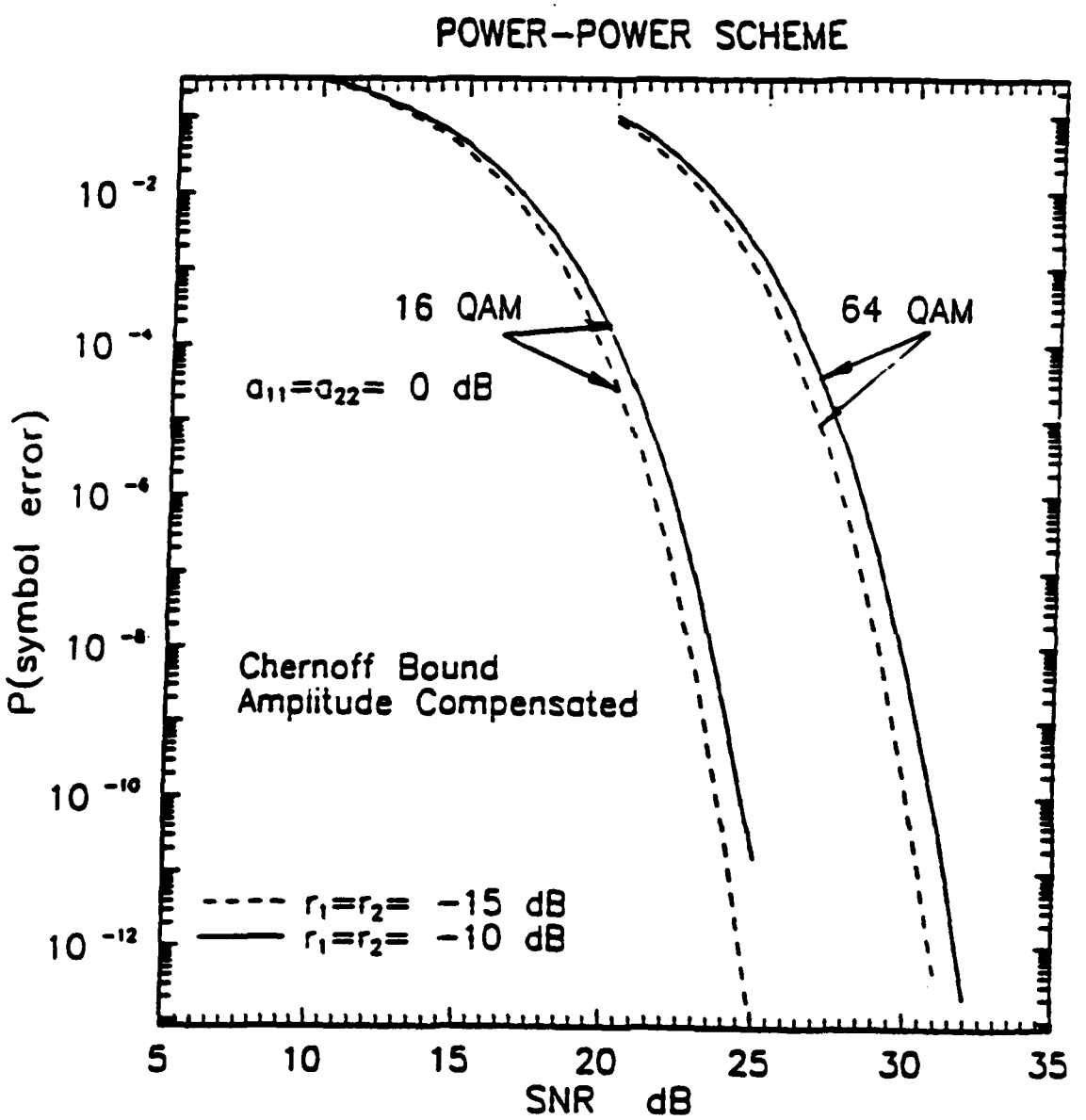


Fig. 14 Power-Power Cross-Pol Interference Canceler, Chernoff Bound comparison
16 QAM vs. 64 QAM

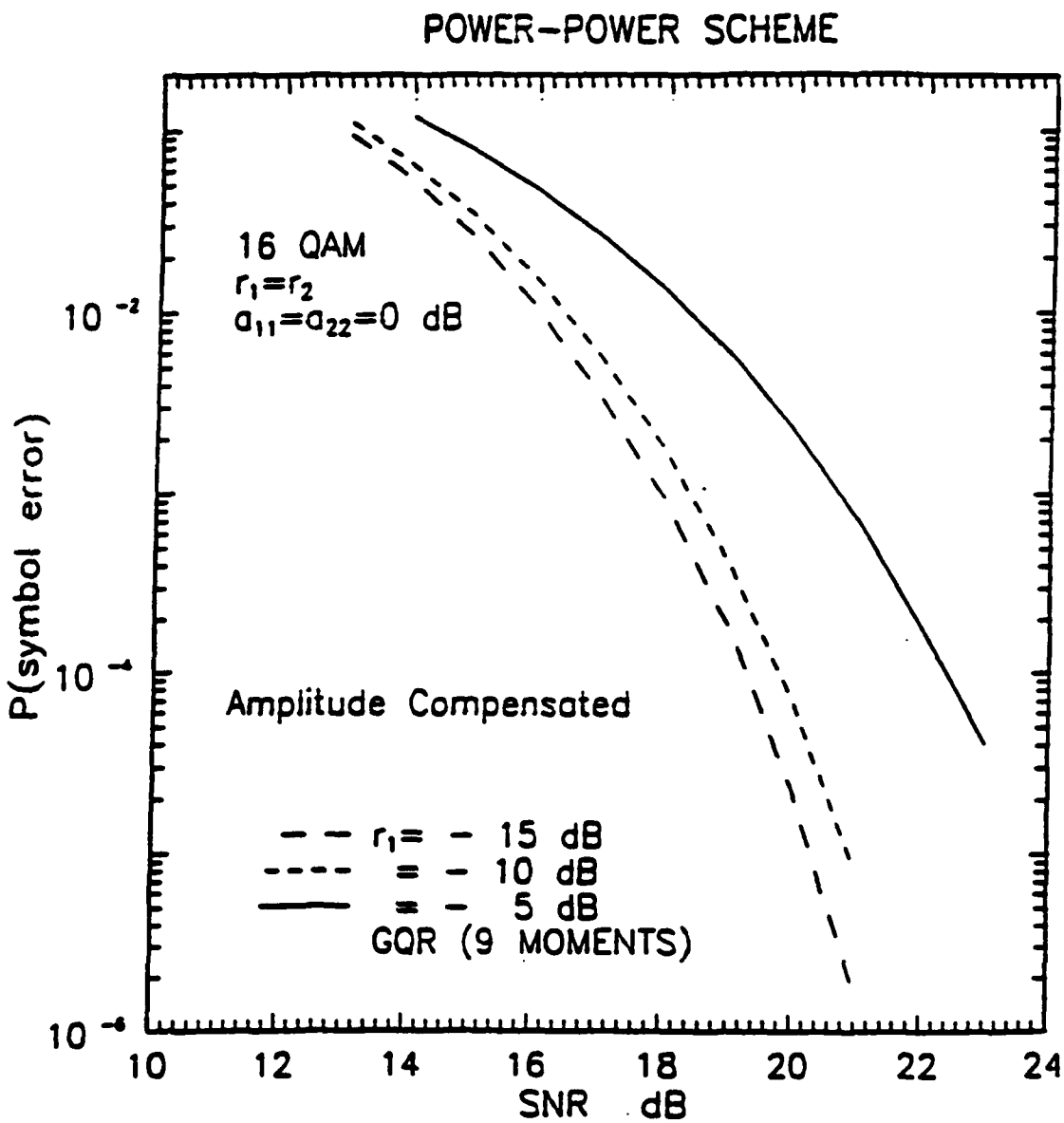


Fig. 15 Power-Power Cross-Pol Interference Canceler, GQR calculation, 16 QAM

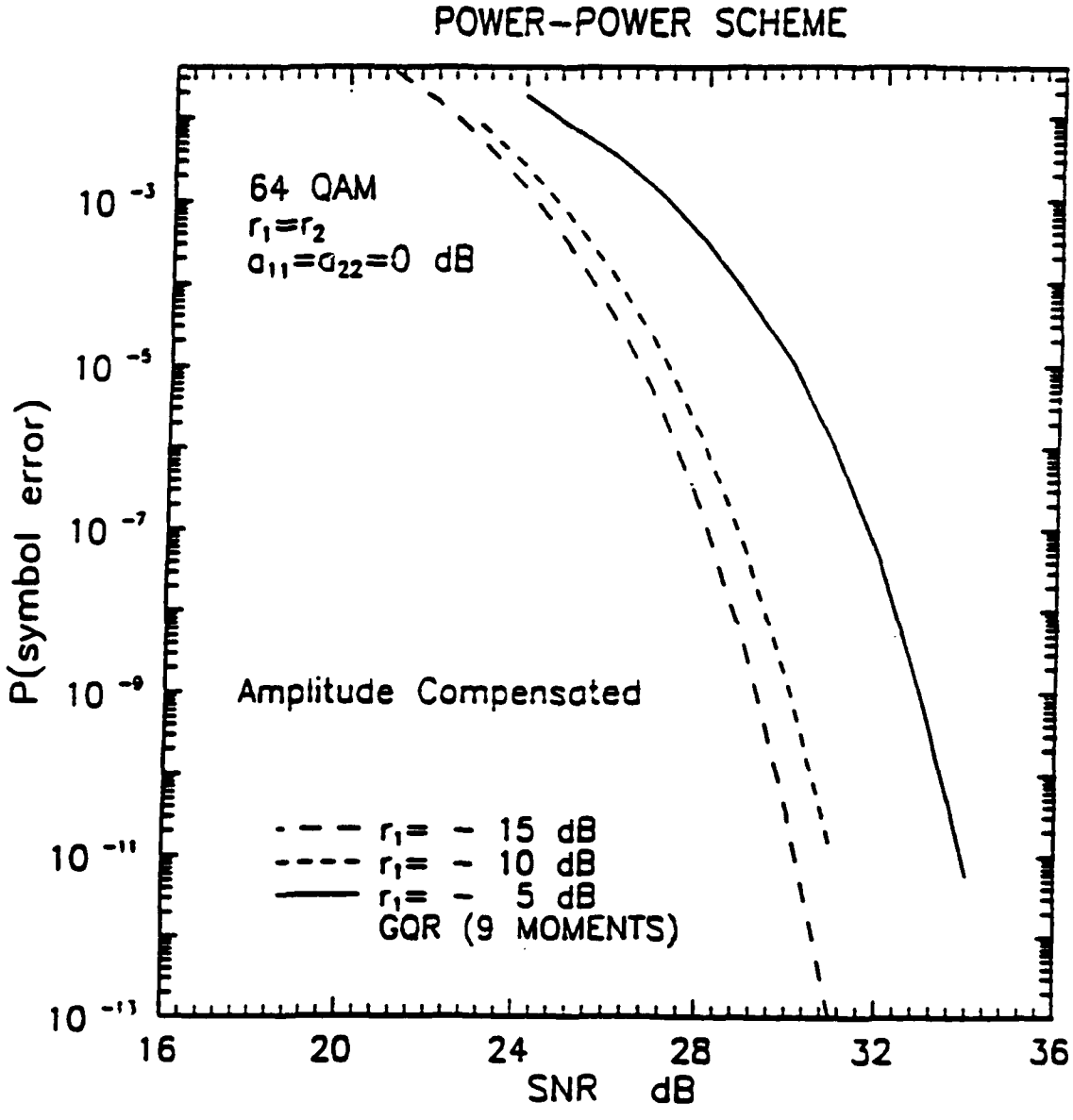


Fig. 16 Power-Power Cross-Pol Interference Canceler, GQR calculation. 64 QAM

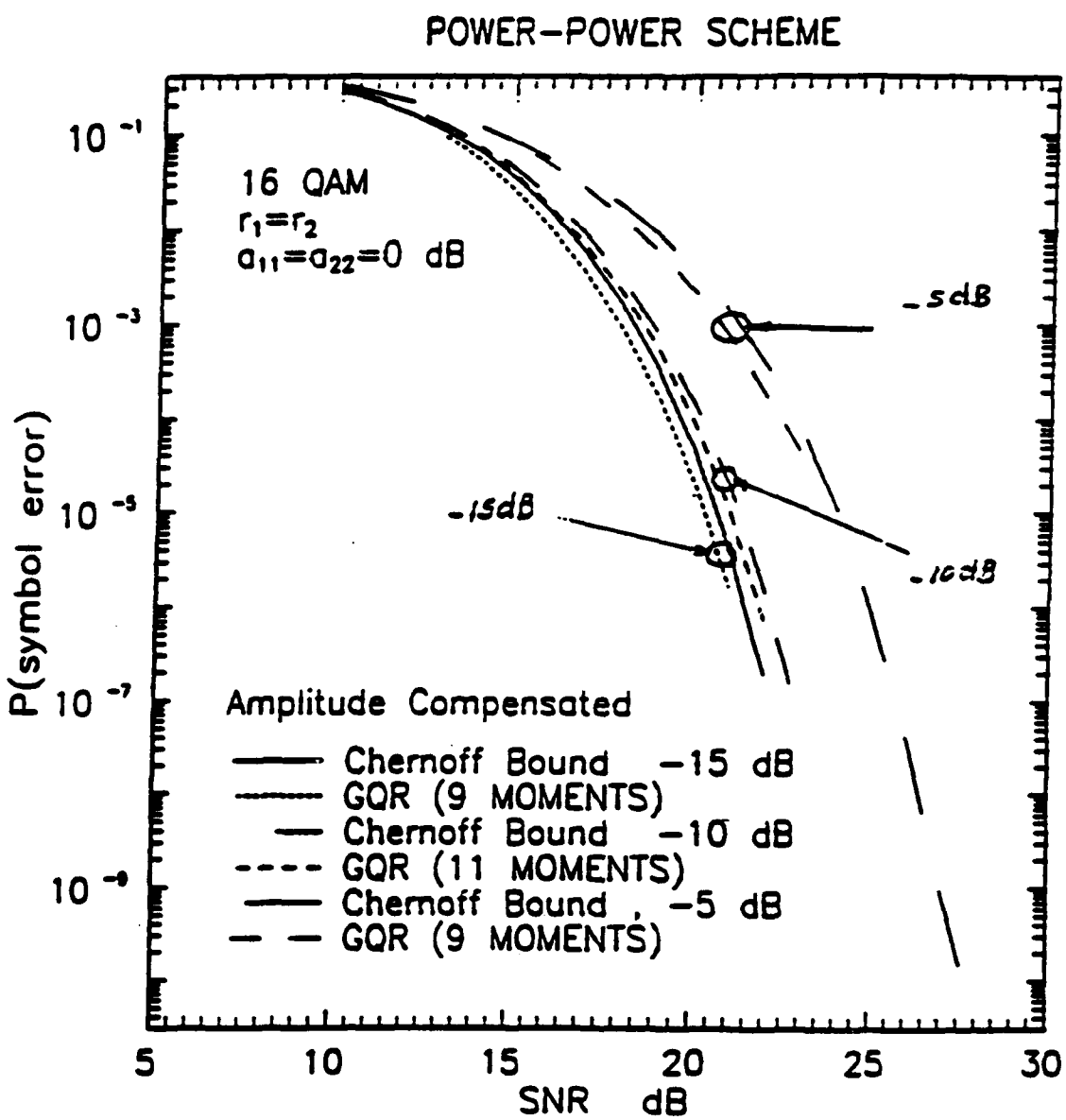


Fig. 17 Power-Power Cross-Pol Interference Canceler, Chernoff Bound and GQR calculation comparison, 16 QAM

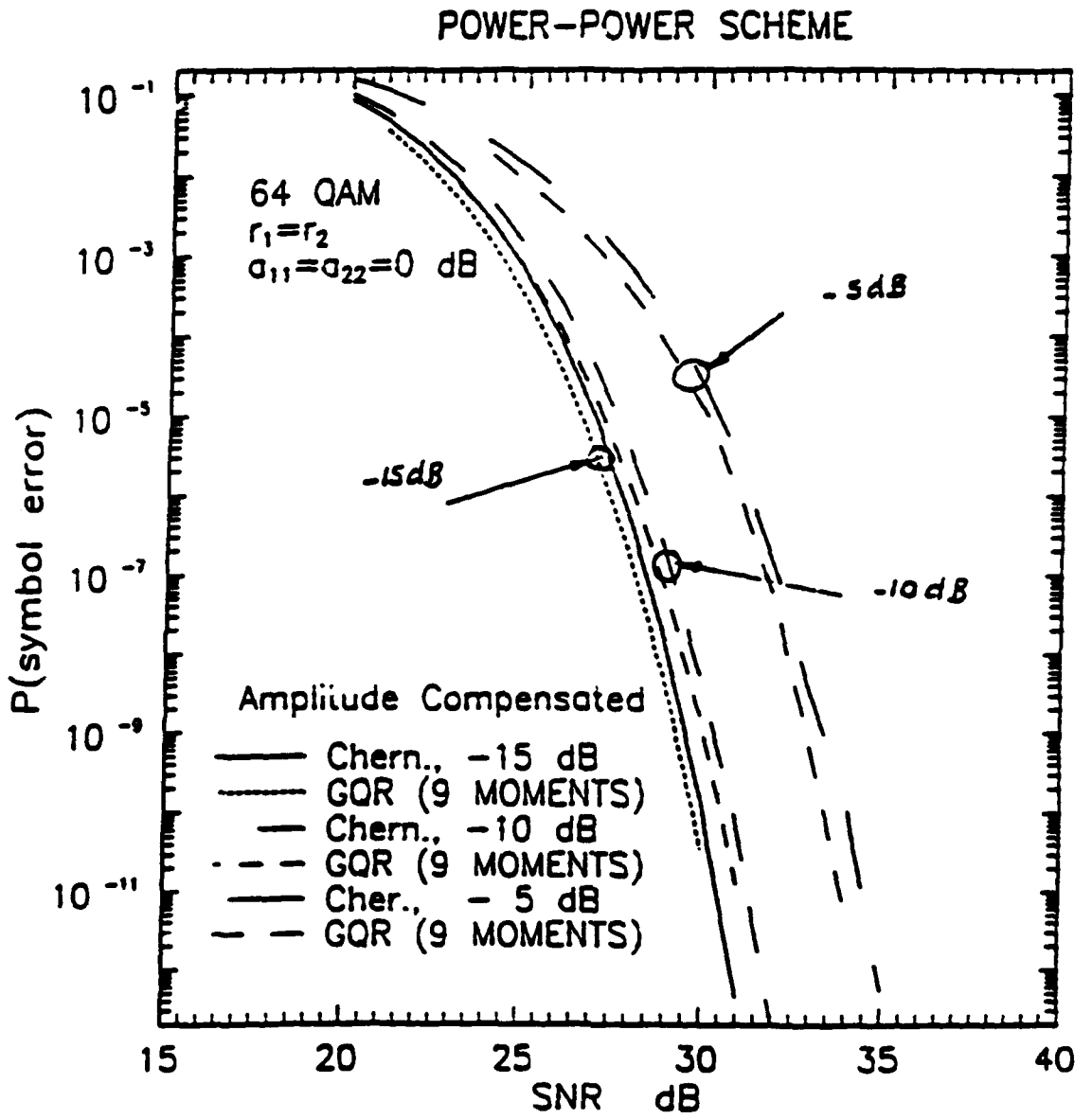


Fig. 18 Power-Power Cross-Pol Interference Canceler, Chernoff Bound and GQR calculation comparison, 64 QAM

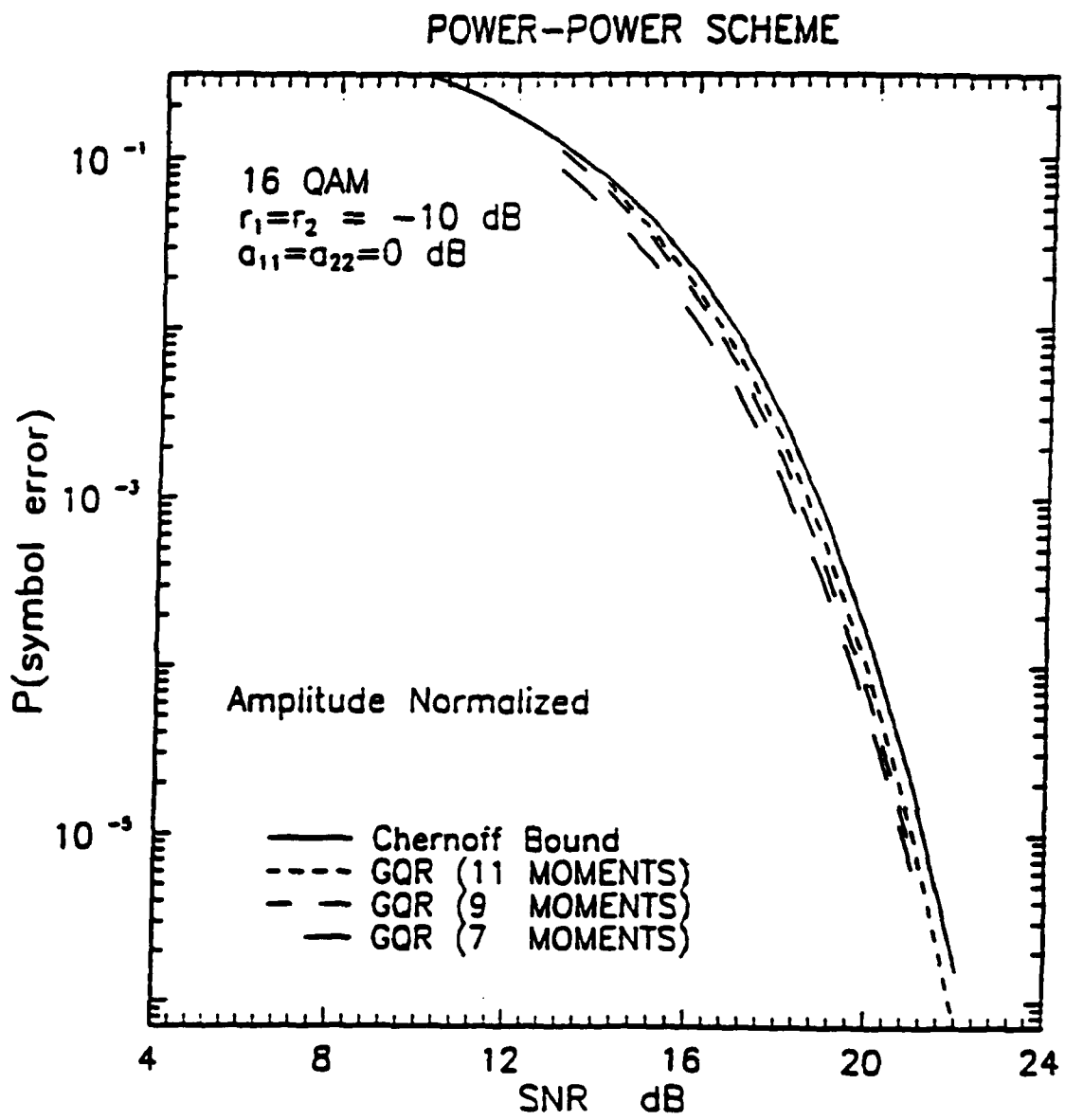


Fig. 19 Power-Power Cross-Pol Interference Canceler, effect of increasing of moments on GQR calculation results. 16 QAM

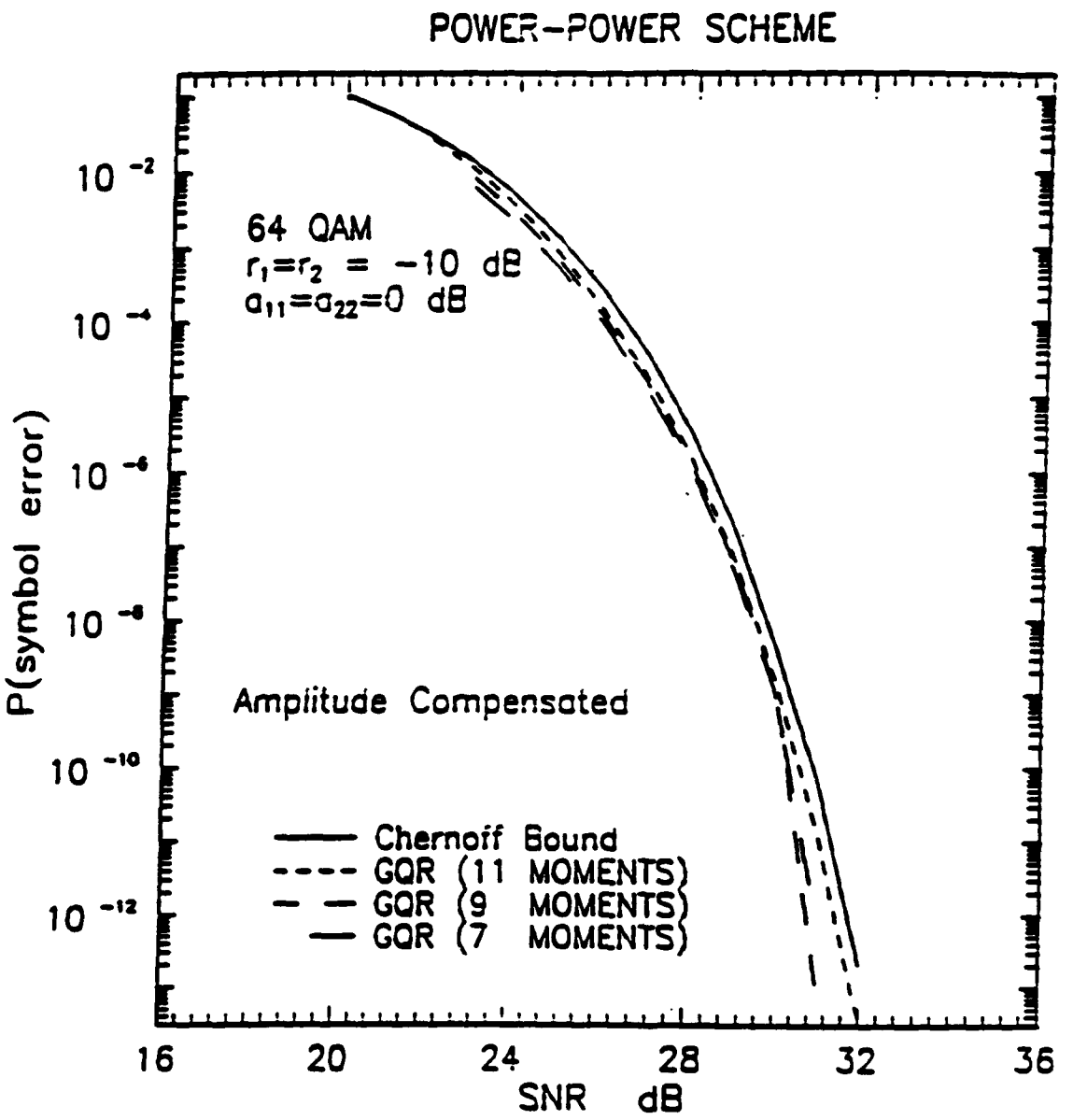


Fig. 20 Power-Power Cross-Pol Interference Canceler, effect of increasing of moments on GQR calculation results. 64 QAM

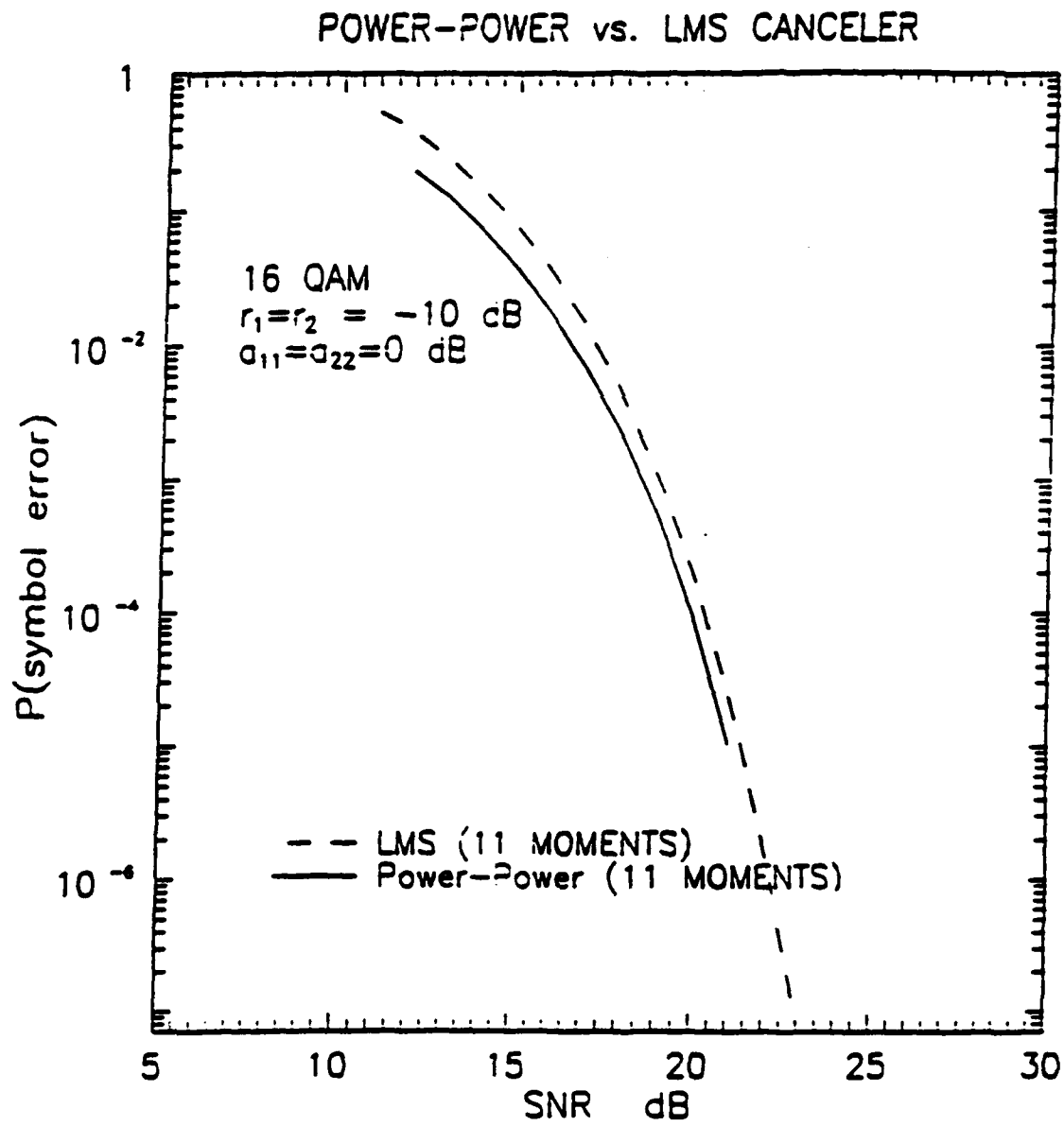


Fig. 21 Performance Comparison of Power-Power with LMS cancelers, GQR calculation, 16 QAM, with amplitude compensation, $r_1 = r_2 = -10$ dB.

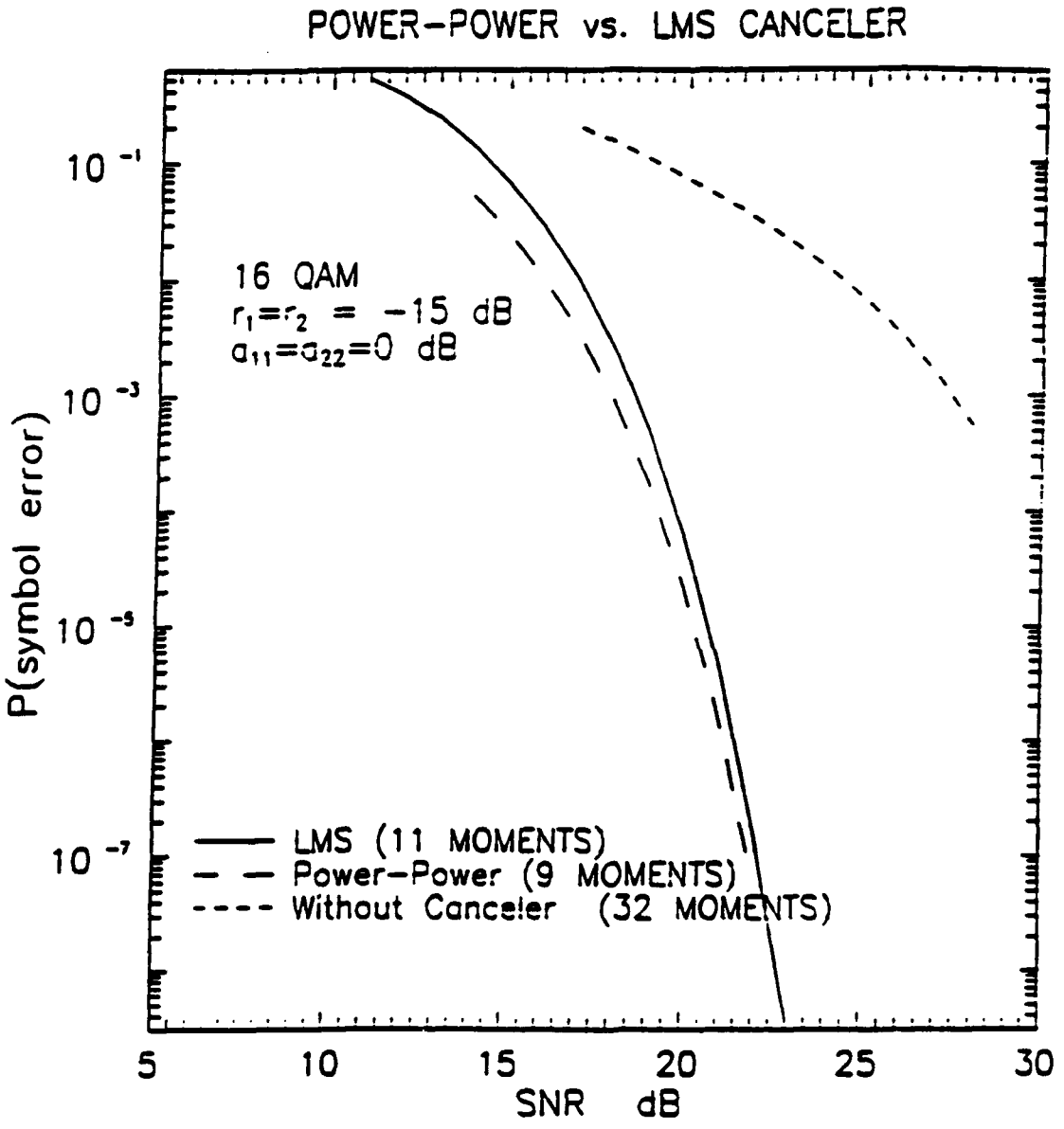


Fig. 22 Performance Comparison of Power-Power with LMS cancelers. GQR calculation, 16 QAM, with amplitude compensation, $r_1 = r_2 = -15$ dB.

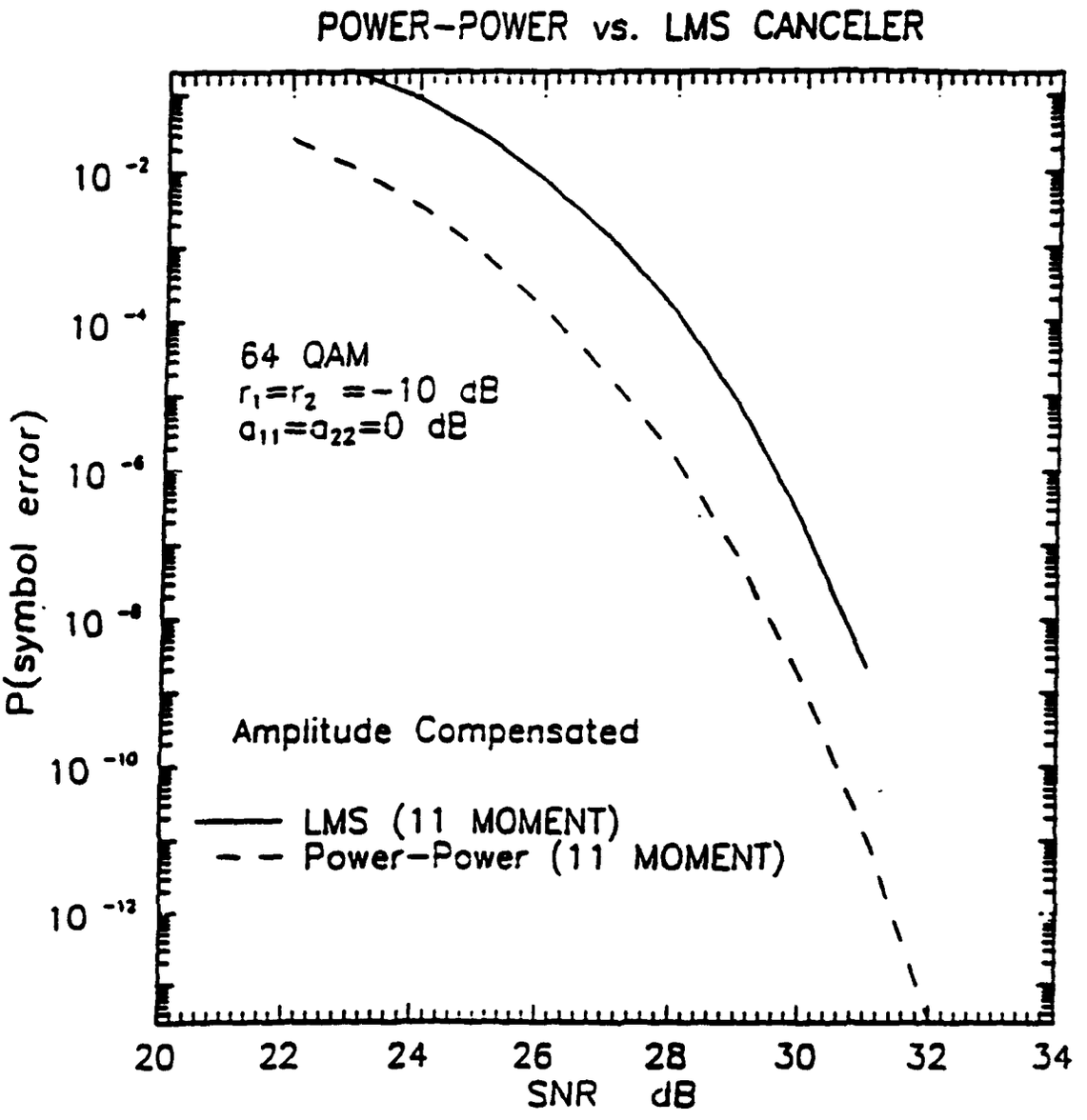


Fig. 23 Performance Comparison of Power-Power with LMS cancelers, GQR calculation, 64 QAM, with amplitude compensation, $r_1 = r_2 = -10$ dB.

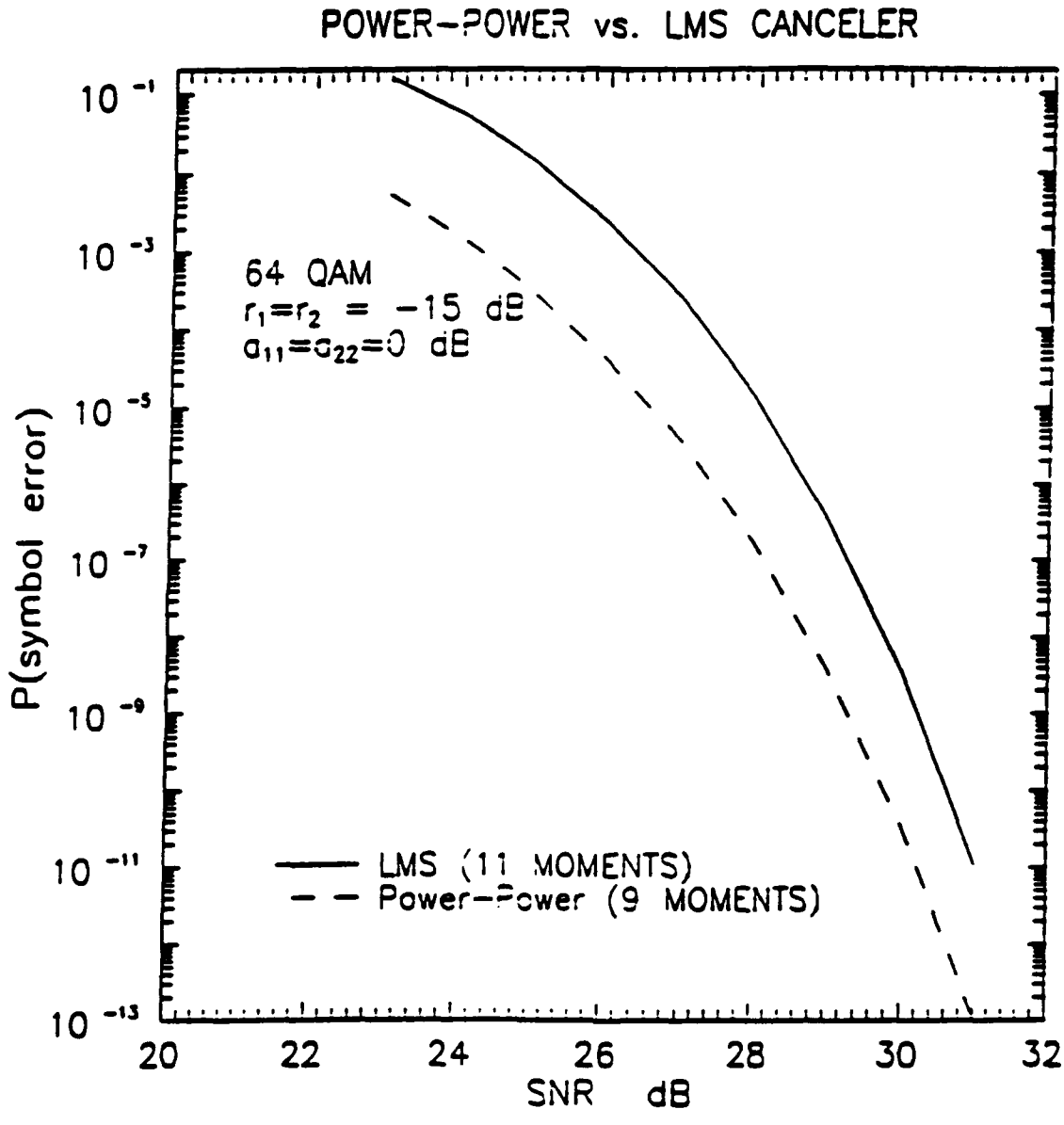


Fig. 24 Performance Comparison of Power-Power with LMS cancelers, GQR calculation, 64 QAM, with amplitude compensation, $r_1 = r_2 = -15$ dB.

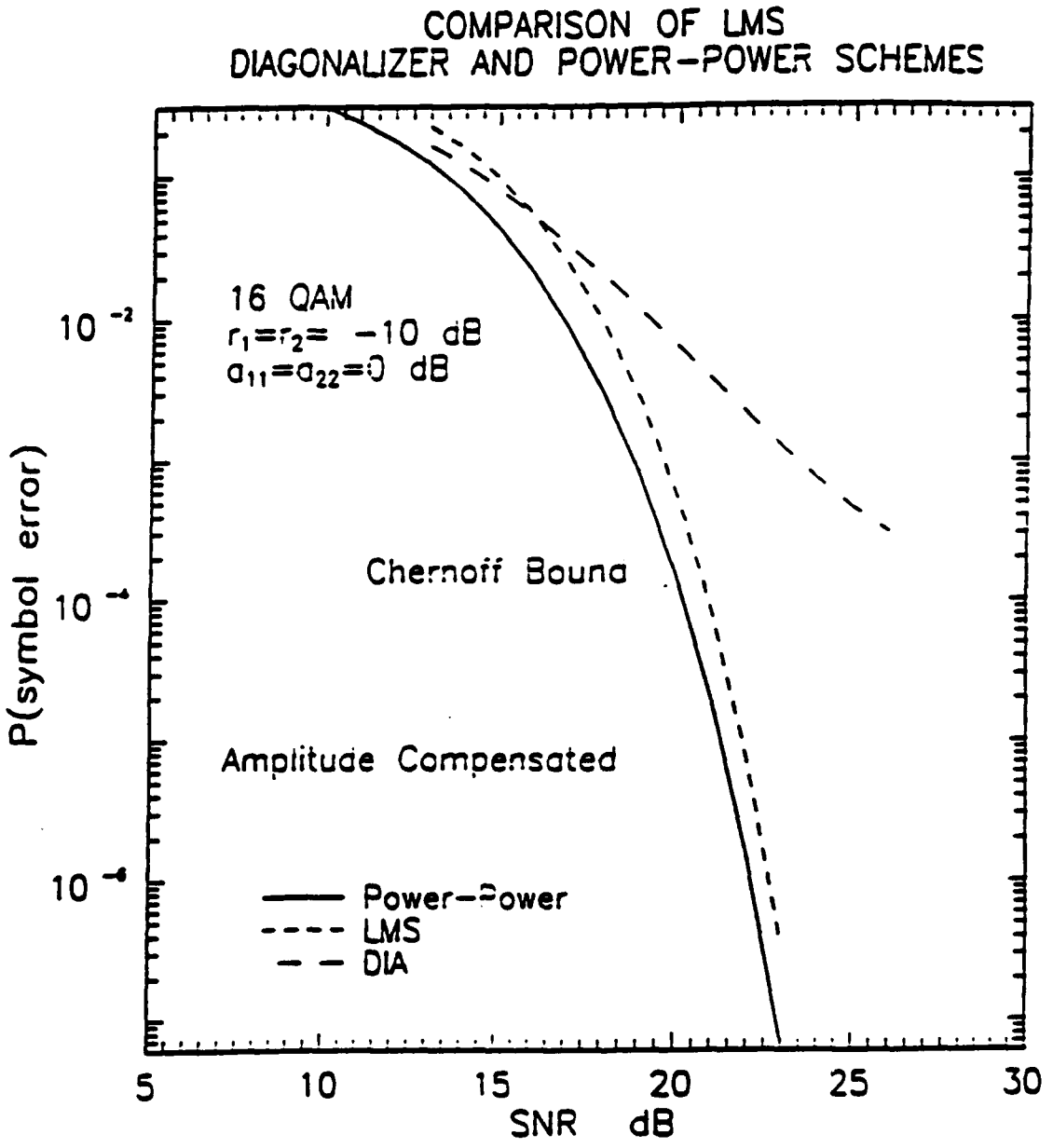


Fig. 25 Performance Comparison of LMS, Diagonalizer and Power-Power cancelers, Chernoff Bound . 16 QAM, with amplitude compensation, $r_1 = r_2 = -10$ dB.

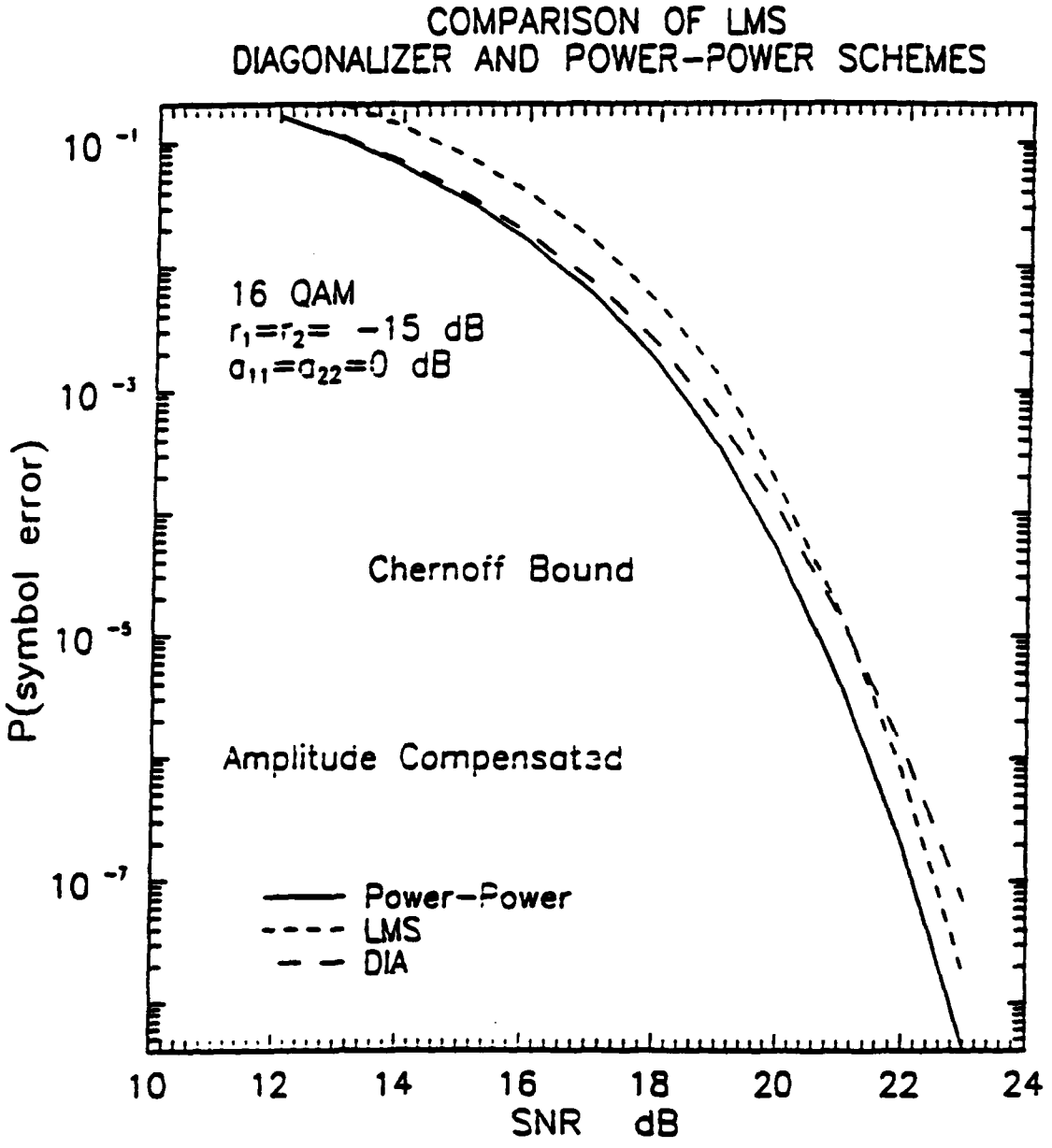


Fig. 26 Performance Comparison of LMS, Diagonalizer and Power-Power cancelers. Chernoff Bound . 16 QAM. with amplitude compensation, $r_1 = r_2 = -15$ dB.

Appendix D

Performance Comparison of Three Bootstrapped Cross-Pol Cancelers for M-ary QAM Signals

Abdulkadir Dinç and Yeheskel Bar-Ness

Abstract

M-QAM dual-polarized transmission became an important method for frequency re-use, particularly in microwave radio communication. However, the orthogonally polarized waves suffered degradation due to cross polarization interference.

Different canceler structures were proposed to mitigate the effect of cross-polarization. Among these are the diagonalizer and the LMS canceler. The bootstrapped algorithms have been suggested in the past for different applications such as: satellite dual-polarized communication, tactical communication and QAM microwave radio, as well as tactical communication. Nevertheless, no attempt has been made yet to quantify these cancelers' probability of error. These issues will be addressed in this appendix and will show that the power-power structure of the bootstrapped cross polarization always outperforms the other cancelers.

1 Introduction

In microwave radio communication networks, the transmission capacity can be doubled by using orthogonally polarized QAM carriers. The orthogonally polarized waves can suffer degradation due to carrier-to-cross polarization interference (C/XPI). Particularly during multipath fading, such degradation could become intolerable.

Many methods have been proposed to cancel the XPI. Among these are the diagonalizer [1] and the least mean square (LMS) cross-pol canceler proposed by Kavehrad [2]. In his paper, Kavehrad compares the error performance of these two cancelers and concludes that the LMS algorithm is substantially better than the diagonalizer.

Still another method for XPI is termed bootstrapped algorithm. It is important to note that the bootstrapped structure requires no reference signal nor any of the decision feedback methods and hence, it can be classified as a blind canceler; a distinct important feature.

The bootstrapped structures have been suggested in the past for different applications, such as in dual-polarized satellite communication [3], QAM microwave radio communication [4] and in tactical communication [5]. The error performance of one scheme of bootstrapped algorithms termed power-power, was presented in [6] and compared with that of LMS and diagonalizer. In this appendix, we intend to extend the performance analysis to the other structures of the bootstrapped blind canceler (BXPC) algorithms for M-ary QAM systems and compare the performance of these structures with each other.

In this study, we derive and compare the average probability of error of the BXPC as a function of its input signal-to-noise ratio for three different structures of this canceler. After this introduction we present the mathematical model of the dual polarized channel in section II. The different structures of the BXPC is described in section III. In section IV, a summary of the performance analysis of the BXPC for dual polarized M-ary QAM system over non-dispersive fading channel is presented.

In section V, the results of these analyses are shown and compared. Also included are the results of the comparison of the performance of BXPC with that of the other cancelers.

2 M-ary QAM Dual-Channel

The model for such a channel has been well presented in [1,2]. Two orthogonal M-ary QAM bandpass signals with the same bandwidth and the same center frequency can be presented as

$$s_i(t) = Re\{\tilde{s}_i(t)\exp(j2\pi f_c t)\} \tag{1}$$

where $Re\{\cdot\}$ stands for the real part, f_c denotes the carrier frequency and $\tilde{s}_i(t), i = 1,2$ is the complex envelope of each of the orthogonal signals, respectively. This complex envelope can be expressed as $\tilde{s}_i(t) = \sum_{k=0}^{\infty} I_k h(t-kT)$, where I_k is a complex information symbol which takes on one of M different statistically independent and equiprobable complex values. The effect of the channel is given by $\mathbf{x}(t) = A\mathbf{s}(t)+\mathbf{n}(t)$, where A is the dual-channel cross coupling matrix. whose elements $a_{ij}, i,j = 1,2$ are complex valued constants that denote the channels co-polarization and cross-polarization responses, $\mathbf{s}(t) = [s_1(t), s_2(t)]^T$ and $\mathbf{n}(t) = [n_1(t), n_2(t)]^T$ is a complex zero mean Gaussian noise vector.

The received signals which are sampled after matched filters, are denoted by

$$\begin{aligned} x_1(n) &= a_{11}I_1(n) + a_{12}I_2(n) + n_1(n), \\ x_2(n) &= a_{21}I_1(n) + a_{22}I_2(n) + n_2(n). \end{aligned} \tag{2}$$

where $x_1(n)$ and $x_2(n)$ are the sampled received signals at the output of the first and second channels respectively. $I_i(n)$ and $n_i(n)$ are the corresponding signals and noises at these outputs. Also $n_1(n)$ and $n_2(n)$ are assumed independent samples of the zero mean complex Gaussian with $E\{|n_i(n)|^2\} = 2\sigma_n^2, i = 1.2$.

3 The Bootstrapped Algorithms

Three configurations of bootstrapped structures, namely, power-power, correlation-correlation and power-correlation, depicted in Figs. 1-3 were partially reported in the open literature [3,4,5] . Each configuration consists of two cross coupled interference cancellation loops. The power-power scheme is so named because it consists of two control loops, each of which tries to minimize the power at its output. The correlation- correlation scheme similarly consists of two control loops, each of which tries to minimize the correlation between its output and its sample interference. The power-correlation scheme is a combination of the other two.

For the power-power scheme depicted in Fig. 1, as well as in the others, one interference cancellation loop attempts to cancel the interference from the signal $I_2(n)$ into the signal $I_1(n)$, while the other loop attempts to cancel the interference from $I_1(n)$ into $I_2(n)$. By virtue of a discrimination technique which slightly enhances the signal component due $I_2(n)$ over that due to $I_1(n)$, the first loop is able to perform a partial cancellation of the interference from $I_2(n)$ into $I_1(n)$. This results in a purer sample of $I_1(n)$ provided to the second interference cancellation loop, allowing that loop to perform a partial cancellation in the other direction, from $I_1(n)$ into $I_2(n)$. The purer sample of $I_2(n)$ is then used by the first loop to improve its cancellation, and so the above cycle is repeated until essentially perfect interference cancellation on both signals has been achieved. A similar argument shows that the other two schemes also lead to perfect interference cancellation. Although it is possible to consider more general terms, we will restrict the channel parameters to

$$\frac{a_{12}}{a_{22}} = r_1 e^{j\phi_1}, \quad \frac{a_{21}}{a_{11}} = r_2 e^{j\phi_2}. \tag{3}$$

where r_1, r_2 denote the magnitude of the normalized XPI constants and ϕ_1, ϕ_2 are the phases of these constants assumed to be independent and uniformly distributed over $[-\pi, \pi]$.

3.1 Power-Power Scheme

From Fig. 1, one can find the output of the canceler

$$y_1(n) = \frac{x_1(n) + x_2(n)w_{12}}{1 - w_{12}w_{21}}, \quad y_2(n) = \frac{x_2(n) + x_1(n)w_{21}}{1 - w_{12}w_{21}}. \tag{4}$$

where $x_1(n)$ and $x_2(n)$ are given in (2). The control algorithm simultaneously minimizes the output powers P_1 and Q_1, by searching for $\partial E\{|y_{1d}(n)|^2\}/\partial w_{12} = 0$ and $\partial E\{|y_{2d}(n)|^2\}/\partial w_{21} = 0$. This can be performed by a recursive steepest descent algorithm, provided that $1 - w_{12}w_{21} \neq 0$. $y_{id}(n)$ is the signal after the discriminator. It can be shown that w_{12} and w_{21} converge in the steady-state to

$$w_{12\text{opt}} = -\frac{a_{12}}{a_{22}} + \epsilon_1, \quad w_{21\text{opt}} = -\frac{a_{21}}{a_{11}} + \epsilon_2. \tag{5}$$

where ϵ_1 and ϵ_2 are the effect of the noise on the steady state weights. They depend on the channel parameters as well as on the signal and noise powers [7]. Substituting (5) into (4), we get for the first output

$$\begin{aligned} y_1(n) \; =& \frac{1}{\Delta_y}\Big[a_{11}I_1(n)[1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \epsilon_1\frac{a_{21}}{a_{11}}] + I_2(n)\epsilon_1 a_{22} \\ & + n_1(n) - n_2(n)\frac{a_{12}}{a_{22}} + n_2(n)\epsilon_1\Big], \end{aligned} \tag{6}$$

where

$$\Delta_y \; \triangleq \; 1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \frac{a_{12}}{a_{22}}\epsilon_2 + \frac{a_{21}}{a_{11}}\epsilon_1 - \epsilon_1\epsilon_2. \tag{7}$$

The co-pol polarized signal at the output of the channel is given by $a_{11}I_1(n)$, and hence it is reasonable to take $y_1(n)a_{11}$ as an estimate of this signal by compensating for the attenuation in the co-pol by a_{11}. Therefore $\hat{I}_1(n) = \frac{y_1(n)}{a_{11}}$ as an estimate of the transmitted signal $I_1(n)$, and the decision parameter for the first output of the canceler is defined as. $Z_1(n) \triangleq \hat{I}_1(n) - I_1(n)$. We will assume the $\epsilon_1\epsilon_2$ in (7) is negligible with respect to the other terms of this equation, and get

$$Z_1(n) \; = \; \frac{1}{\Delta_y}\Big[-I_1(n)\epsilon_2\frac{a_{12}}{a_{22}} + I_2(n)\epsilon_1\frac{a_{22}}{a_{11}} + \frac{n_1(n)}{a_{11}}$$

$$+\frac{n_2(n)}{a_{11}}\left(\frac{-a_{12}}{a_{22}}+\epsilon_1\right)\Big], \tag{8}$$

where

$$\Delta_y \approx 1-\frac{a_{12}a_{21}}{a_{11}a_{22}}+\frac{a_{12}}{a_{22}}\epsilon_2+\frac{a_{21}}{a_{11}}\epsilon_1.$$

The probability of error is given by $P_1(e)=P\{|Z_1(n)|>c\}$ with c as the half of the distance between any two symbols in the corresponding signal space. Note that $Z_1(n)$ depends on the random signal and noise variables $I_i(n)$ and $n_i(n)$ $i=1,2$ and on the random variables ϕ_1 and ϕ_2 through the channel parameter and through the weight perturbation ϵ_1 and ϵ_2. Therefore, in order to be able to calculate the probability of error, we must find the real and imaginary part of $Z_1(n)$. Algebraic manipulation [7] leads to

$$\begin{aligned} Z_{1R} &= I_{1R}Y_1+I_{1I}Y_2+I_{2R}Y_3+I_{2I}Y_4+n_{1R}Y_5 \\ &\quad +n_{1I}Y_6+n_{2R}Y_7+n_{2I}Y_8, \\ Z_{1I} &= -I_{1R}Y_2+I_{1I}Y_1-I_{2R}Y_4+I_{2I}Y_3-n_{1R}Y_6 \\ &\quad +n_{1I}Y_5-n_{2R}Y_8+n_{2I}Y_7. \end{aligned} \tag{9}$$

where we dropped the sampling time n for simplicity, and Y_m, $m=1,2,..8$ depend only on the random variables ϕ_1 and ϕ_2. Expressions for Y_m are given in the appendix.

3.2 Correlation-Correlation Scheme

The outputs $y_1(n)$ and $y_2(n)$ from Fig. 2 are as follows

$$\begin{aligned} y_1(n) &= x_1(n)+x_2(n)w_{12}, \\ y_2(n) &= x_2(n)+x_1(n)w_{21}. \end{aligned} \tag{10}$$

With this scheme, the algorithm simultaneously minimizes the correlation between two outputs. It searches for $\partial|E\{y_{1d}(n)y_2^*(n)\}|^2/\partial w_{12}=0$ and $\partial|E\{y_{2d}(n)y_1^*(n)\}|^2/\partial w_{21}=0$. As in the previous scheme, it can be shown that w_{12} and w_{21} converge in the steady

state to the same values as in (5). Following the same steps as in the power-power scheme, we get the decision parameter

$$\begin{aligned} Z_1(n) &= I_1(n)[-\frac{a_{12}a_{21}}{a_{22}a_{11}} + \frac{\epsilon_1 a_{21}}{a_{11}}] + \frac{1}{a_{11}}\Big[I_2(n)a_{22}\epsilon_1 \\ &\quad + n_1(n) + n_2(n)[\frac{-a_{12}}{a_{22}} + \epsilon_1]\Big]. \end{aligned} \tag{11}$$

Algebraic manipulations lead to

$$\begin{aligned} Z_{1R} &= I_{1R}Y_{c1} + I_{1I}Y_{c2} + I_{2R}Y_{c3} + I_{2I}Y_{c4} + n_{1R}Y_{c5}, \\ &\quad + n_{2R}Y_{c6} + n_{2I}Y_{c7}, \\ Z_{1I} &= -I_{1R}Y_{c2} + I_{1I}Y_{c1} - I_{2R}Y_{c4} + I_{2I}Y_{c3} + n_{1I}Y_{c5}. \\ &\quad - n_{2R}Y_{c7} + n_{2I}Y_{c6}. \end{aligned} \tag{12}$$

where again Y_{cm} $m = 1.2,..7$ depends only on the random variables ϕ_1 and φ_2. Expressions for Y_{cm} are given in the appendix.

3.3 Power-Correlation Scheme

The outputs $y_1(n)$ and $y_2(n)$ from Fig. 3 are as follows

$$\begin{aligned} y_1(n) &= x_1(n)(1 + w_{12}w_{21}) + x_2(n)w_{12}, \\ y_2(n) &= x_2(n) + x_1(n)w_{21}. \end{aligned} \tag{13}$$

The control algorithm simultaneously minimizes the output power $E\{|y_{1d}(n)|^2\}$ and the square magnitude of the correlation of output $y_{2d}(n)$ with $y_1(n)$. It searches for $\partial E\{|y_{1d}(n)|^2\}/\partial w_{12} = 0$ and $\partial |E\{y_{2d}(n)y_1^*(n)\}|^2/\partial w_{21} = 0$. It can be shown that w_{12} and w_{21} converge in the steady state to

$$w_{12\text{opt}} = -\frac{a_{12}}{a_{22}[1 - \frac{a_{12}a_{21}}{a_{22}a_{11}}]} + \epsilon_3, \quad w_{21\text{opt}} = -\frac{a_{21}}{a_{11}} + \epsilon_4. \tag{14}$$

Unlike the other two schemes of bootstrapped cancelers. the power-correlator scheme is not symmetrical; its outputs and hence, the decision parameters for each output are

different. From (14) with the substitution of $w_{12\text{opt}}$ and $w_{21\text{opt}}$ in (13), we obtain

$$\begin{aligned}
y_1(n) &= I_1(n)a_{11}[1-\frac{a_{12}}{a_{22}k}\epsilon_4+\epsilon_3\epsilon_4]+I_2(n)a_{22}\cdot \\
&\quad [-(\frac{a_{12}}{a_{22}})^2\frac{1}{k}\epsilon_4+k\epsilon_3+\epsilon_3\epsilon_4\frac{a_{12}}{a_{22}}]+n_1(n)[1+\frac{a_{12}a_{21}}{a_{22}a_{11}}\frac{1}{k} \\
&\quad -\frac{a_{21}}{a_{11}}\epsilon_3-\frac{a_{12}}{a_{22}}\frac{\epsilon_2}{k}+\epsilon_3\epsilon_4]+n_2(n)[\frac{-a_{12}}{a_{22}k}+\epsilon_1], \\
y_2(n) &= I_1(n)\epsilon_4 a_{11}+I_2(n)a_{22}[1-\frac{a_{21}a_{12}}{a_{11}a_{22}}+\epsilon_4\frac{a_{12}}{a_{22}}] \\
&\quad +n_2(n)+n_1(n)[\frac{-a_{21}}{a_{11}}+\epsilon_4]. \qquad (15)
\end{aligned}$$

and the corresponding decision parameters

$$\begin{aligned}
Z_1(n) &= -I_1(n)\frac{a_{12}}{a_{22}k}\epsilon_4+\frac{1}{a_{11}k}\Big[I_2(n)a_{22}[-(\frac{a_{12}}{a_{22}})^2\epsilon_4 \\
&\quad +\epsilon_3k^2]+n_1(n)[1-\frac{a_{21}}{a_{11}}\epsilon_3k-\frac{a_{12}}{a_{22}}\epsilon_4] \\
&\quad +n_2(n)[-\frac{a_{12}}{a_{22}}+k\epsilon_3]\Big], \\
Z_2(n) &= \frac{1}{a_{22}}\Big[I_1(n)a_{11}\epsilon_4+I_2(n)[-\frac{a_{12}a_{21}}{a_{11}a_{22}}+\epsilon_4\frac{a_{12}}{a_{22}}] \\
&\quad +n_2(n)+n_1(n)[-\frac{a_{21}}{a_{11}}+\epsilon_4]\Big]. \qquad (16)
\end{aligned}$$

where $k \triangleq 1-\frac{a_{12}a_{21}}{a_{11}a_{22}}$. The real and imaginary parts of $Z_1(n)$ and $Z_2(n)$ are given by

$$\begin{aligned}
Z_{1R} &= I_{1R}Y_{11}+I_{1I}Y_{12}+I_{2R}Y_{13}+I_{2I}Y_{14}+n_{1R}Y_{15} \\
&\quad +n_{1I}Y_{16}+n_{2R}Y_{17}+n_{2I}Y_{18}, \\
Z_{1I} &= -I_{1R}Y_{12}+I_{1I}Y_{11}-I_{2R}Y_{14}+I_{2I}Y_{13}-n_{1R}Y_{16} \\
&\quad +n_{1I}Y_{15}-n_{2R}Y_{18}+n_{2I}Y_{17}. \qquad (17)
\end{aligned}$$

$$\begin{aligned}
Z_{2R} &= I_{1R}Y_{21}+I_{1I}Y_{22}+I_{2R}Y_{23}+I_{2I}Y_{24}+n_{1R}Y_{25} \\
&\quad +n_{1I}Y_{26}+n_{2R}Y_{27},
\end{aligned}$$

$$Z_{2I} = -I_{1R}Y_{22} + I_{1I}Y_{21} - I_{2R}Y_{24} + I_{2I}Y_{23} - n_{1R}Y_{26} + n_{1I}Y_{25} + n_{2I}Y_{27}. \quad (18)$$

where as before Y_{1l}, $l = 1, 2, ..8$ and Y_{2m}, $m = 1, 2, ..7$ [7] depends only on the random variables ϕ_1 and ϕ_2.

4 Performance Analysis

Using the decision parameter of equation (9), (12), (17) and (18) for the three structures, respectively, we have for the average symbol error probability

$$P_1(e) = \frac{1}{2}\left[1 - \frac{1}{\sqrt{M}}\right]\{P(|Z_{1R}| > c) + P(|Z_{1I}| > c)\}. \quad (19)$$

The least upper bound on $P(|Z_{1R}| > c)$, called the Chernoff bound [8], would then be given by

$$P_i(e) \leq (1 - \frac{1}{\sqrt{M}})\frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \cdot \exp\left[\frac{-3\,(SNR)}{2(M-1)}\frac{1}{(SNR)U_i(\phi_1,\phi_2) + W_i(\phi_1,\phi_2)}\right] d\phi_1 d\phi_2, \quad (20)$$

where, $SNR = \frac{M-1}{3}\frac{c^2}{\sigma_n^2}$ [8], and $U_i(\phi_1, \phi_2)$ and $W_i(\phi_1, \phi_2)$, $i = 1, 2$ differ for different schemes [7].

The power-power scheme

$$U_i(\phi_1, \phi_2) = Y_1^2 + Y_2^2 + Y_3^2 + Y_4^2,$$
$$W_i(\phi_1, \phi_2) = Y_5^2 + Y_6^2 + Y_7^2 + Y_8^2. \quad (21)$$

where Y_m, $m = 1, 2, ..8$ is defined in (25).

The correlator-correlator scheme

$$\begin{aligned} U_i(\phi_1,\phi_2) &= Y_{c1}^2 + Y_{c2}^2 + Y_{c3}^2 + Y_{c4}^2 \\ W_i(\phi_1,\phi_2) &= Y_{c5}^2 + Y_{c6}^2 + Y_{c7}^2 \end{aligned} \tag{22}$$

where Y_m, $m = 1,2,..7$ is defined in (27).

Due to the symmetric structure of the power-power and correlator-correlator schemes, the error probabilities at both outputs of each corresponding scheme are the same for each structure.

The power-correlator scheme for output-1

$$\begin{aligned} U_1(\phi_1,\phi_2) &= Y_{11}^2 + Y_{12}^2 + Y_{13}^2 + Y_{14}^2 \\ W_1(\phi_1,\phi_2) &= Y_{15}^2 + Y_{16}^2 + Y_{17}^2 + Y_{18}^2 \end{aligned} \tag{23}$$

where Y_{1l}, $l = 1,2,..8$ *is defined in* [7].

For output-2

$$\begin{aligned} U_2(\phi_1,\phi_2) &= Y_{21}^2 + Y_{22}^2 + Y_{23}^2 + Y_{24}^2 \\ W_2(\phi_1,\phi_2) &= Y_{25}^2 + Y_{26}^2 + Y_{27}^2 \end{aligned} \tag{24}$$

where Y_{2m} $m = 1,2,..7$ is defined in [7].

5 Results and Conclusion

The Chernoff upper bound (20) on the average probability of error as a function of signal-to-noise ratio (SNR) is calculated for various cross coupling constants for 16 and 64 QAM signals and for the three different structures of the blind bootstrapped cross polarization canceler.

The performance of these three structures is depicted in Fig. 4, 5, 6 and 7 for the 16 QAM with r= -15 dB and -10 dB. The bounds are higher for 64 QAM indicating

a possibility of higher error rates for the same SNR. In Fig. 8, we compare the performances of these three structures of BXPC with that of the LMS cross polarization canceler.

From these figures, we conclude that the power-power scheme outperforms the correlator-correlator and power-correlator schemes, particularly, when the cross coupling is high , such as -10 dB. Also noticeable is the fact that the performance at the two outputs of the power-correlator scheme are not the same. The performance of one output is the same as that of the power-power scheme and the other output performance is close to the performance of the correlator-correlator structure. In Fig. 8, it is clearly shown that the power-power scheme performs better than the others, while LMS performs better than that of the correlator-correlator and that of output-2 of the power-correlator scheme.

Appendix

The Power-Power Scheme

$$
\begin{aligned}
Y_1 &= -\frac{K_R Z_{DR} + K_I Z_{DI}}{|Z_D|^2},\ Y_2 = \frac{K_I Z_{DR} - K_R Z_{DI}}{|Z_D|^2}, \\
Y_3 &= \frac{a_{22}}{a_{11}} \frac{\epsilon_{1R} Z_{DR} + \epsilon_{1I} Z_{DI}}{|Z_D|^2}, \\
Y_4 &= \frac{a_{22}}{a_{11}} \frac{\epsilon_{1R} Z_{DI} - \epsilon_{1I} Z_{DR}}{|Z_D|^2}, \\
Y_5 &= \frac{Z_{DR}}{a_{11}|Z_D|^2},\ Y_6 = \frac{Z_{DI}}{a_{11}|Z_D|^2}, \\
Y_7 &= \frac{(\epsilon_{1R} - r_1 cos o_1) Z_{DR} + (\epsilon_{1I} - r_1 sin \phi_1) Z_{DI}}{a_{11}|Z_D|^2}, \\
Y_8 &= \frac{(r_1 sin \phi_1 - \epsilon_{1I}) Z_{DR} + (\epsilon_{1R} - r_1 cos \phi_1) Z_{DI}}{a_{11}|Z_D|^2}.
\end{aligned} \tag{25}
$$

$$
\begin{aligned}
Z_D &= Z_{DR} + j Z_{DI},\ K = K_R + j K_I, \\
V &= V_R + j V_I,\ \epsilon_1 = \epsilon_{1R} + j \epsilon_{1I} \\
\epsilon_2 &= \epsilon_{2R} + j \epsilon_{2I}
\end{aligned} \tag{26}
$$

where

$$Z_D \approx 1 - \frac{a_{12}a_{21}}{a_{11}a_{22}} + \frac{a_{12}}{a_{22}}\epsilon_2 + \frac{a_{21}}{a_{11}}\epsilon_1,$$
$$K \triangleq \frac{a_{12}}{a_{22}}\epsilon_2, \quad V \triangleq \frac{a_{21}}{a_{11}}\epsilon_1.$$

The Correlator-Correlator Scheme

$$\begin{aligned}
Y_{c1} &= V_R - r_1 r_2 cos(\phi_1 + \phi_2), \\
Y_{c2} &= -[V_I - r_1 r_2 sin(\phi_1 + \phi_2)], \\
Y_{c3} &= \frac{a_{22}}{a_{11}}\epsilon_{1R}, \quad Y_{c4} = -\frac{a_{22}}{a_{11}}\epsilon_{1I}, \quad Y_{c5} = \frac{1}{a_{11}}, \\
Y_{c6} &= \frac{1}{a_{11}}[\epsilon_{1R} - r_1 cos\phi_1], \\
Y_{c7} &= -\frac{1}{a_{11}}[\epsilon_{1I} - r_1 sin\phi_1],
\end{aligned} \tag{27}$$

and

$$\begin{aligned}
\epsilon_1 &= \frac{1}{Z_D}a_{11}^2\Big[r_1 e^{j\phi_1} + r_2 e^{-j\phi_2} - r_1^2 r_2 e^{j(2\phi_1+\phi_2)} \\
&\quad -r_1 r_2^2 e^{j\phi_1}\Big].E\{|n(n)|^2\}E\{|I_1(n)|^2\}(\delta_{21} - \delta_{11}) \\
\epsilon_2 &= \frac{1}{Z_D^*}a_{22}^2\Big[r_1 e^{-j\phi_1} + r_2 e^{j\phi_2} - r_1 r_2^2 e^{j(\phi_1+2\phi_2)} \\
&\quad -r_1^2 r_2 e^{j\phi_2}\Big].E\{|n(n)|^2\}E\{|I_2(n)|^2\}(\delta_{12} - \delta_{22})
\end{aligned}$$

where $\delta_{ij} \triangleq \frac{E\{|y_{jd}(n)|^2\}}{E\{|y_j(n)|^2\}}$ is the discrimination constant which is the ratio of the mean square of j th output signal at the output of the i th discriminator to the j th signal. and $y_{jd}(n)$ is the signal after the discriminator.

References

[1] Amitay N., "Signal-to-Noise Ratio Statistics for Nondispersive Fading in Radio Channels with Cross Polarization Interference Cancellation," *IEEE Trans. Commun.*, Vol.27, pp 498-502 Feb.1979.

[2] Kavehrad M., "Performance of Cross-Polarized M-ary QAM Signals Over Nondispersive Fading Channels," AT&T *Bell Lab. Tech. J.*, Vol 63, pp 499-521 March 1984.

[3] Bar-Ness Y. , Carlin, J.W and Steinberger, M.L. , "Bootstrapping Adaptive Cross-Pol Canceller for Satellite Communications," *Int. Conf. on Comm.*, June 1982 Philadelphia, PA

[4] Carlin, J. W., Bar-Ness, Y., Gross, S., Steinberger, M. S and Studdiford, W. E., "An IF Cross-Pol Canceller for Microwave Radio," *Journal on Selected Areas in Communication - Advances in Digital Communication by Radio* , vol. SAC-5, No. 3, pp. 502-514, April 1987

[5] Bar-Ness Y., "Bootstrapped Algorithm for Interference Cancellation," *AFCEA - IEEE Technical Conf. Tactical Comm.* Fort Wayne, Indiana, May 1988

[6] Dinç A. and Bar-Ness Y., "Performance Comparison of LMS Diagonalizer and Bootstrapped Adaptive Cross-Pol Cancellers For M-ary QAM," *Proceedings of MILCOM 90*, paper 3.7, Monterey, CA Oct.1990

[7] Dinç A., 1991. "The Theory of Bootstrapped Algorithms and Their Applications to Cross Polarization Interference Cancellation," Ph.D. Dissertation, ECE Dept., New Jersey Institute of Tech., Newark, NJ

[8] I. Korn, *Digital Communication*, Van Nostrand Reinhold Company, pp. 355-430, 1985

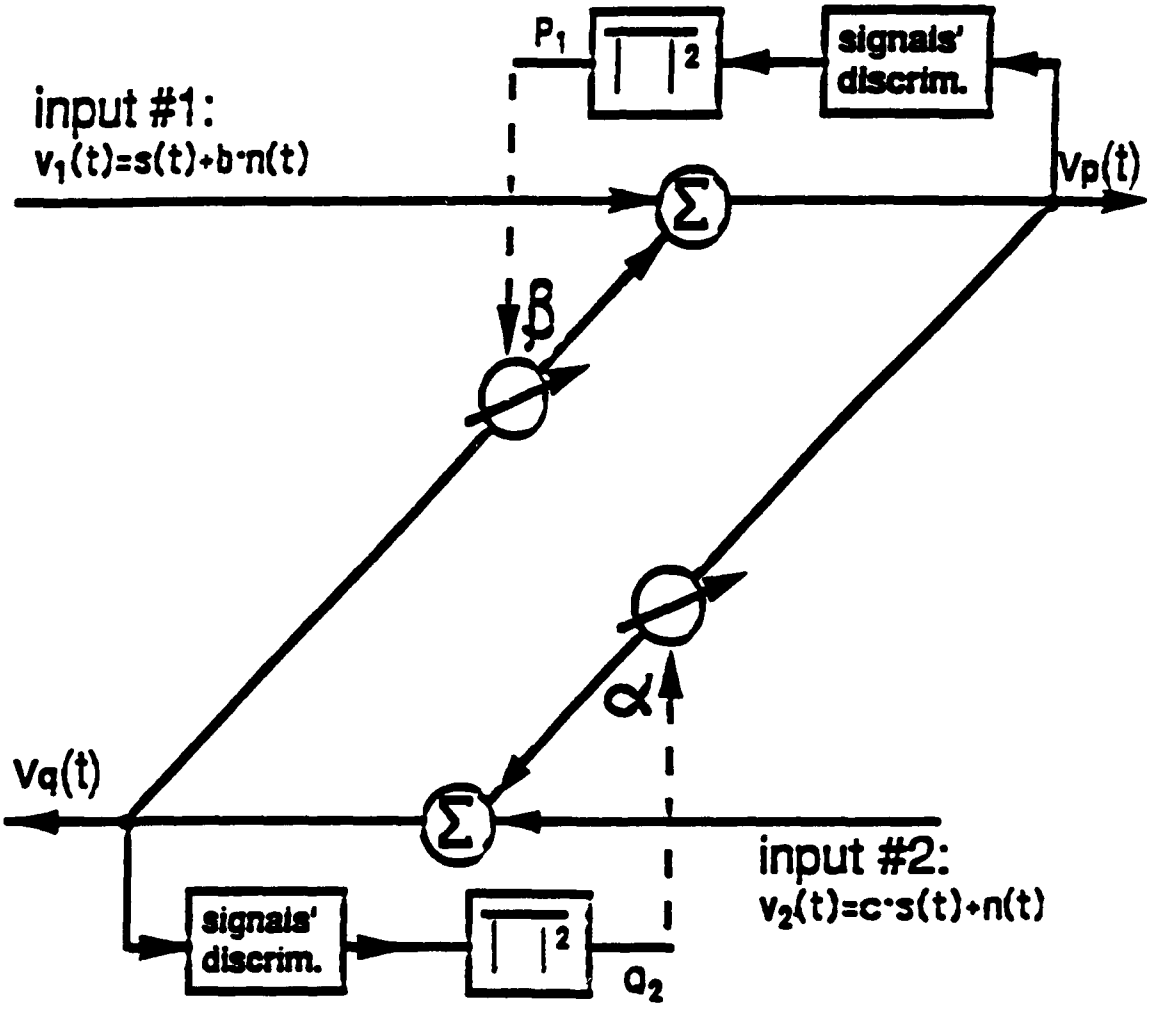


Fig. 1 Power-Power Cross-Polarization Canceler

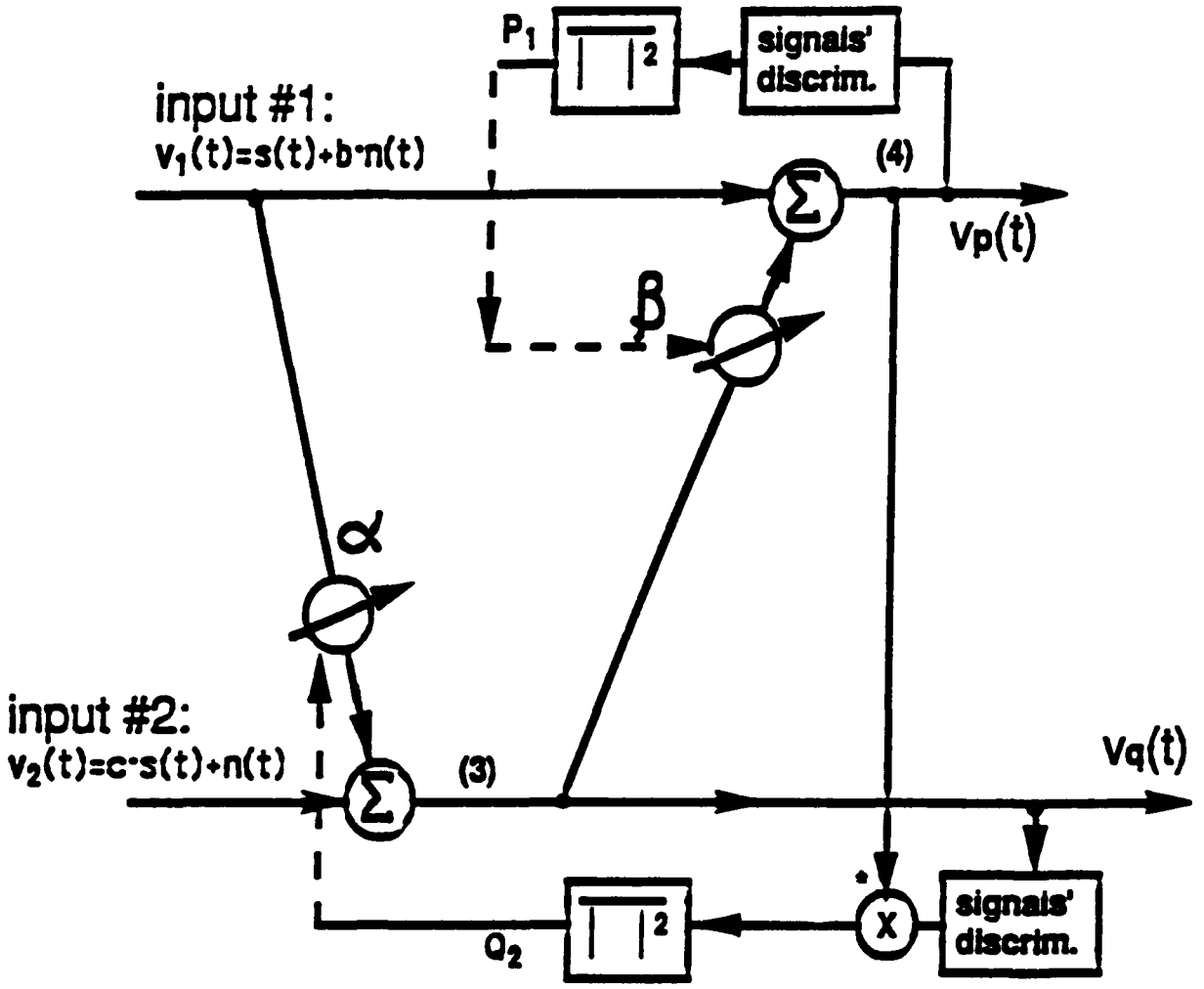


Fig. 2 Power-Correlator Cross-Polarization Canceler

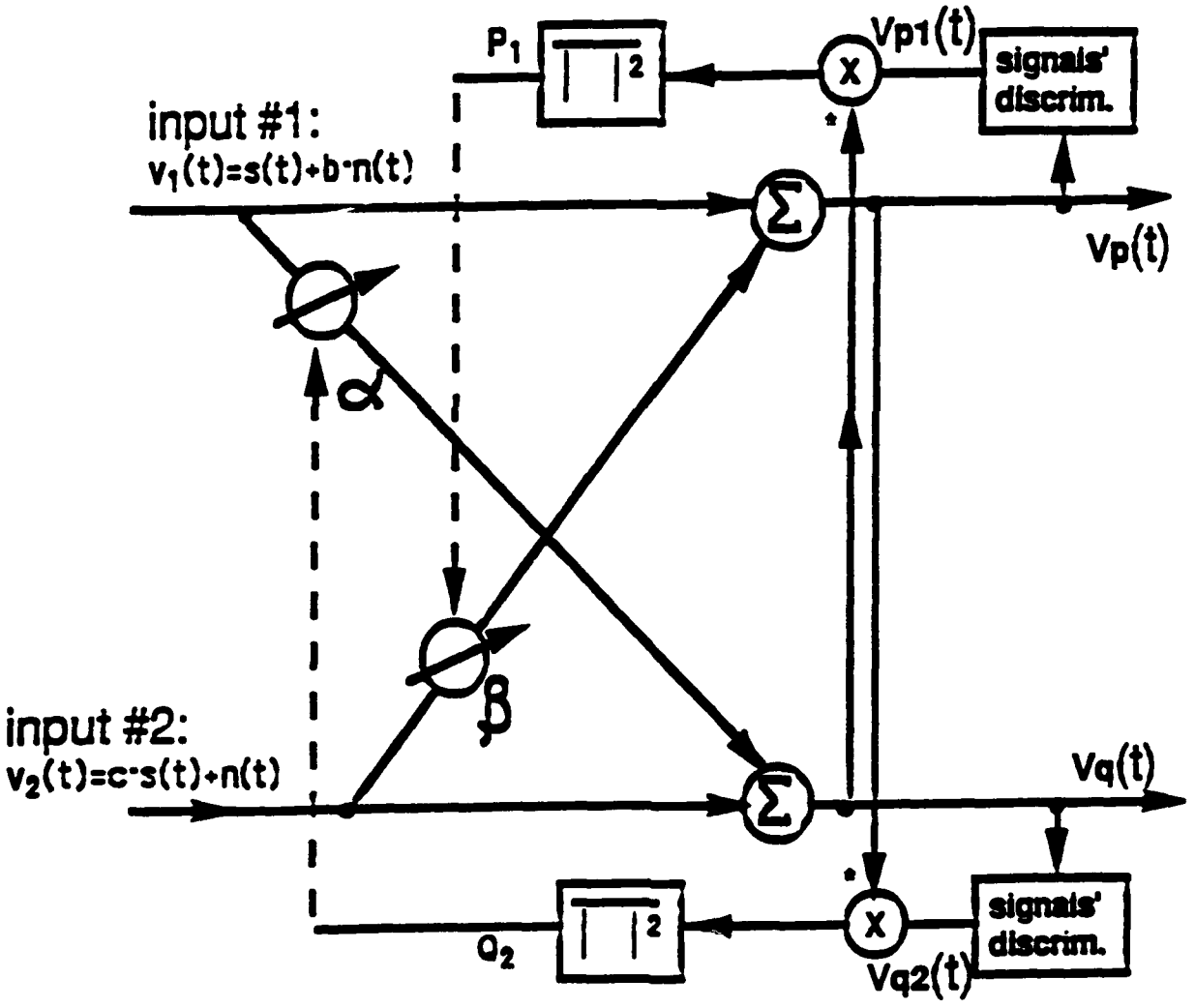


Fig. 3 Correlator-Correlator Cross-Polarization Canceler

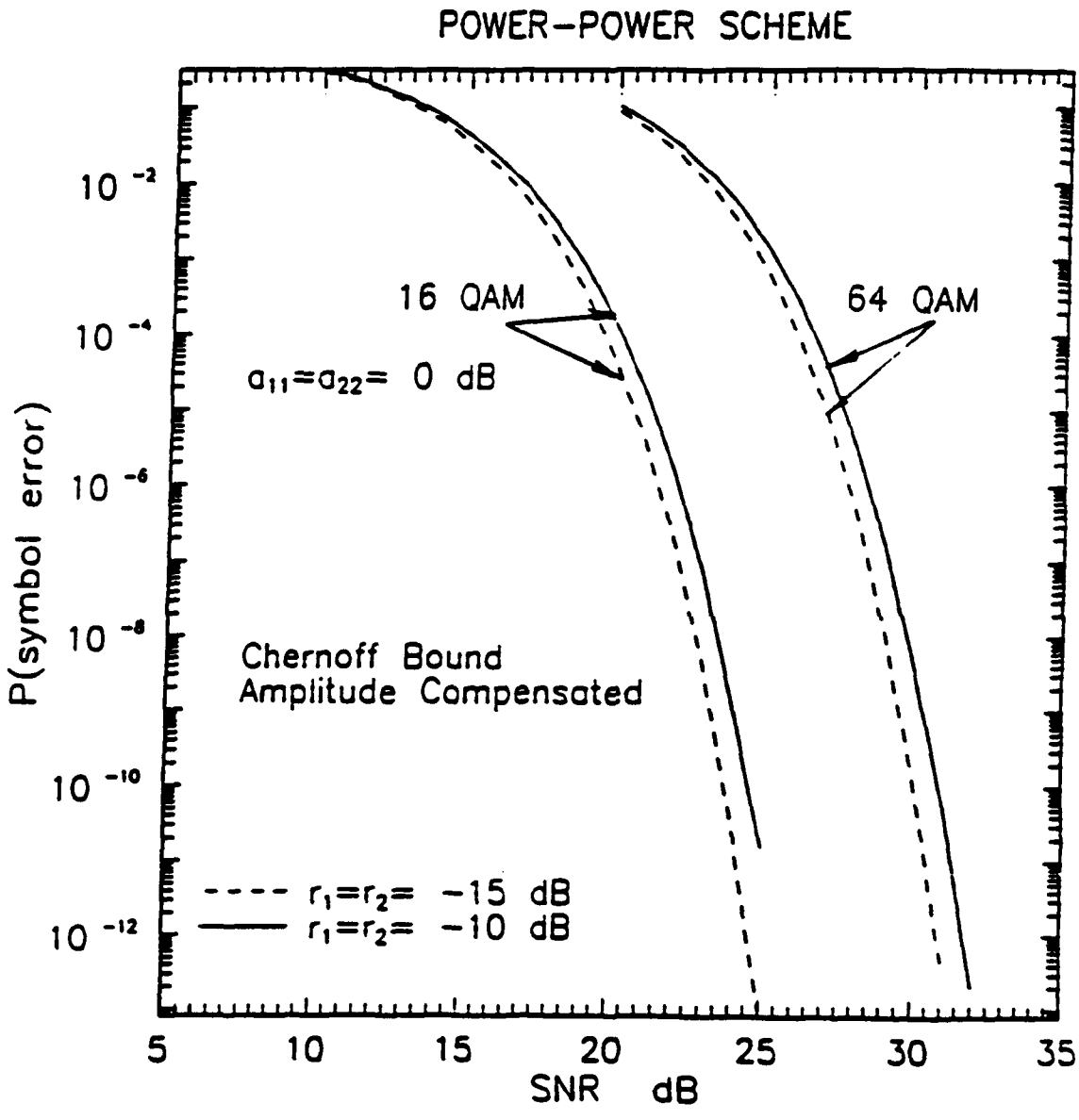


Fig 4. Power-Power Cross-Pol Canceler, Chernoff Bound comparison 16 QAM vs. 64 QAM

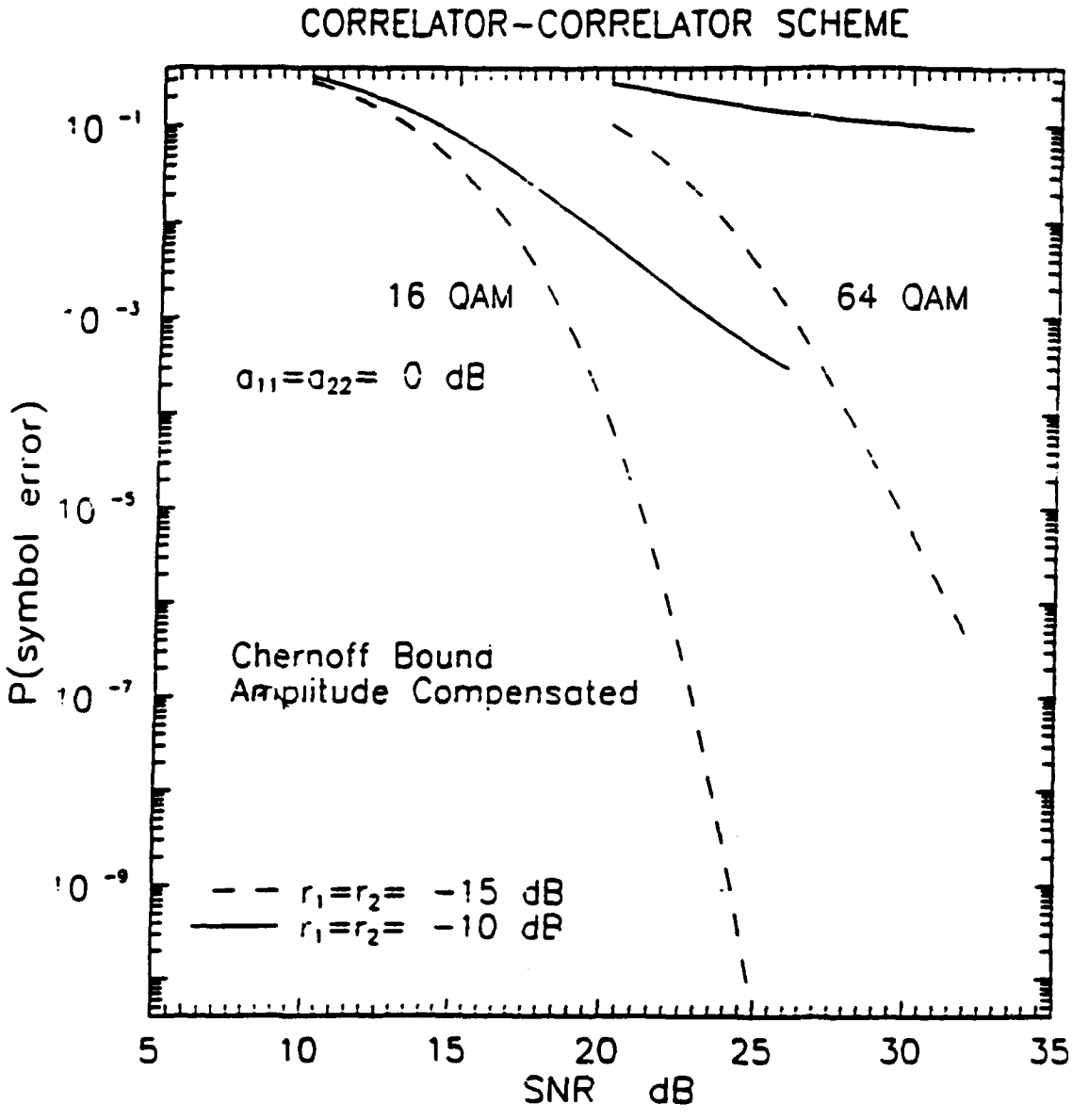


Fig 5. Correlator-Correlator Cross-Pol Canceler. Chernoff Bound comparison 16 QAM vs. 64 QAM

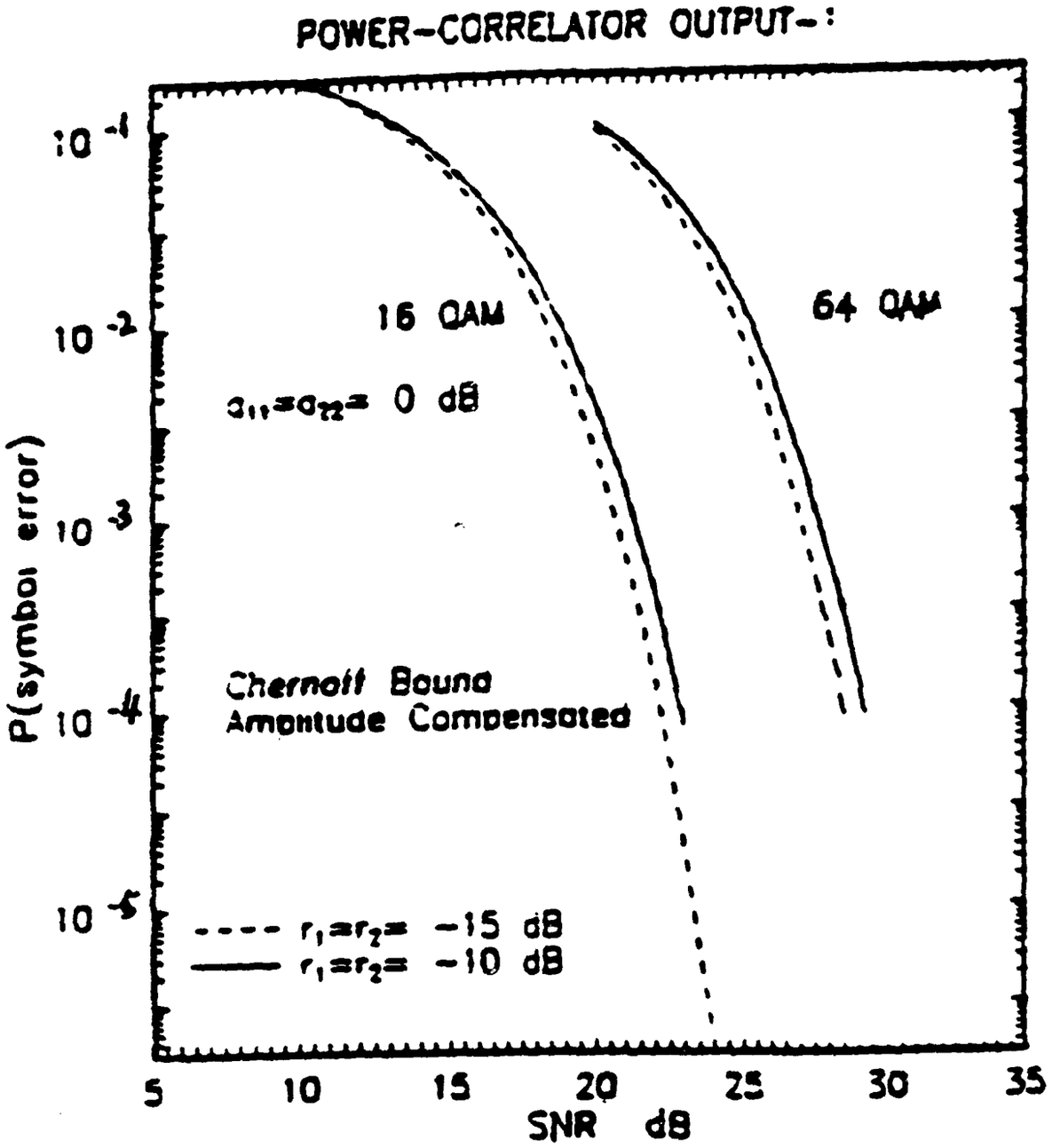


Fig 6. Power-Correlator Cross-Pol Canceler output-1. Chernoff Bound comparison 16 QAM vs. 64 QAM

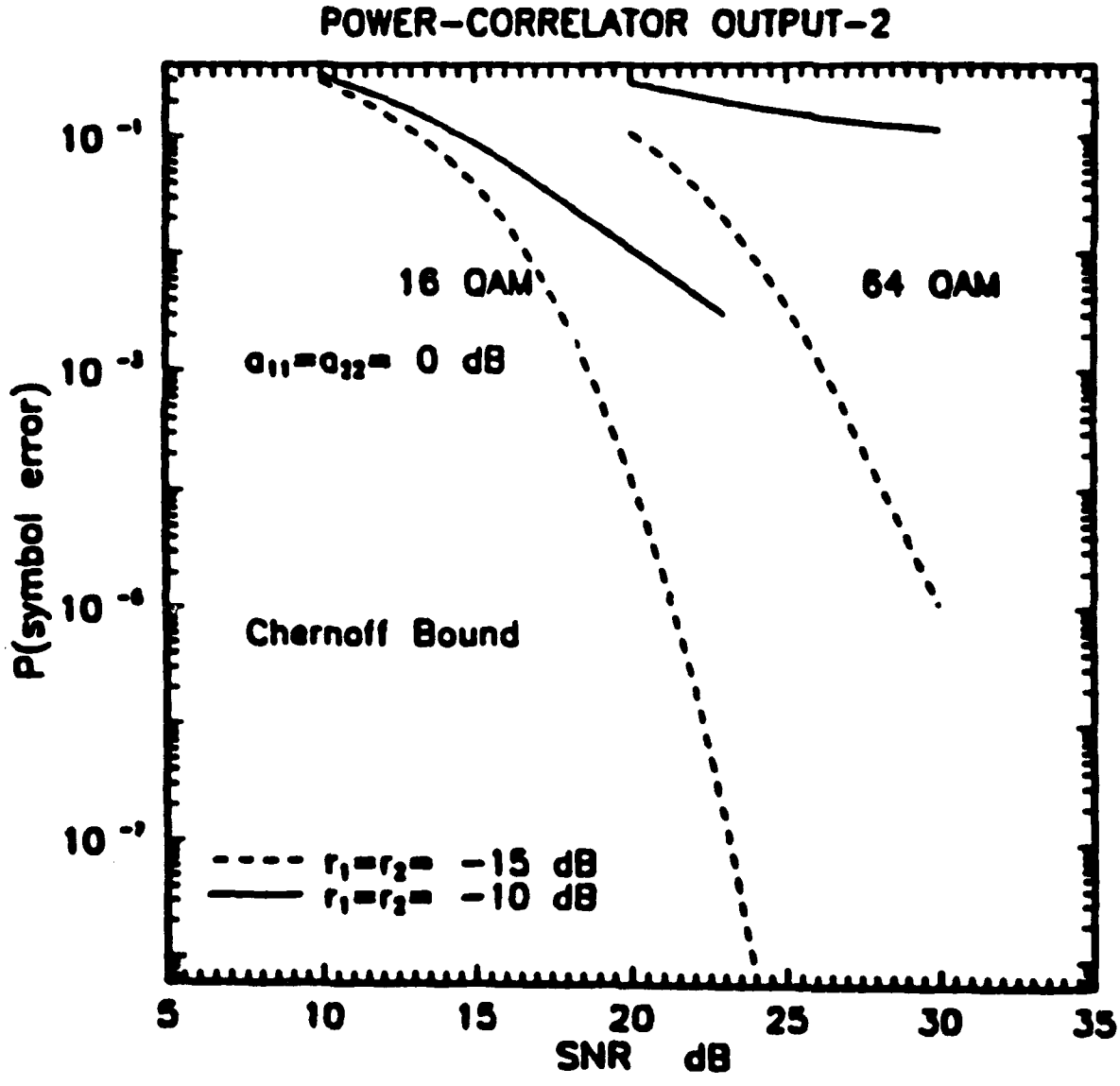


Fig 7. Power-Correlator Cross-Pol Canceler output-2. Chernoff Bound comparison 16 QAM vs. 64 QAM

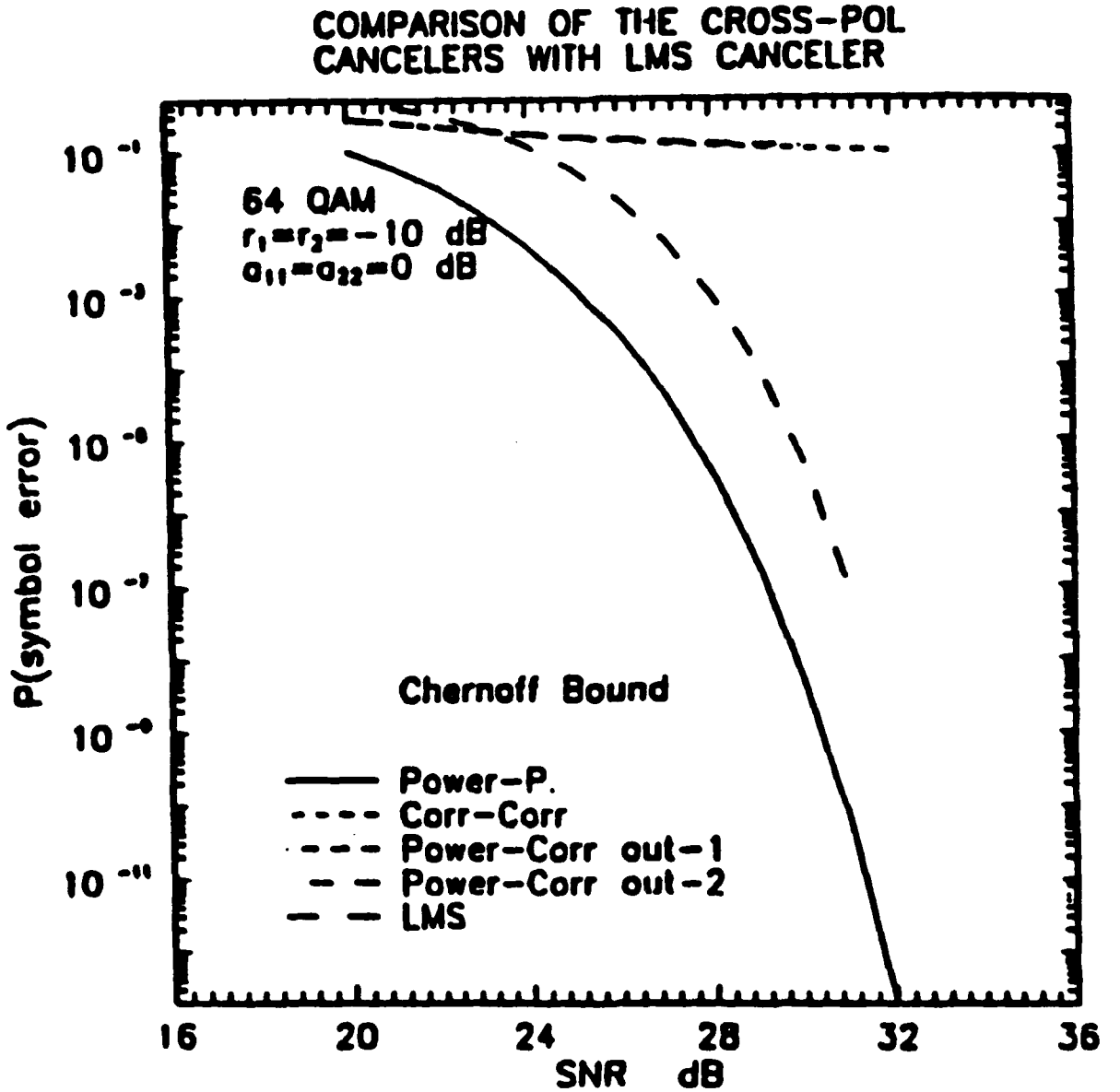


Fig 8. Comparison of the Cross-pol Cancelers with LMS Canceler

Appendix E

Bootstrap: A Fast Blind Adaptive Signal Separator

Abdulkadir Dinç and Yeheskel Bar-Ness

Abstract

In this report, we propose a new fast multidimensional adaptive algorithm for multi signal separation. The method is essentially based on multi-power-inversion schemes [1]. It separates multi uncorrelated signals imposed on each other. The two dimensional version of this adaptive algorithm, named bootstrapped canceler, has been applied to digital communication to mitigate the effect of cross-polarization in dual polarized M-ary QAM signals [3]. It has been shown that in previous appendices that the power-power bootstrapped algorithm performs better than the least mean square (LMS) algorithm in separating the two uncorrelated signals [5]. In this appendix, we propose a multi dimensional nonlinear learning algorithm which is based on minimization of output signal correlations and we also investigate the learning process of this bootstrapped algorithm compared with that of the LMS algorithm for different eigenvalue spreads. It has been found from the computer simulations that the bootstrapped algorithm converges very rapidly with respect to the LMS algorithm. It is important to notice from the computer simulations that the learning process of the bootstrapped algorithm is almost independent of the eigenvalue spread.

1 Introduction

Multi channel interference can be a performance limiting factor in signal processing systems. Removal of these interferences can be accomplished by use of adaptive signal separators. The bootstrapped algorithm, which does not require a training sequence, has been proposed for two dimensional interference cancellation with different adaptive learning schemes [1]. Several adaptive signal separators which require a training sequence have been proposed in the literature applied to communication systems [2].

Our objective in this report is to extend the bootstrapped algorithm to multi dimensional signal separation applications and to investigate its convergence for different eigenvalue spreads.

In section II. the description of the channel model is given. We discuss the adaptive signal separators in section III. In section IV the convergence of LMS and bootstrapped algorithms are compared by a computer simulation. The results of the convergence comparisons are presented in section V. Finally, the conclusion is presented in section VI.

2 Channel Model and Problem Statement

In matrix notation, a discrete time model of a M-dimensional interference channel is given by

$$\mathbf{x}(n) = \mathbf{A}\mathbf{I}(n) + \mathbf{n}(n) \tag{1}$$

where $\mathbf{A}$ is the channel matrix, $\mathbf{I}$ is the information vector assumed to be an independent and identically distributed sequence and $\mathbf{n}$ is a white Gaussian noise sequence,

and $\mathbf{x}$ is the received signal vectors respectively.

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ a_{21} & \cdots & a_{2M} \\ . & \cdots & . \\ a_{M1} & \cdots & a_{MM} \end{bmatrix}, \quad \mathbf{I}(n) = \begin{bmatrix} I_1(n) \\ I_2(n) \\ . \\ I_M(n) \end{bmatrix} \mathbf{n}(n) = \begin{bmatrix} n_1(n) \\ n_2(n) \\ . \\ n_M(n) \end{bmatrix} \tag{2}$$

The channel is assumed to be slowly time varying and non-dispersive. Hence, the channel interference coefficients a_{ij} $i,j = 1,2..M$ are assumed to vary slowly with respect to the signal rate and assumed to be less than one while the diagonal coefficients a_{ii} $i = 1,2..M$ are assumed to be close to one.

Our objective is to find a multi dimensional bootstrapped adaptive algorithm structure that will diagonalize the channel matrix $\mathbf{A}$ (that is to find the inverse of $\mathbf{A}$) without requiring a training sequence and demonstrate its convergence with that of an LMS algorithm.

3 LMS and Bootstrapped Adaptive Signal Separator

3.1 Multidimensional LMS Adaptive Signal Separator

The traditional LMS algorithm which minimizes the error $E\{e_1^2 + e_2^2 + ..e_M^2\}$ at the output of the separator (Fig. 1) can be used as a multi dimensional signal separator and solution of the optimum weights are given by

$$\mathbf{W}_{opt} = \mathbf{R}^{-1}\mathbf{P} \tag{3}$$

and

$$\hat{\mathbf{I}}(n) = \mathbf{W}_{opt}^T \mathbf{x}(n) \tag{4}$$

where $\hat{\mathbf{I}}(n)$ is the estimate of $\mathbf{I}(n)$ information vector.

$$\mathbf{W}_{opt} = \begin{bmatrix} w_{11opt} & \cdots & w_{M1opt} \\ & w_{22opt} & w_{M2opt} \\ \vdots & \cdots & \vdots \\ w_{1Mopt} & . & w_{MMopt} \end{bmatrix}, \quad \mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}, \quad \mathbf{P} = E\{\mathbf{x}(n)\mathbf{I}^T(n)\} \tag{5}$$

where using (1) in (5)

$$\mathbf{R} = \mathbf{A}\mathbf{R}_I\mathbf{A}^T + \mathbf{R}_n \tag{6}$$

where

$$\mathbf{R}_I = \begin{bmatrix} E\{I_1^2\} & 0\cdots & & 0 \\ 0 & E\{I_2^2\} & \cdots & 0 \\ \vdots & & & . \\ 0 & & . & E\{I_M^2\} \end{bmatrix} \quad \mathbf{R}_n = \begin{bmatrix} \sigma_1^2 & & \cdots & 0 \\ 0 & & \cdots & \vdots \\ \vdots & \cdots & & 0 \\ 0 & & . & \sigma_M^2 \end{bmatrix} \tag{7}$$

The recursive weight updating algorithm to search for the optimum weights is given by,

$$\begin{aligned} w_{ij}(n+1) &= w_{ij}(n) - \mu e_i(n)x_j(n), \;\; j = 1,2..M\}_{i=1}^{M} \\ \epsilon_i(n) &= y_i(n) - I_i(n) \end{aligned} \tag{8}$$

where $I_i(n)$ is the reference signal.

3.2 Multidimensional Bootstrapped Signal Separator

From Fig. 2, it can be easily shown that the output of the bootstrapped algorithm is given by;

$$\mathbf{y}(n) = \mathbf{w}_B^{-1}\mathbf{x}(n) = \mathbf{w}_B^{-1}[\mathbf{A}\mathbf{I}(n) + \mathbf{n}(n)] \tag{9}$$

where,

$$\mathbf{w}_B = \begin{bmatrix} 1 & -w_{B12} & \cdots & -w_{B1M} \\ -w_{B21} & 1 & & -w_{B2M} \\ . & . & \cdots & . \\ -w_{BM1} & -w_{BM2} & . & 1 \end{bmatrix} \tag{10}$$

provided that the determinant of $\mathbf{w}_B$ is not zero. we suggest the following bootstrapped recursive algorithm that provides $\mathbf{w}_{Bopt} = \mathbf{A}^{-1}$ in a no noise condition.

Our approach is to minimize the correlations at the outputs. simultaneously. For simplicity, we show the optimal weights for two dimensional signal separation.

Using (1) in (9), one can easily show that,

$$\begin{aligned} y_1(n) &= \frac{1}{\Delta}[AI_1(n) + BI_2(n)] \\ y_2(n) &= \frac{1}{\Delta}[CI_1(n) + DI_2(n)] \end{aligned} \tag{11}$$

where

$$\begin{aligned} A &\triangleq a_{11} + a_{21}w_{B12}, \quad B = a_{12} + a_{22}w_{B12}, \quad C = a_{21} + a_{11}w_{B21} \\ D &= a_{22} + a_{12}w_{B21}, \quad \Delta = 1 - w_{B12}w_{B21} \end{aligned} \tag{12}$$

3.3 Cube Nonlinearity in Controlling Bootstrapped Algorithm

The recursive algorithm to search for the optimum weights is given by,

$$\{w_{Bij}(n+1) = w_{Bij}(n) - \mu f(y_i(n))y_j(n), \;\; j = 1,2..M \;\; i \neq j\}_{i=1}^{M} \tag{13}$$

where $f(y) = y^3$ is an odd nonlinear function.

To simplify the analysis, the optimum weights are found in a no noise environment. By taking the expected value denoted by $E\{(.)\}$ of both sides of (13), and using (11) in (13), we can write,

$$\begin{aligned} E\{y_1^3(n)y_2(n)\} &= E\{\frac{1}{\Delta^4}[I_2^3(n)B^3 + 3I_1(n)I_2^2(n)AB^2 + \\ &\quad 3I_1^2(n)I_2(n)A^2B + I_1^3(n)A^3][I_1(n)C + I_2(n)D] \end{aligned} \tag{14}$$

$$E\{I_i(n)I_j(n)\} = \delta(i-j), \quad E\{I_i(n)\} = 0 \tag{15}$$

Using (15) in (14), and assuming $E\{I_1^2(n)\} = E\{I_2^2(n)\}$, and $E\{I_1^4(n)\} = E\{I_2^4(n)\}$.

$$E\{y_1^3(n)y_2(n)\} = \frac{1}{\Delta^4}[3E\{I_1^2(n)\}E\{I_2^2(n)\}BA(BC + AD) + E\{I_1^4(n)\}(A^3C + B^3D]$$

similarly

$$E\{y_2^3(n)y_1(n)\} = \frac{1}{\Delta^4}[3E\{I_1^2(n)\}E\{I_2^2(n)\}DC(BC + AD) + E\{I_1^4(n)\}(AC^3 + BD^3] \tag{16}$$

The optimum weight vector $\mathbf{w}_{Bopt}$ for the bootstrapped algorithm, is the solution of the nonlinear equations in (16). For i.i.d binary data , $I_i(n) = \pm 1$, from (16), we can write,

$$\begin{aligned} E\{y_1^3(n)y_2(n)\} &= \frac{1}{\Delta^4}[(3B^2 + A^2)AC + (B^2 + 3A^2)BD] \\ \text{similarly} & \\ E\{y_2^3(n)y_1(n)\} &= \frac{1}{\Delta^4}[(3D^2 + C^2)AC + (D^2 + 3C^2)BD] \end{aligned} \tag{17}$$

The desired optimum solution requires $A = B = C = D \neq 0$, simultaneously and $\Delta \neq 0$. Two equilibrium solutions that make the equations in (16) equal to zero are $B = C = 0$ or $A = D = 0$. These optimum solutions are, $\mathbf{w}_{Bopt1} = -[\frac{a_{12}}{a_{22}}, \frac{a_{21}}{a_{11}}]$, and $\mathbf{w}_{Bopt2} = -[\frac{a_{22}}{a_{12}}, \frac{a_{11}}{a_{21}}]$. The required optimum solution which provides a stable equilibrium point is $\mathbf{w}_{Bopt1}$. The detail steady state analysis of the solution of the optimum weights can be found in [1].

3.4 Use of Supervised (reference) Signal in Controlling Bootstrapped Algorithm

This analysis is done in order to compare the learning process of the least mean square (LMS) algorithm which uses a supervised signal with that of the bootstrapped algorithm . The weight updating algorithm is given by;

$$\begin{aligned} \{w_{Bij}(n+1) &= w_{Bij}(n) - \mu e_i(n)y_j(n). \;\; j = 1, 2..M \;\; i \neq j\}_{i=1}^{M} \\ e_i &= y_i(n) - I_i(n) \end{aligned} \tag{18}$$

where $I_i(n)$ is the reference signal.

By taking the expected value of the gradient and using (11) and (12) in (18), we get

$$E\{\epsilon_1(n)y_2(n)\} = \frac{1}{\Delta^2}[E\{I_1^2(n)\}(A - \Delta)C + E\{I_2^2(n)BD]$$

similarly

$$E\{e_2(n)y_1(n)\} = \frac{1}{\Delta^2}[E\{I_1^2(n)\}AC + E\{I_2^2(n)\}B(D-\Delta)] \quad (19)$$

Similarly, for the optimum $\mathbf{w}_{Bopt1}$ solution to exist $A = B = C = D \neq 0$, simultaneously and $\Delta \neq 0$ required. At the required equilibrium point $B = C = 0$, the optimum weight is $\mathbf{w}_{Bopt1} = -[\frac{a_{12}}{a_{22}}, \frac{a_{21}}{a_{11}}]$.

4 Simulation

The inputs to the channel are assumed to be binary $\pm$ 1. The two and three dimensional interference channels are modeled by **A** matrix (2). The outputs from the channel are corrupted by additive Gaussian noise with SNR=40 dB. The received signals are then input to the LMS in one hand and bootstrapped adaptive signal separators in the other hand

The results for 500 Monte Carlo runs are given for a two dimensional interference channel for different **A** matrix coefficients essentially for different eigenvalue spreads. We take signal attenuations $a_{ii} = 1$ and interference constants a_{ij} to be all the same constant value. By setting all the weights initially to zero, and providing a constraint $w_{Bij} < 1$ to search for the optimum w_{Bopt1} in the bootstrapped blind separator. the learning processes of the LMS algorithm and the bootstrapped algorithm are compared for different signal-to-interference ratio, $(SIR = 10log(\frac{a_{i2}^2 E\{I_2^2(n)\}+..+a_{ij}^2 E\{I_j^2(n)\}}{a_{ii}E\{I_i^2(n)\}})\ i \neq j$.

In a two dimensional interference case , the eigenvalue spread $(\frac{\lambda_{max}}{\lambda_{min}})$ of the (2x2) input signal correlation matrix **R** in (5) for $a_{11} = a_{22} = 1$ and interference constants $a_{12} = a_{21}$, $E\{I_1^2(n)\} = E\{I_2^2(n)\} = 1$ and zero noise power are given in Table 1.

SIR	a_{12}	λ_{max}	λ_{min}	λ_{max} / λ_{min}	μ
20	0.1	1.21	0.81	1.494	0.826
13.97	0.2	1.44	0.64	2.25	0.694
10.46	0.3	1.69	0.49	3.449	0.592
7.96	0.4	1.96	0.36	5.444	0.51
6.02	0.5	2.25	0.25	9.0	0.44
4.42	0.6	2.56	0.16	16	0.391
3.01	0.7	2.89	0.09	32.11	0.346
1.94	0.8	3.24	0.04	81.0	0.309
0.91	0.9	3.61	0.01	361	0.277
0	1.0	4.0	0.0	∞	0.255

Table 1.

5 Results

In this section, we present the results of the computer simulations for different SIR and convergence constants μ. We have done the experiment in two parts. In the first part, we compared the learning processes of LMS in (3) and bootstrapped algorithms in (18) for the separation of two dimensional signal sources by using the reference signals. The results of the experiment are given in Figs. 3 and 4, for different $SIR = 10log(\frac{a_{12}^2}{a_{11}^2})$ (from Table 1) and for chosen convergence constants μ The convergence constants for these two algorithms are chosen to be less than their maximum values. In Fig. 5, the learning process of bootstrapped algorithms is depicted for SIR=7.96 dB with different convergence constants.

In the second part of the experiment, we show the comparison of the learning processes of bootstrapped blind separation algorithm in (13) with respect to the same updating algorithm by use of the reference signal in (18) for the three signal separation in Fig. 6.

6 Conclusion

In this report, it is shown experimentally that the learning process of the bootstrapped signal separator in Fig. 1 with fewer of weights is faster than the LMS algorithm in Fig. 2. This can be easily seen from Figs. 3 and 4 especially at high SIR. At the steady state, the residue power with blind separation (using cube nonlinearity) is more than with respect to the use of reference signal in the bootstrapped algorithm. As the recursive weight updating algorithm (13) is a nonlinear process, the search for the global optimum weights is essential for the multi-dimension bootstrapped blind separator.

References

[1] Bar-Ness, Y., "Bootstrapped Adaptive Cross-Pol Interference Cancelling Techniques- Steady State Analysis," *Bell Labs. Report,* Jan. 26, 1982.

[2] Kavehrad, M., "Performance of Cross-Polarized M-ary QAM Signals Over Nondispersive Fading Channels." *AT&T Bell Lab. Tech. J.,* Vol. 63, pp. 499-521, March 1988

[3] Carlin, J. W., Bar-Ness, Y., Gross, S., Steinberger, M. S and Studdiford, W. E., "An IF Cross-Pol Canceller for Microwave Radio," *Journal on Selected Areas in Communication - Advances in Digital Communication by Radio* , vol. SAC-5, No. 3 pp. 502-514, April 1987

[4] Bar-Ness Y., "Bootstrapped Algorithm for Interference Cancellation," *AFCEA - IEEE Technical Conference Tactical Communication* Fort Wayne, Indiana, May 1988

[5] Dinç A. and Bar-Ness Y., "Performance Comparison of LMS, Diagonalizer and Bootstrapped Adaptive Cross-Pol Cancellers For M-ary QAM " *Proceedings of MILCOM 90*, paper 3.7, Monterey, CA Oct.1990

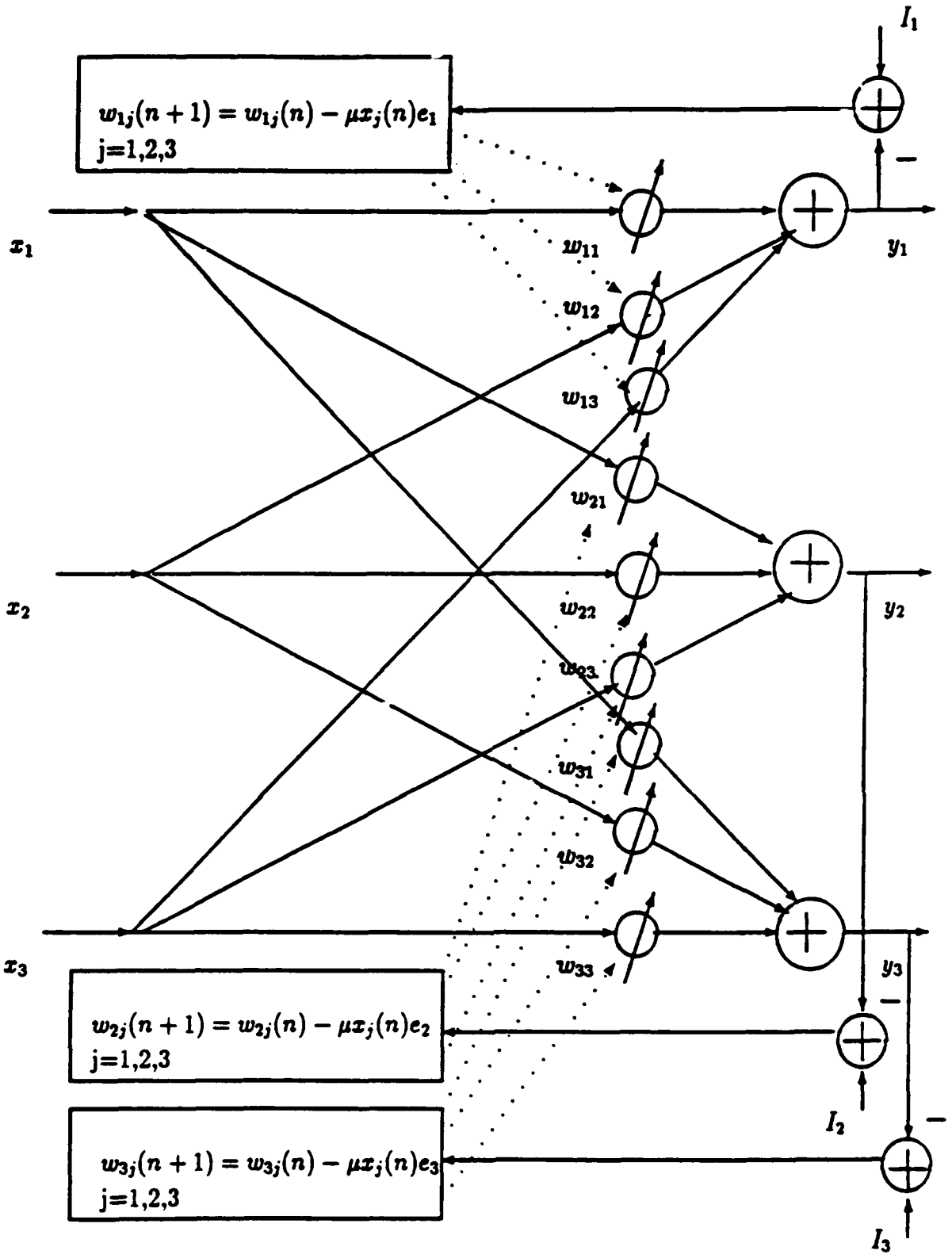


Fig. 1 Multi-Input/Output LMS Signal Separator for 3 input/output

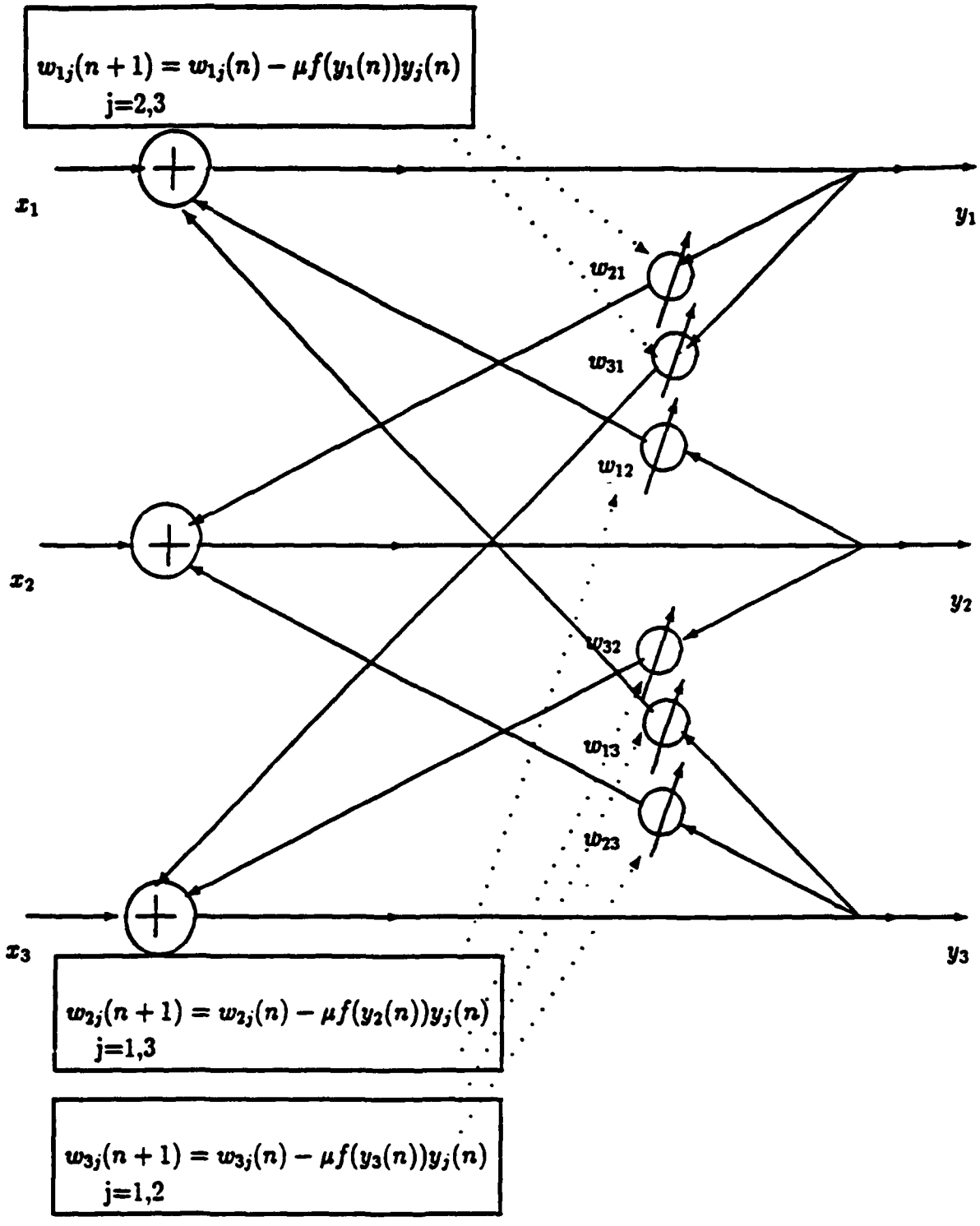


Fig. 2 Multi-Input/Output Backward/Backward Bootstrapped Signal Separator

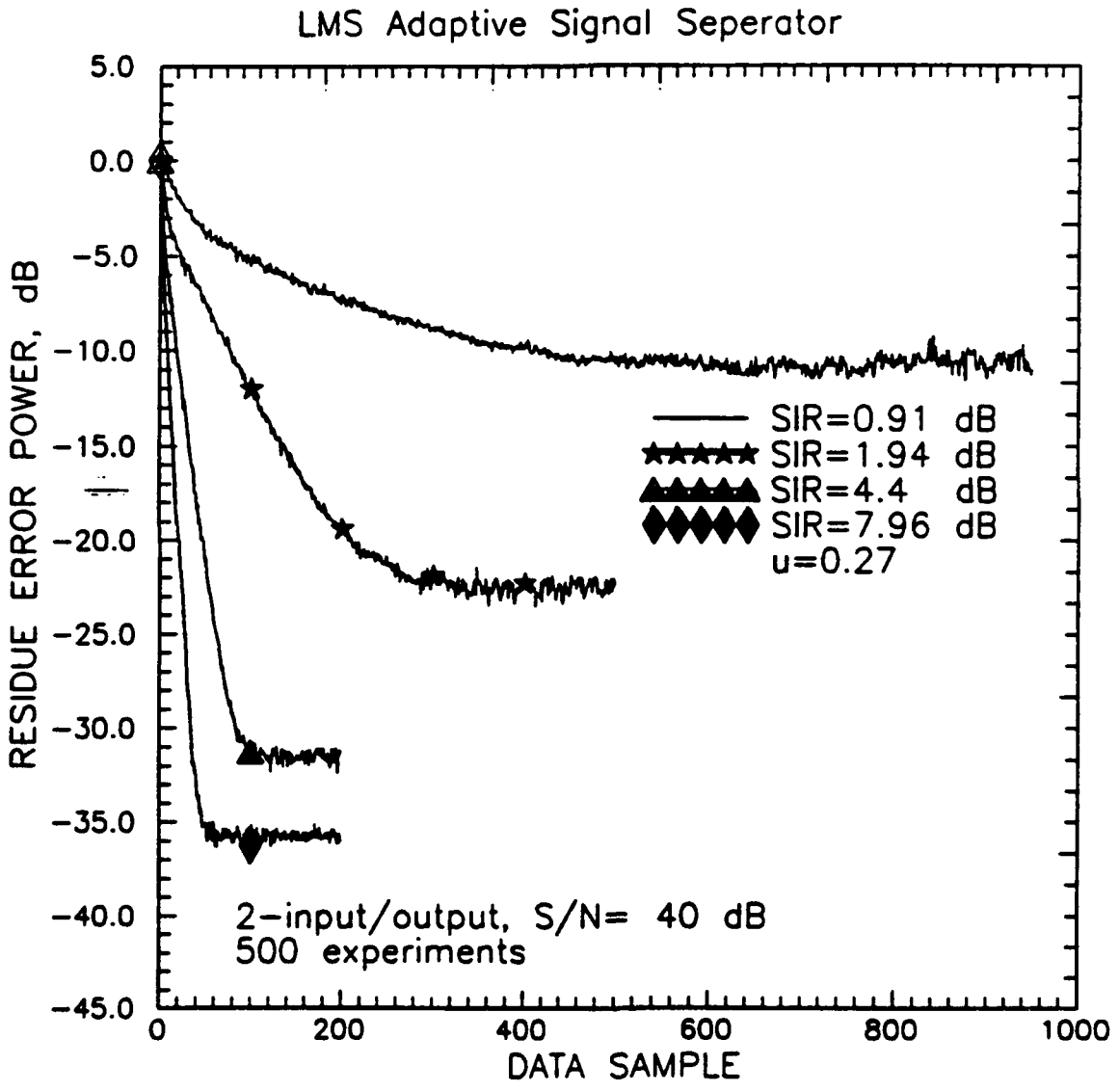


Fig. 3 LMS Signal Separator for Different convergence constants and eigenvalue spreads (cross coupling)

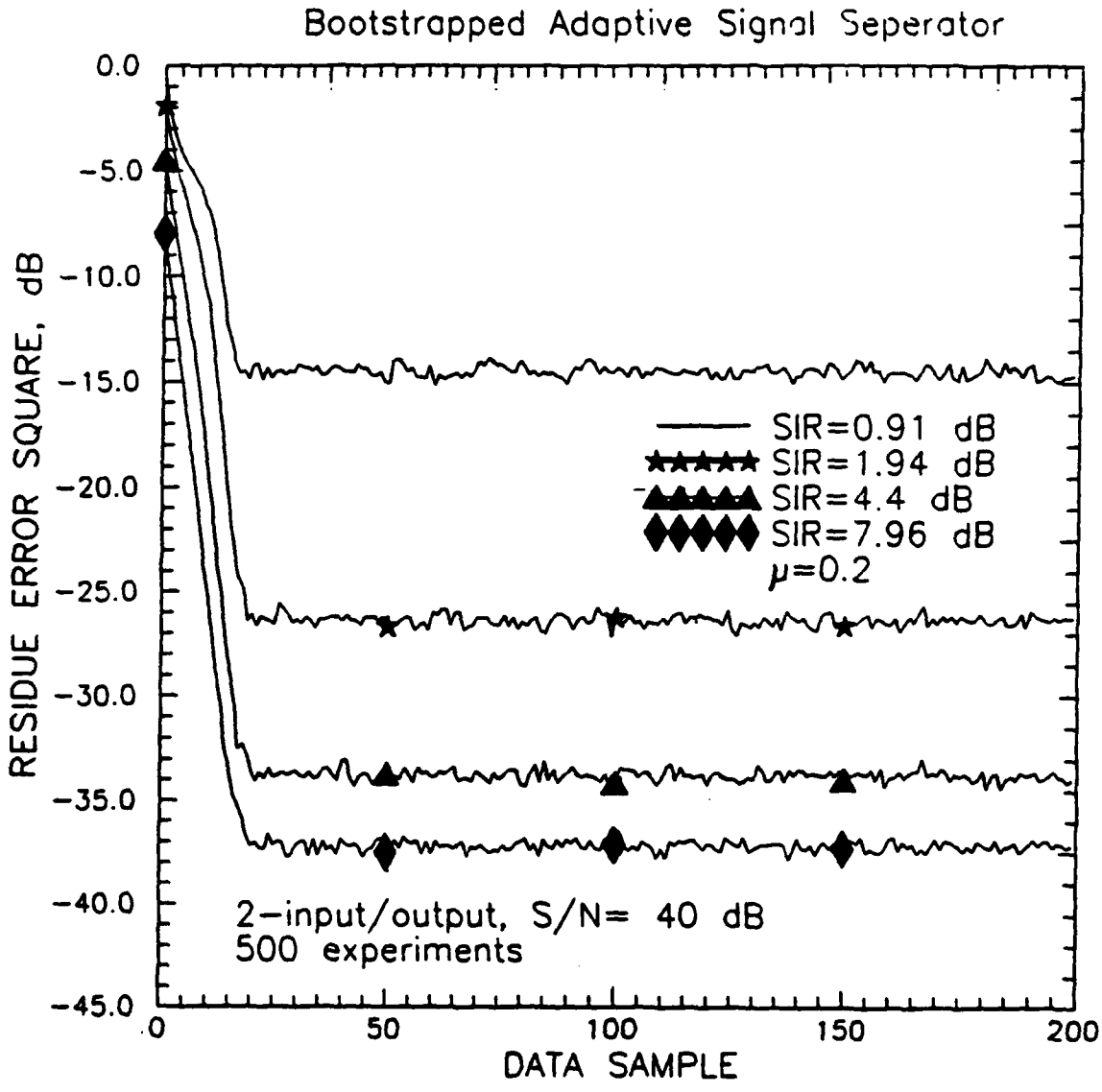


Fig. 4 Bootstrapped Signal Separator with different eigenvalue spreads (cross coupling)

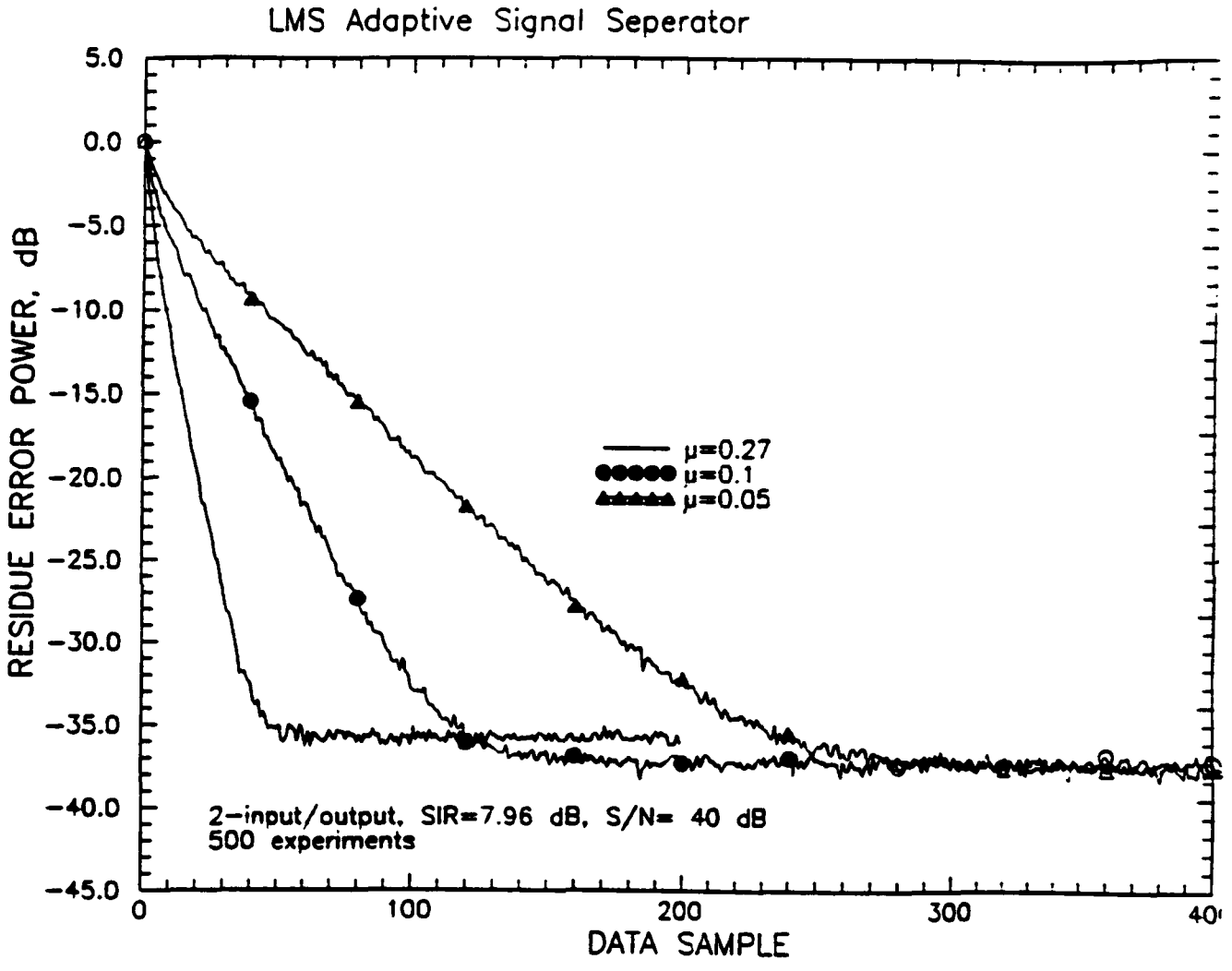


Fig. 5 Bootstrapped Signal Separator for Different convergence constants with fixed (cross coupling)

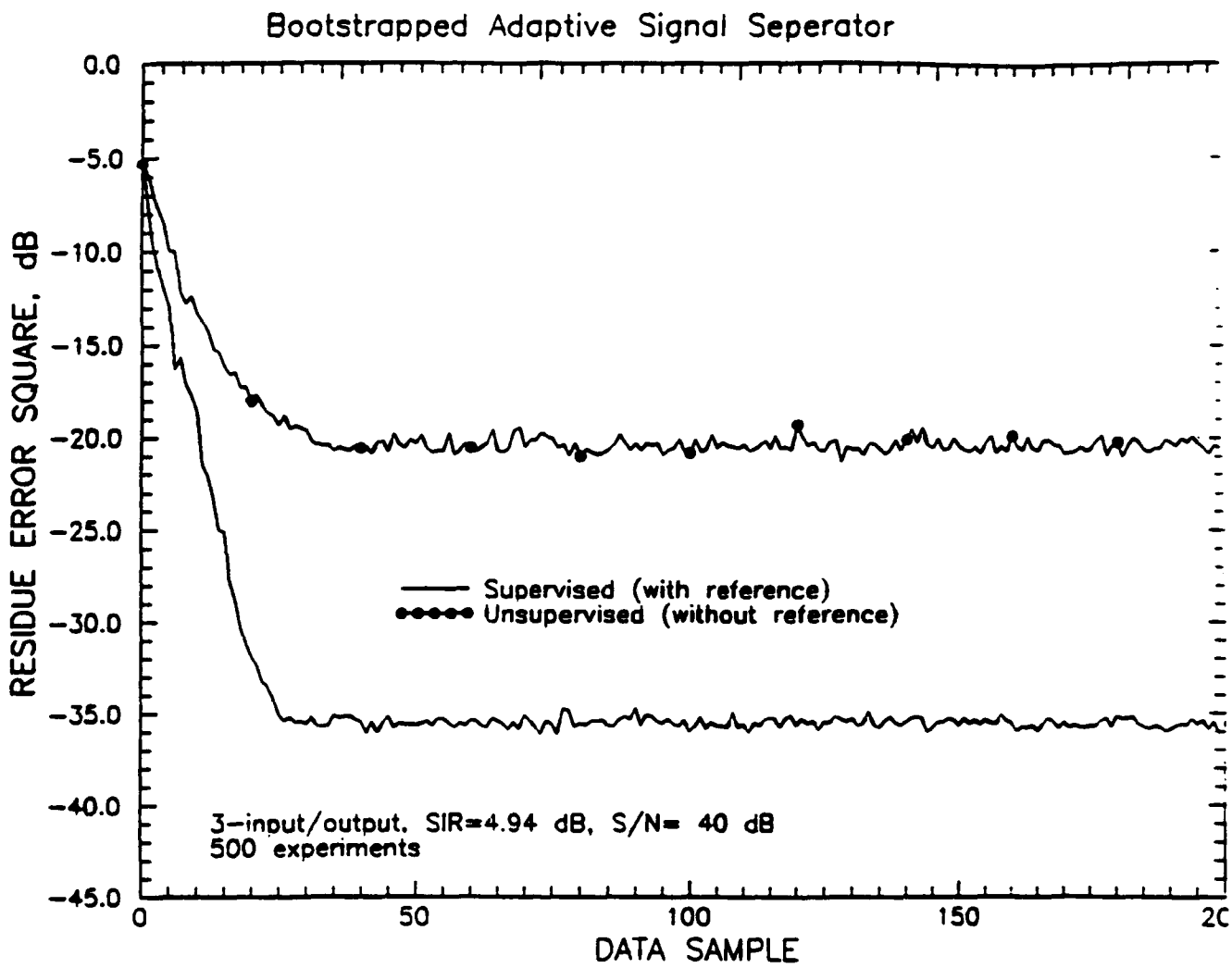


Fig. 6 Bootstrapped Signal Separator for with Supervised and Unsupervised learning

Appendix F

Convergence and Performance Comparison of Three Different Structures of Bootstrap Blind Adaptive Algorithm for Multi-signal Co-Channel Separation

Abdulkadir Dinç and Yeheskel Bar-Ness

Abstract

Multi-signal co-channel interferences can cause major limitations on the performance of communication systems. The multidimensional least mean square (LMS) algorithm can be used to cancel these interferences. However, such an algorithm requires the availability of the reference (supervisory) inputs. In some applications, there is a need for signal separation rather than interference cancellation. That is, the system must contain more than a single output, (each delivers one signal,) as clean as possible from other signals (interferences).

In this appendix, we intend to extend the previously reported three structures of bootstrap blind signal separators to the multi-signal co-channel case, study their performances, their depth of cancellations, speed of convergences and their dependency on eigenvalue spreads. We also present simulation results comparing the performance of these three structures for two and three signals separation under white Gaussian noise environment and for different signal to interference ratios. Particularly, it is shown analytically, as well as experimentally (by simulation), that the use of equalization (automatic gain control (AGC)) at the output of these structures improves the depth of interference cancellation dramatically.

1 Introduction

Although one can design the LMS algorithm to perform signal separation, the complexity of such an approach increases rapidly as the number of signals to be separated increases. Using the bootstrap algorithm, one can perform such separation *without* the need for reference signals in a form of training sequences, decision feedbacks, or other methods. Three different structures proposed in [1], termed power-power, correlator-correlator and power-correlator, have also been reported in the open literature [2]. Its separation capability for two signals was shown and simulation results using the power-power structure were included. Successful application of the power-correlator structure for satellite communication was reported in [3-4]. The use of the power-power structure in cross-polarization cancellation for M-ary QAM was reported in [5]. Error probability performance was estimated and compared to that of LMS and other cancelers. It was proven analytically and demonstrated by simulation and in practical hardware implementation that for the bootstrap algorithms to converge to a state of signal separation. there is a need for the inclusion of nonlinearity, termed signal discriminator.

In addition to being capable of high blind signal separation. it was shown that the power-power structure of the bootstrap algorithm converges faster than the LMS algorithm and is independent of the eigenvalue spread of the input correlation matrix [6].

In this work, we investigate the convergence properties and performance of three structures of the bootstrap blind algorithm with different eigenvalue spreads and extend it to multidimensional signal separation.

In section II a multi-signal co-channel description is given. In section III, the three different structures of the bootstrap blind algorithm are given. In section IV, the convergence and the steady state performance of the three structures of bootstrap blind adaptive algorithms are compared using computer simulations. The conclusion

is presented in section V.

2 Multi-signal Co-Channel Model

In matrix notation, the discrete time model of an M-signal co-channel is given by

$$\mathbf{x}(n) = \mathbf{A}\mathbf{I}(n) + \mathbf{n}(n), \tag{1}$$

where **A** is the channel matrix, **I** is the information vector assumed to be independent and identically distributed sequence, while **n** is a white Gaussian noise and **x** is the received signal vectors, respectively. The channel is assumed to be slowly time varying and non-dispersive. Hence, the channel matrix **A** can be approximated by interference coefficients a_{ij} $i \neq j$, $i.j = 1,2,..M$, assumed to vary slowly with respect to the signal $I_i(n)$ rate and to be less than unity in magnitude. The diagonal coefficients a_{ii} $i = 1,2...M$ are assumed without loss of generality to be unity.

3 Bootstrap Blind Adaptive Signal Separators

3.1 Multidimensional Power-Power Scheme

It can easily be shown from Fig. 1 that the output is given by

$$\mathbf{y}(n) = \mathbf{w}_1^{-1}\mathbf{x}(n) = \mathbf{w}_1^{-1}[\mathbf{A}\mathbf{I}(n) + \mathbf{n}(n)] \tag{2}$$

where,

$$\mathbf{w}_1 = \begin{bmatrix} 1 & -w_{12} & \cdots & -w_{1M} \\ -w_{21} & 1 & & -w_{2M} \\ . & . & \cdots & . \\ -w_{M1} & -w_{M2} & . & 1 \end{bmatrix}. \tag{3}$$

$$\begin{aligned} y_i(n) \;=\; & \frac{1}{\Delta}\Bigg[\Bigg(a_{ii} + \sum_{j\neq i}^{M} w_{ij}a_{ji} + \sum_{j\neq i}^{M} w_{i(M-j+1)}w_{(M-j+1)j}a_{ji} - a_{ii}w_{M(M-1)}w_{(M-1)M}\Bigg)I_i(n) \\ & + \sum_{l\neq i}^{M}\Bigg(a_{il} + \sum_{j\neq i}^{M} w_{ij}a_{jl} + \sum_{j\neq i}^{M} w_{i(M-j+1)}w_{(M-j+1)}a_{jl} - a_{il}w_{M(M-1)}w_{(M-1)M}\Bigg)I_l(n) \\ & + n_{oi}(n)\Bigg] \end{aligned} \tag{4}$$

The first term in (4) represents the desired signal at the i th output, the second term is the residual interference from all other signals, while $n_{oi}(n)$ is the resultant Gaussian noise at the output. Δ is the determinant of the matrix $\mathbf{w}_1$ in (3). Notice that the desired signal is somewhat distorted and needs equalization. Due to the complexity of the general expression, we only write the output for the case of two dimensional signals and no noise environment. Therefore, using (1), (2) and (3), we get for the outputs

$$\begin{aligned} y_1(n) &= \frac{1}{\Delta}[AI_1(n) + BI_2(n)] \\ y_2(n) &= \frac{1}{\Delta}[CI_1(n) + DI_2(n)] \end{aligned} \tag{5}$$

where

$$\begin{aligned} A &\triangleq a_{11} + a_{21}w_{12}, \quad B \triangleq a_{12} + a_{22}w_{12} \\ C &\triangleq a_{21} + a_{11}w_{21}, \quad D \triangleq a_{22} + a_{12}w_{21} \\ \Delta &\triangleq 1 - w_{12}w_{21} \end{aligned} \tag{6}$$

Notice that the desired signal is somewhat distorted and needs equalization.

A Recursive Algorithm to Search for Optimal Weights

A recursive algorithm is used to search for optimal weights that result in signal separation for the three different structures of bootstrapped algorithms. The algorithm simultaneously minimizes the estimates of the output correlations $f[y_i(n)]y_j(n)$ $i,j = 1,2,..M$ $i \neq j$, where $f[.]$ is an odd memoryless nonlinear transformation. It can be shown that such nonlinearity satisfies the need for signal discrimination and is sufficient to make it converge to a state of signal separation. Using steepest descent recursion, we get

$$w_{ij}(n+1) = w_{ij}(n) - \mu f[y_i(n)]y_j(n), \quad i,j = 1,2..M \quad i \neq j. \tag{7}$$

where μ is the stability convergence constant. In order to have convergence in the mean of the weights w_{ij}, we must require $E\{f[y_i(n)]y_j(n)\} = 0, \;\; i,j = 1,2,..M \;\; i \neq j$. In the case of the two dimensional signal, we must have

$$E\{f[y_1(n)]y_2(n)\} = E\{f[y_2(n)]y_1(n)\} = 0. \tag{8}$$

Using (4) in (7), with $f[y] = y^3$, we get

$$\begin{aligned} E\{y_1^3(n)y_2(n)\} &= \frac{1}{\Delta^4}[3E\{I_1^2(n)\}E\{I_2^2(n)\}BA(BC+AD) \\ &\quad +E\{I_1^4(n)\}(A^3C+B^3D], \\ E\{y_2^3(n)y_1(n)\} &= \frac{1}{\Delta^4}[3E\{I_1^2(n)\}E\{I_2^2(n)\}DC(BC+AD) \\ &\quad +E\{I_1^4(n)\}(AC^3+BD^3]. \end{aligned} \tag{9}$$

For the sake of simplicity, we take $E\{I_1^2(n)\} = E\{I_2^2(n)\}$ and $E\{I_1^4(n)\} = E\{I_2^4(n)\}$. Furthermore, for $I_i(n)$, taking values ± 1 with equal probability, (8) becomes

$$\begin{aligned} E\{y_1^3(n)y_2(n)\} &= \frac{1}{\Delta^4}[(3B^2+A^2)AC+(B^2+3A^2)BD], \\ E\{y_2^3(n)y_1(n)\} &= \frac{1}{\Delta^4}[(3D^2+C^2)AC+(D^2+3C^2)BD]. \end{aligned} \tag{10}$$

For the two equations in (9) to equal zero, it is necessary and sufficient to have either $B = C = 0$ or $A = D = 0$. These conditions result in two equilibrium points

$$\mathbf{w}_{opt1} = -[\frac{a_{12}}{a_{22}}, \frac{a_{21}}{a_{11}}] \;, \;\; \mathbf{w}_{opt2} = -[\frac{a_{22}}{a_{12}}, \frac{a_{11}}{a_{21}}]. \tag{11}$$

It can be shown that $\mathbf{w}_{opt1}$ is a stable equilibrium point. Using $\mathbf{w}_{opt1}$ in (4), we get the optimal separator outputs

$$y_{1opt}(n) = a_{11}I_1(n)\ , \quad y_{2opt}(n) = a_{22}I_2(n). \tag{12}$$

With the use of an automatic gain control (AGC), (amplitude compensation) at the output of the separator, we obtain

$$y_{iAGC} \triangleq \frac{y_{iopt}(n)}{a_{ii}} = I_i(n), \quad i = 1, 2 \tag{13}$$

which depicts a total compensated signal separation.

3.2 Multidimensional Correlator-Correlator Scheme

From Fig. 2, we write

$$\mathbf{y}(n) = \mathbf{w}\mathbf{x}(n), \tag{14}$$

where

$$\mathbf{w}_2 = \begin{bmatrix} 1 & \cdots & w_{1j} & \cdots \\ \vdots & & 1 & \cdot \\ w_{i1} & \cdot & \cdot & w_{(M-1),M} \\ \cdot & & w_{M,(M-1)\cdot} & 1 \end{bmatrix}. \tag{15}$$

Using (1) and (14) in (13), the output of this separator is given by

$$y_i(n) = \left[a_{ii} + \sum_{j\neq i}^{M} w_{ij}a_{ji}\right] I_i(n) + \sum_{l\neq i}^{M}\left[a_{il} + \sum_{j\neq i}^{M} w_{ij}a_{jl}\right] I_l(n) + n_{oi}(n) \tag{16}$$

The first term in (15) represents the desired signal at the i th output and the second term is the residual interference from all other signals, while $n_{oi}(n)$ is the resultant Gaussian noise at the output. For the case of the two dimensional signal and no noise environment, we have from (15)

$$y_1(n) = AI_1(n) + BI_2(n)$$
$$y_2(n) = CI_1(n) + DI_2(n) \tag{17}$$

where A, B, C and D as in (6). Using the same argument as in (7), we conclude that for the two dimensional case the optimum weights for the correlator-correlator scheme are also given by (10). Finally, using $\mathbf{w}_{opt1}$ in (16), we get the optimal separator outputs

$$y_{1opt}(n) = a_{11}[1 - \frac{a_{12}a_{21}}{a_{22}a_{11}}]I_1(n)\ , \quad y_{2opt}(n) = a_{22}[1 - \frac{a_{12}a_{21}}{a_{22}a_{11}}]I_2(n). \tag{18}$$

If the channel matrix can be approximated by a constant, then AGC can be implemented at the output leading to

$$y_{iAGC} = \frac{y_{iopt}(n)}{a_{ii}([1 - \frac{a_{12}a_{21}}{a_{22}a_{11}}]} = I_i(n)\ , \quad i = 1, 2. \tag{19}$$

3.3 Multidimensional Power-Correlator Scheme

From Fig. 3, the output is given by

$$\mathbf{y}(n) = \mathbf{w}\mathbf{x}(n), \tag{20}$$

where

$$\mathbf{w} = \begin{bmatrix} (1 + \quad w_{1j}w_{j1} \quad + w_{1M}w_{M1})_{j\neq 1} & \quad w_{1j} & \\ & & \\ w_{i1} & \quad (1 + \quad w_{ij}w_{j1} \quad + w_{iM}w_{M1})_{j>1} & \\ & & \\ & & w_{M,M-1} \quad w_{M-1,M} \\ & & \quad 1 \end{bmatrix}. \tag{21}$$

Similarly, using (19) and (20) together with (1), we obtain

$$y_i(n) = \left[a_{ii}(1+\sum_{j>i}^{M} w_{ij}w_{ji})+\sum_{j\neq i}^{M} w_{ij}a_{ji}\right] I_i(n)+\sum_{l\neq i}^{M}\left[a_{il}(1+\sum_{j>i}^{M} w_{ij}w_{ji})+\sum_{j\neq i}^{M} w_{ij}a_{jl}\right] I_l(n)+n_{oi}(n) \tag{22}$$

It is similar to correlator-correlator i th output in (15) but, notice the added term in the desired response, $I_i(n)$. As is shown later, this term causes reduction in the output equalization requirement without effecting its bandwidth response [7]. In the case of no noise and two dimensional signal separation, we write from (21)

$$y_1(n) = AI_1(n) + BI_2(n)\,, \quad y_2(n) = CI_1(n) + DI_2(n), \tag{23}$$

where

$$\begin{aligned} A &\triangleq (1 + w_{12}w_{21})a_{11} + a_{21}w_{12}\,, \quad B \triangleq (1 + w_{12}w_{21})a_{12} + a_{22}w_{12}\,, \\ C &\triangleq a_{21} + a_{11}w_{21}\,, \quad D \triangleq a_{22} + a_{12}w_{21}. \end{aligned} \tag{24}$$

Similarly, using (22) in (7), we get the equilibrium points

$$\mathbf{w}_{opt1} = -\left[\frac{a_{12}}{a_{22}[1-\frac{a_{12}a_{21}}{a_{22}a_{11}}]}, \frac{a_{21}}{a_{11}}\right], \quad \mathbf{w}_{opt2} = -\left[\frac{a_{22}[1-\frac{a_{12}a_{21}}{a_{22}a_{11}}]}{a_{12}}, \frac{a_{11}}{a_{21}}\right]. \tag{25}$$

As before, we can show that only $\mathbf{w}_{opt1}$ is a stable equilibrium point. Substituting $\mathbf{w}_{opt1}$ of (24) into (22), we get

$$y_{1opt}(n) = a_{11}I_1(n)\,, \quad y_{2opt}(n) = a_{22}[1 - \frac{a_{12}a_{21}}{a_{22}a_{11}}]I_2(n), \tag{26}$$

and with the use of a suitable AGC at the outputs of the canceler, we have

$$y_{1AGC} = \frac{y_{1opt}(n)}{a_{11}} = I_1(n)\,, \quad y_{2AGC} = \frac{y_{2opt}(n)}{a_{22}[1 - \frac{a_{12}a_{21}}{a_{22}a_{11}}]} = I_2(n). \tag{27}$$

Notice that for the power-correlator structure, one of the outputs requires an AGC normalization by a_{ii}, the desired signal response, similar to that required in power-power structure. Particularly, for $a_{ii} = 1$, AGC is not needed. The other output needs an AGC which depends on the interfering signals coupling in a manner similar to that required in the correlator-correlator case.

4 Simulation and Results

In this section, we present the Monte Carlo simulation results for two and three signal separators based on the three different structures of bootstrap and LMS algorithms. The block diagram for the simulation set up is given in Fig. 4. The channel input $I_i(n)$ $i = 1, 2, 3$ are random bipolar independent sequences. Channel parameters were chosen to present different desired signal-to-interference ratios (SIR) at the output. The signal-to-interference ratio is defined by $SIR = 10log[\frac{\sum_{j\neq i}^{M} a_{ij}^2 E\{I_j^2(n)\}}{a_{ii}^2 E\{I_i^2(n)\}}]$. Without loss of generality, we take $a_{ii} = 1$ and cross coupling a_{ij} to be the same for all i and j, $i \neq j$. Different a_{ij} causes the canceler's input correlation matrix to have different eigenvalue spread. White Gaussian noise is added to the output of the channel, with signal-to-noise ratio (SNR) of 40 dB. Such high SNR is used in the simulation to enable better examination of cancellation depth. The blind bootstrap separator in one hand or the LMS separator in the other are used to cancel cross channel interferences. Finally, wherever needed, AGC is added to the output of the separators.

By setting all the weights initially to zero, and providing a constraint $w_{ij} < 1$ to search for the optimal weight w_{ijopt1}, we obtain learning curves from the average results for 500 runs. This is done for two and three dimensional interference chan-

nels. In Fig. 5, we depict such learning curves for power-power, correlator-correlator and power-porrelator bootstrap separators and the LMS separator with a two signal channel. Since the first two are symmetric structures, the results from only one output are shown. For the second, we show both outputs, $y_1(n)$ and $y_2(n)$, as they are shown in Fig. 3. As channel coupling parameters, we used $a_{ij} = 0.8$, and as a convergence constant μ we took values 0.12, 0.08, 0.08 and 0.2 for power-power, correlator-correlator, power-correlator and the LMS signal separators, respectively. These values were chosen to be slightly less than the maximum values allowed for stability. From these results, we note that power-power results in a smaller steady state interference residue than that of the correlator-correlator separator or output 2 of the power-correlator. However, this residue is larger than that of LMS or output 1 of the power-correlator. The correlator-correlator residue equals that of output 2 of the power-correlator. Nevertheless, all bootstrap separators converge faster than the LMS separator.

In Fig. 6, we depict the learning curves of the three separators with channel coupling parameters $a_{ij} = 0.4$ instead. Residues in this case behave similar to Fig. 5 except for the fact that for correlator-correlator, the residue is slightly larger than the residue at output 2 of the power-correlator. The convergence constant μ was taken to be equal to 0.2, which is slightly less than the maximum allowable for the stability for this case of channel parameters. Comparing Fig. 5 to Fig. 6, we notice that when the channel coupling parameter is smaller, the convergence is faster. To examine the effect of AGC on the separator's performance, we show Fig. 7, the learning curves of correlator-correlator with and without AGC. In Fig. 8, we do the same for output 2 of the power-correlator separator. Note that these are the only cases which require equalization via AGC. Fig. 9 and Fig. 10 deal with the case of the three signals channel. In Fig. 9, we compare the learning curves of the correlator-correlator separator with and without AGC to that of the power-power

separator without AGC. We again use $a_{ij} = 0.4$ and μ to be the same as in Figs. 6, 7 and 8. From this curve, we notice that by adding AGC, the correlator-correlator learning curve becomes similar to that of the power-power separator. In comparing the results of this figure to those obtained with the two-signal channel, we clearly notice a higher residue with the former, due to the added interfering signal. In Fig. 10, we compare the learning curves of the power-power separator to the three outputs of the power-correlator separator when AGC is added to the outputs. It is quite clear from the results with the two and three signals' channel that, the speed of convergence is practically the same in both cases.

5 Conclusion

In this appendix, we have extended the previously reported three structures of bootstrap blind adaptive separators to the multi-signal channel case. We suggested a recursive weight updating algorithm for the three structures termed power-power, correlator-correlator and power-correlator. The optimum weights for these separators were found analytically in the absence of noise. The signal separation process was shown via simulation by the outputs learning curve. It was shown that the different bootstrap separators converge to their steady states almost with the same speed for a two or three signals channel. They all converge faster than LMS regardless of the value of channel coupling parameters. The steady state interference residues of the three separators are different, lowest for power-power and highest for correlator-correlator. One output of power-correlator results in residue similar to that of the correlator-correlator. However, adding AGC to the correlator-correlator outputs or one of the outputs of the power-correlator reduces the amount of residue, so that when AGC was added (where it was needed), all separators behaved similarly.

References

[1] Bar-Ness, Y., "Bootstrapped Adaptive Cross-Pol Interference Cancelling Techniques- Steady State Analysis," *Bell Labs. Report,* Jan. 26, 1982.

[2] Bar-Ness, Y., "Bootstrapped Algorithm for Interference Cancellation," *AFCEA - IEEE Technical Conference Tactical Communication* Fort Wayne, Indiana, May 1988.

[3] Carlin, J. W., Bar-Ness, Y.. Gross, S., Steinberger, M. L and Studdiford, W. E., "An IF Cross-Pol Canceller for Microwave Radio," *Journal on Selected Areas in Communication - Advances in Digital Communication by Radio* , vol. SAC-5, No. 3 pp. 502-514, April 1987.

[4] Bar-Ness, Y., Carlin, J. W.. Steinberger, M. L, "Bootstrapping Adaptive Interference Cancellers: Some Practical Limitations," *The Globecom Conf.* Nov. 1982.

[5] Dinç, A. and Bar-Ness, Y.. "Performance Comparison of LMS, Diagonalizer and Bootstrapped Adaptive Cross-Pol Cancellers For M-ary QAM," *Proceedings of MILCOM '90*, paper 3.7, Monterey, CA Oct.1990.

[6] 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.

[7] Bar-Ness, Y., "Bootstrapped Adaptive Cross-Pol Interference Cancelling Techniques- Bandwidth Complexity Tradeoff," Bell Labs. Report, also *to be submitted to IEEE Trans. on Signal Processing*

[8] Dinç, A. and Bar-Ness, Y.. "this appendix to be presented" *to be presented at MILCOM '92*, paper U-141. San Diego, CA Oct. 1992.

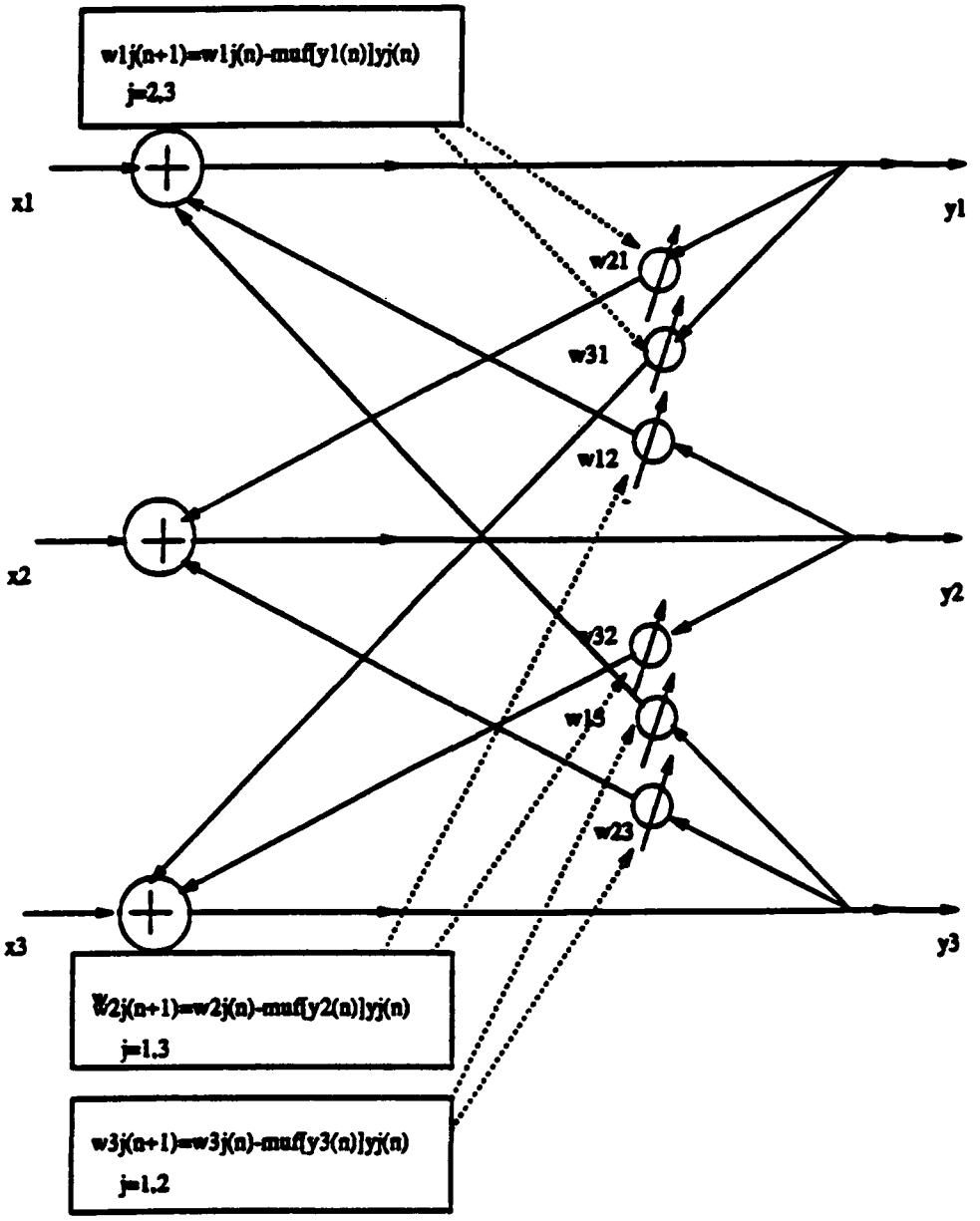


Fig. 1 Multi-Inputs Multi-Outputs Backward/Backward Bootstrapped (Power-Power) Signal Separator

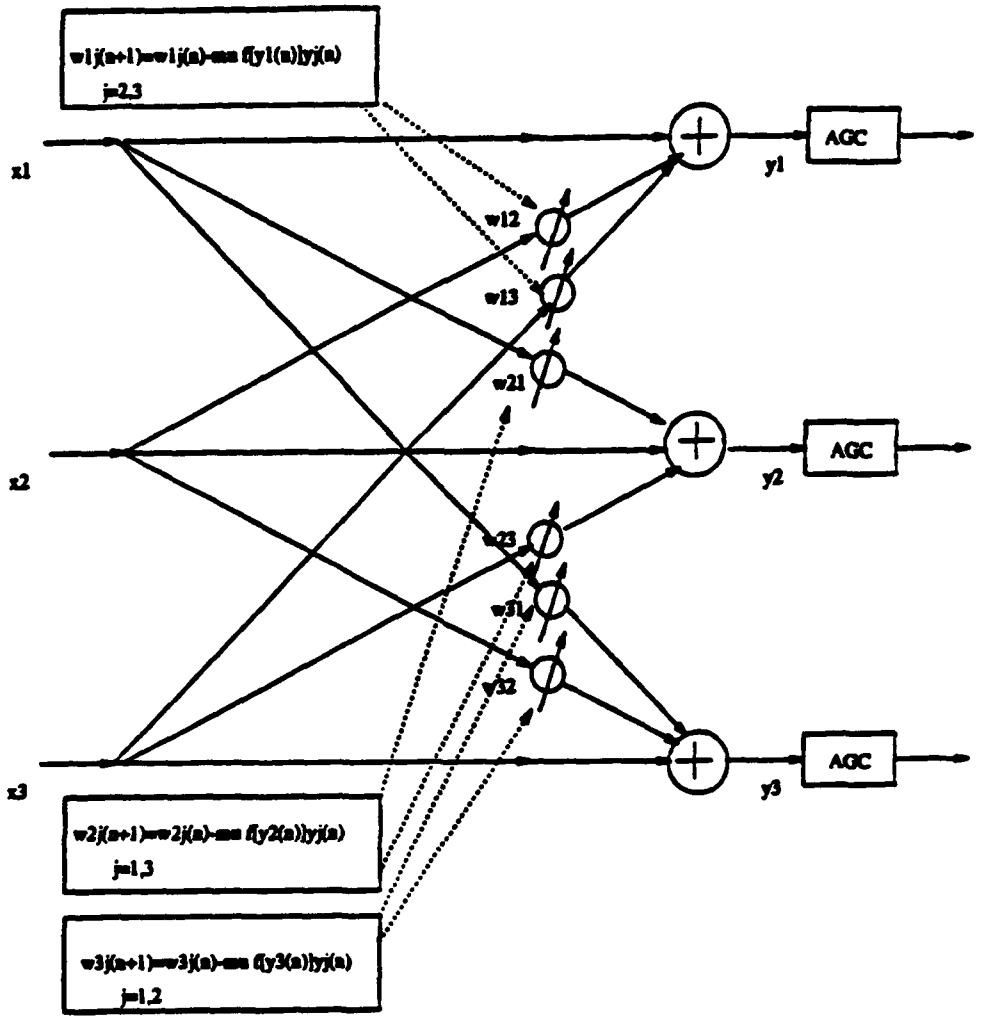


Fig. 2 Multi-Inputs Multi-Outputs Forward/Forward Bootstrapped (Correlator-Correlator) Signal Separator

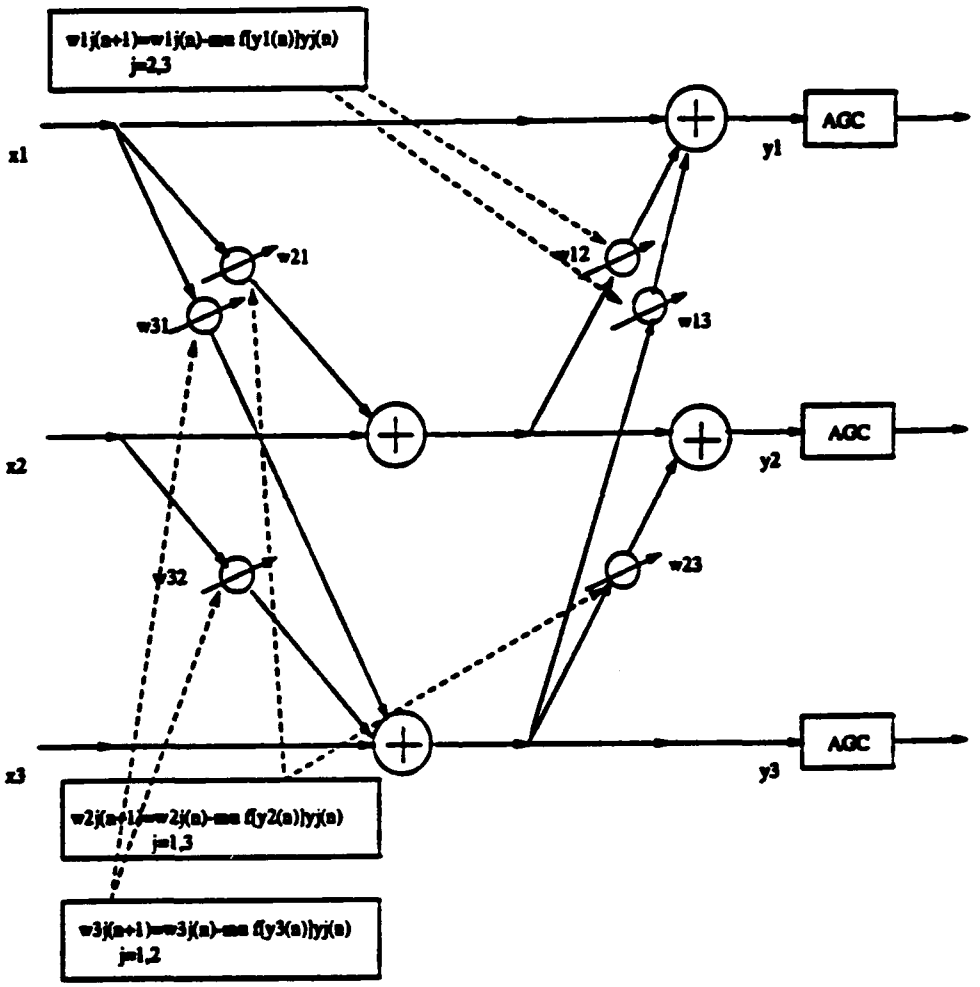


Fig. 3 Multi-Inputs Multi-Outputs Forward/Backward (Power-Correlator) Signal Separator

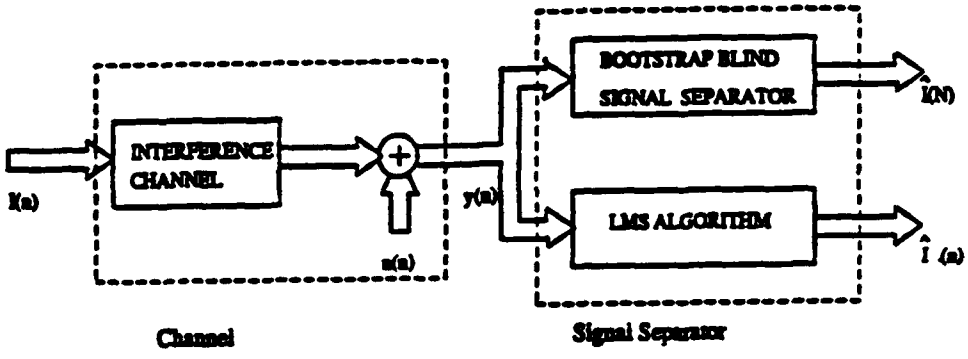


Fig. 4 Multi-Input/Output Interference Channel and Signal Separator

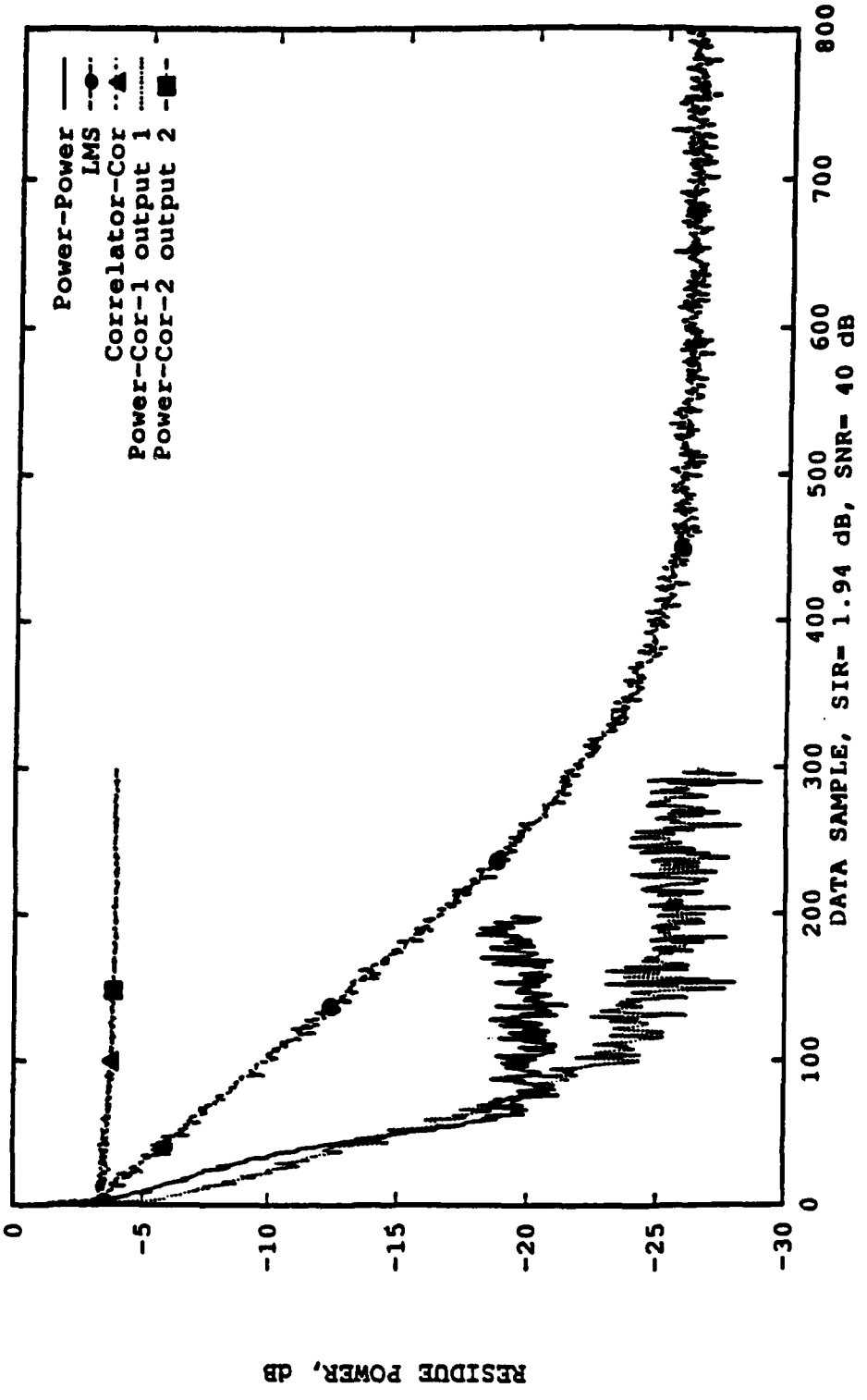


Fig. 5 Convergence Comparison of the Signal Separators for two-inputs/two-outputs

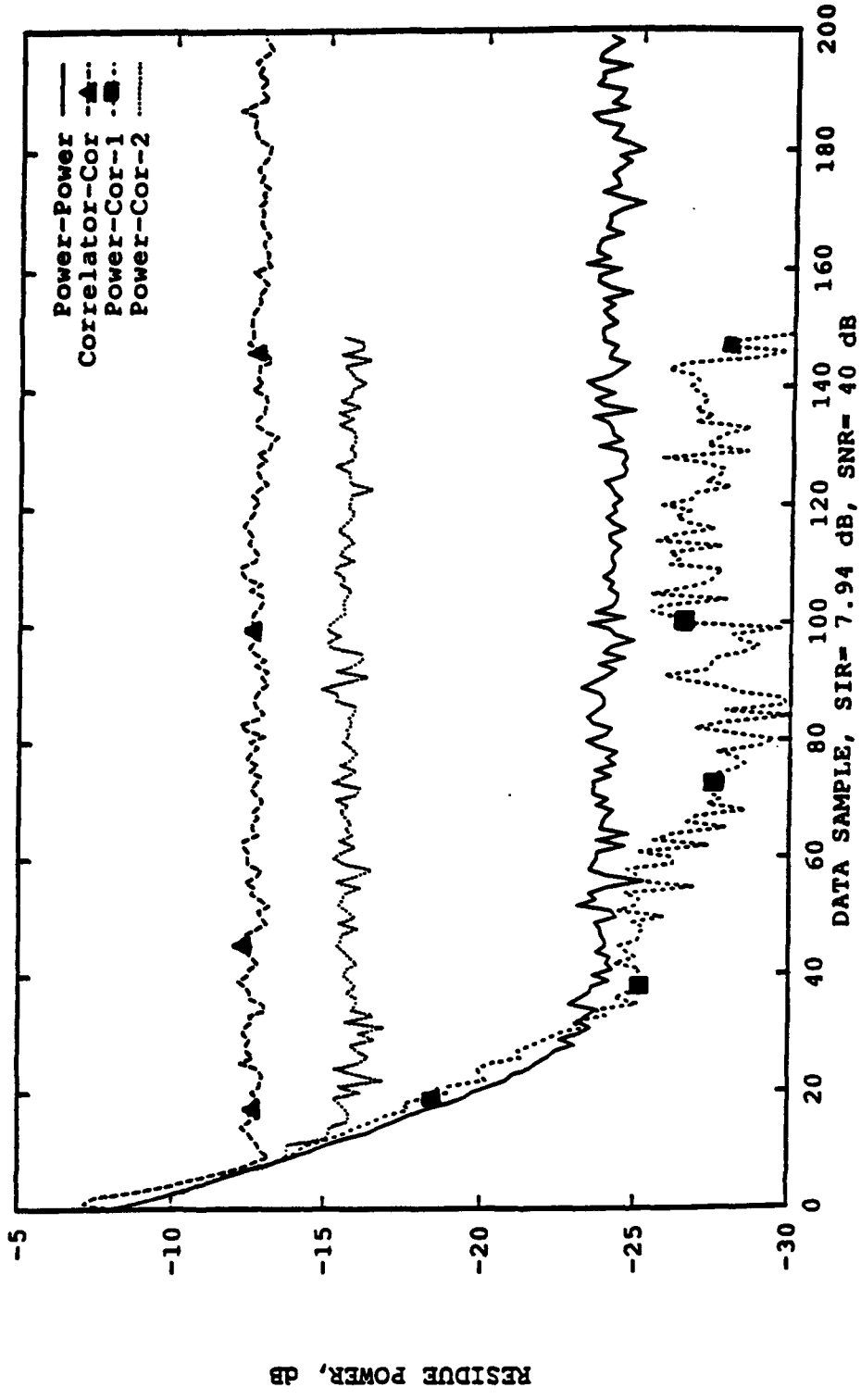


Fig. 6 Convergence Comparison of the Bootstrapped Signal Separators for two-inputs/two-outputs

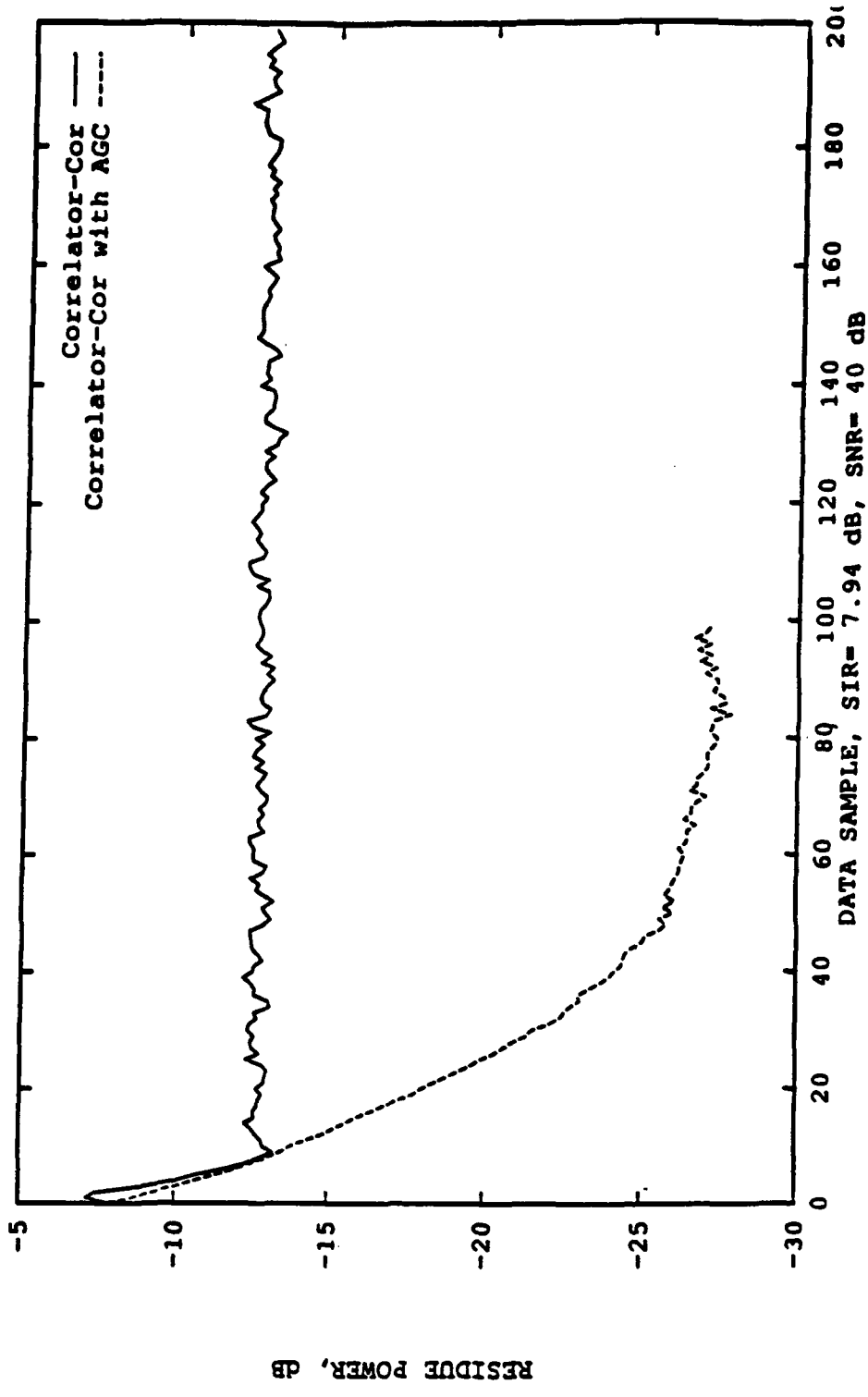


Fig. 7 Convergence Comparison of the Correlator-Correlator with AGC for two-inputs/two-outputs

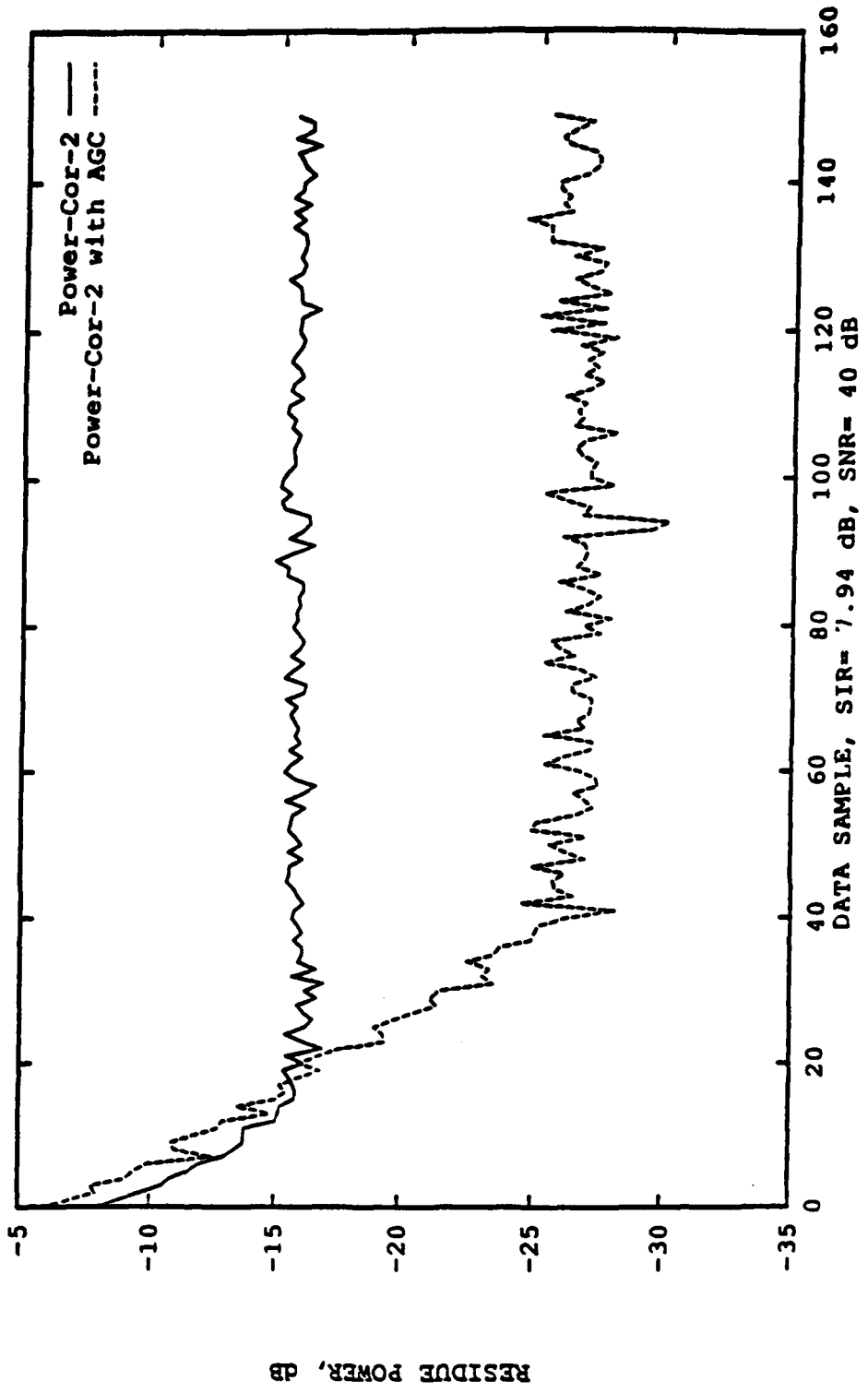


Fig. 8 Convergence Comparison of the Power-Correlator with AGC for two-inputs/two-outputs

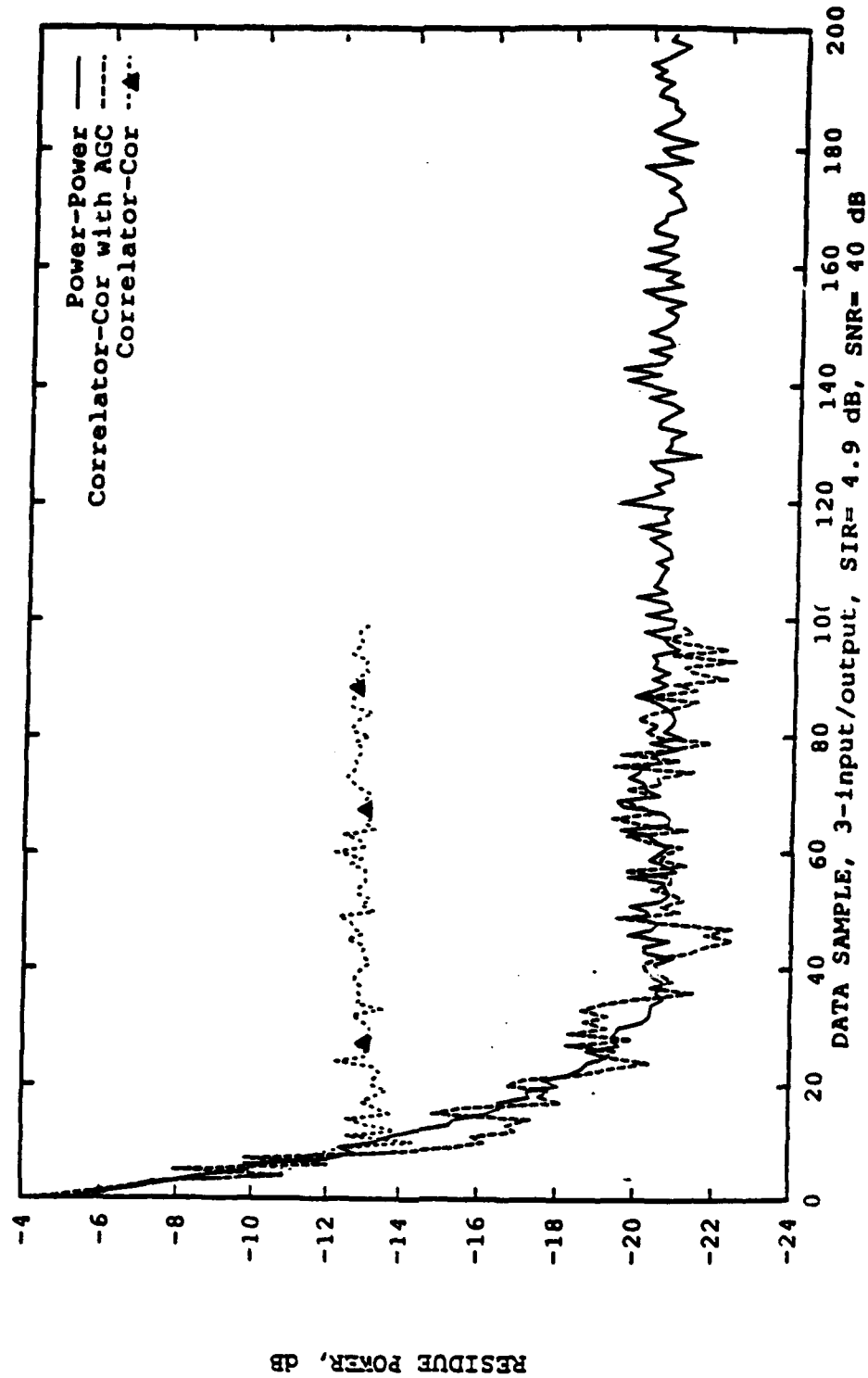


Fig. 9 Convergence Comparison of Power-Power and Correlator-Correlator for Three-inputs Three-outputs

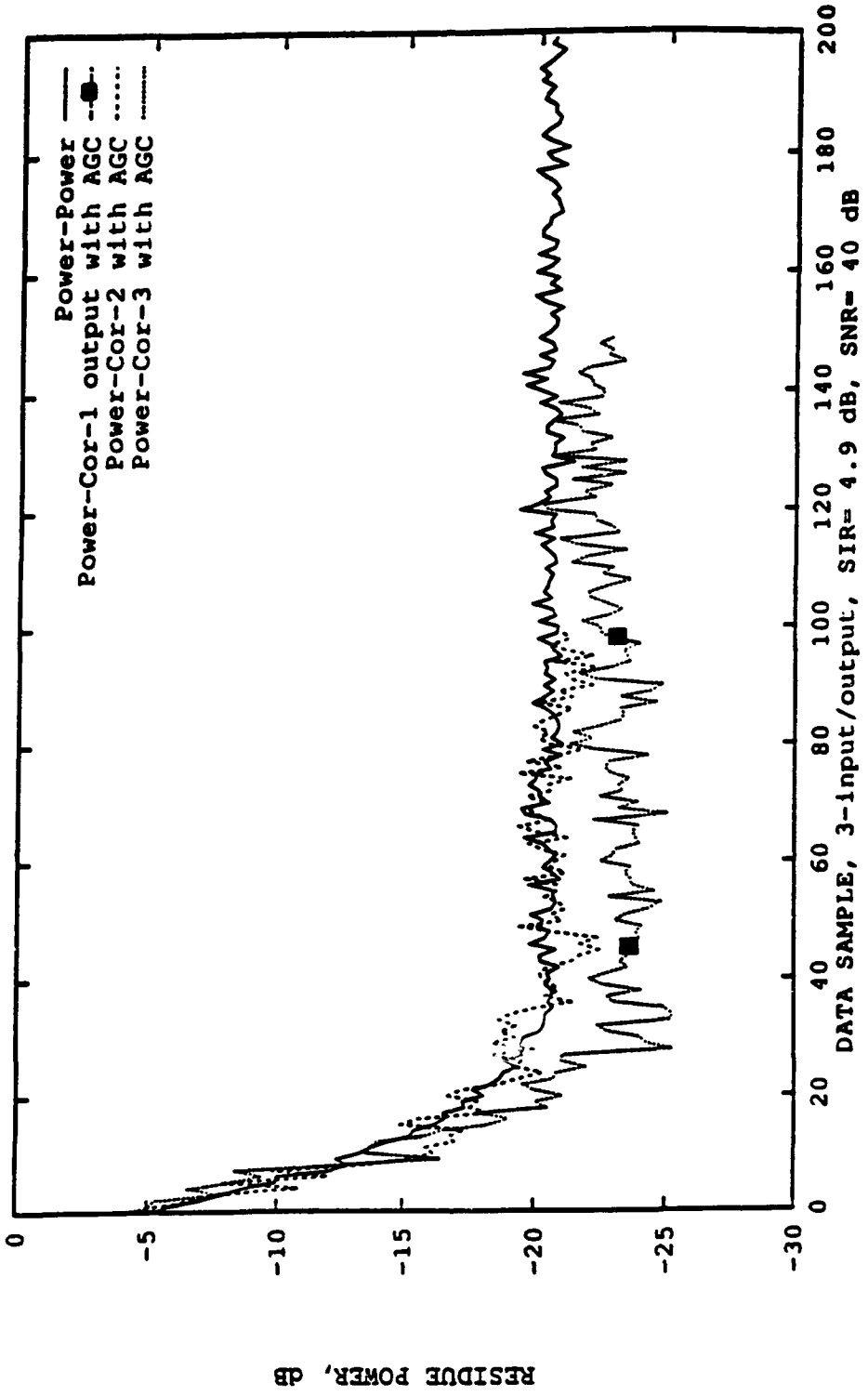


Fig. 10 Convergence Comparison of the Bootstrapped Signal Separators for Three-inputs Three-outputs

Appendix G

Bootstrapped Spatial Separation of Wideband Superimposed Signals

Abstract

Bootstrapped algorithms were developed and used for separation of two signals, when two versions of their weighted sum is given. In this paper we apply the bootstrap principle to the separation of N signals transmitted from point sources at different, unknown locations, when received by an array of M sensors. We present a general structure of the separation scheme which consists of delay elements and summation only. Its input is the M-sensor output signals and its output is the estimates of the N source signals. We show that if the source locations are known, this system provides a least squares estimate of the source signals. If not, it can adaptively converge to the least squares solution, provided that some prior information about the source signals is available. In particular, we present a detailed study of the bootstrapped algorithm for the separation of two sources received by two sensors. A simplified version of the algorithm is presented and the idea of adaptively controlling the unknown delays is discussed. We show that the proposed algorithm is a powerful tool for the decomposition of spatially mixed wideband signals.

1 Introduction and Background

This appendix deals with the scenario in which N point sources are received by M omni-directional sensors. The received signal at the output of each of the M sensors can be modeled by:

$$z_m(t) = \sum_{n=1}^{N} s_n(t - \tau_m(\theta_n)) + e_m(t) \quad m = 1, ..., M \quad ; \quad |t| \leq T/2 \tag{1}$$

where: $s_n(t)$ is the signal radiated from the n-th source; θ_n represents the coordinates (location) of the n-th source: and $e_m(t)$ is the additive noise at the m-th sensor. In the special case of a two dimensional array and far-field sources, θ_n is the source bearing. $\tau_m(\theta_n)$, the travel time of the n-th source from the array origin to the m-th sensor is given by:

$$\tau_m(\theta_n) = \frac{1}{c}[x_m sin\theta_n + y_m cos\theta_n] \tag{2}$$

where c is the propagation velocity of the signal wavefront and (x_m, y_m) are the Cartesian coordinates of the m-th sensor. If the array is an equally spaced linear array (ESLA) then y_m =0 and x_m=$(m-1)d$, m=1,...,M where d is the separation between successive sensors and the plane origin is assumed at the coordinates of the first sensor. Thus, (2) becomes:

$$\tau_m(\theta_n) = (m-1)\frac{d}{c} sin\theta_n \tag{3}$$

In general there are MxN (or (M-1)xN) delays which are a function of the N source locations θ_1,...θ_N . For an ESLA there are only N different delays which carry all the spatial information.

The above model can match many practical applications in different fields. In passive sonar the source signals are wideband, noise-like random processes and the unknown source location vector. $\underline{\theta} = (\theta_1, .., ...\theta_N)^T$, has to be estimated. In the active

case the source signals are basically known, but the aim is the same - estimation of the source locations. However, in a communication system, one is usually interested in the source signals themselves and not in their locations which are sometimes known.

We assume broadband signals, which are more general than the usually assumed narrowband case, where a delay can be regarded as a phase shift. The received sensor data (1) can be represented in the frequency domain:

$$Z_m(\omega_k) = \frac{1}{T}\int_{-T/2}^{T/2} z_m(t)e^{-j\omega_k t}dt \;\; ; m = 1, ..., M \;\; ; k = 1, ...L \tag{4}$$

$Z_m(\omega_k)$ is the Fourier coefficient of the output from the m-th sensor at frequency $(2\pi/T)k$. The processing bandwidth is, therefore, $B = L/T$. At each frequency, the array output is given by the m-dimensional vector:

$$\underline{Z}(\omega_k) = (Z_1(\omega_k), ..., Z_M(\omega_k))^T \tag{5}$$

Following the model of (1), $\underline{Z}(\omega_k)$ can be written as

$$\underline{Z}(\omega_k) = \sum_{n=1}^{N} S_n(\omega_k)\underline{a}(\omega_k, \theta_n) + \underline{E}(\omega_k) = A(\omega_k\underline{\theta})\underline{S}(\omega_k) + \underline{E}(\omega_k) \tag{6}$$

where $\underline{S}(\omega_k) = (S_1(\omega_k), ..., S_N(\omega_k))^T$ and $\underline{E}(\omega_k = (E_1(\omega_k), ..., E_M(\omega_k))^T$. $S_n(\omega_k)$ and $E_m(\omega_k)$ are the Fourier coefficients of the n-th source signal and the noise in the m-th sensor, respectively. Also,

$$A(\omega_k, \underline{\theta}) = [\underline{a}_1(\omega_k) : ... : \underline{a}_N(\omega_k)] \tag{7}$$

where

$$\underline{a}(\omega_k, \theta_n) = \underline{a}_n(\omega_k) = (e^{j\omega_k\tau_1(\theta_n)}, ..., e^{j\omega_k\tau_M(\theta_n)})^T \tag{8}$$

If $M \geq N$ the least-squares estimate of the frequency domain vector of source signals, $S(\omega)$, given the data vector $\underline{Z}(\omega)$, is [1]:

$$\underline{\hat{S}}(\omega_k) = [A^*(\omega_k, \underline{\theta})A(\omega_k, \underline{\theta})]^{-1}A^*(\omega_k, \underline{\theta})\underline{Z}(\omega_k) \tag{9}$$

For the special case of $M = N = 2$ it easily can be verified that

$$[A^*(\omega_k, \underline{\theta})A(\omega_k, \underline{\theta})]^{-1}A^*(\omega_k, \underline{\theta})\underline{Z}(\omega_k) \;=\; \frac{1}{sin^2\omega\Delta}\begin{pmatrix} e^{j\omega_k D_1} - cos\omega_k\Delta e^{j\omega_k D_2} & e^{-j\omega_k D_1} - cos\omega_k\Delta e^{-j\omega_k D_2} \\ e^{j\omega_k D_2} - cos\omega_k\Delta e^{j\omega_k D_1} & e^{-j\omega_k D_2} - cos\omega_k\Delta e^{-j\omega_k D_1} \end{pmatrix} \tag{10}$$

where $\Delta = D_1 - D_2$. In (8) we also assume that the array origin reference is the mid location between the elements, so that $\tau_1(\theta_1) = -\tau_2(\theta_1) = D_1$, $\tau_1(\theta_2) = -\tau_2(\theta_2) = D_2$. We see that, even if D_1 and D_2 are known, the implementation of the least squares solution of (9) requires filtering the array outputs $z_i(t)$ using filters having transfer functions of the form: $cos\omega\Delta$ or $1/sin^2\omega\Delta$, as well as pure delays (see Fig. 1). The implementation of the trigonometric filters is difficult, especially when δ is unknown and is to be estimated adaptively. One possible approach to deal with this implementation problem is to *approximate* the trigonometric filters by FIR (or IIR) filters. In the sequel, we show that applying the bootstrap principle to this problem results in an *exact* implementation of (9) which uses only delay elements and summations, in a feedback configuration.

2 The Bootstrapped Algorithm

Bootstrapped systems are multi-input multi-output feed-back systems in which each output "helps" to improve the other outputs. The idea is to assume that a system output is indeed the desired response and to use it, via feedback, to get other outputs. This approach has been successfully applied in satellite communication to improve separation of cross-pol signals [2]. In our problem, given any N-1 source signals one can get a good estimate of the remaining signal, say $s_n(t)$ from any of the sensor outputs using:

$$\hat{s}_{nm}(t) = \hat{s}_n(t - \tau_m(\theta_n)) = z_m(t) - \sum_{(i=1)_{i\neq n}}^{N} s_i(t - \tau_m(\theta_i))) \tag{11}$$

If no noise exists, then the left hand side of (11) is indeed a delayed version of $s_n(t)$. However, since noise is never zero, averaging over the estimated version of a certain signal from all sensor outputs will improve SNR output. Therefore, we have:

$$\hat{s}_n(t) = \frac{1}{M} \sum_{m=1}^{M} \hat{s}_{nm}(t + \tau_m(\theta_i)) \tag{12}$$

The outputs of the proposed algorithm are N signals, $y_1(t), ..., y_N(t)$ which are desired to be the best possible estimates of the N source signals $s_1(t), ..., s_N(t)$. Following the bootstrapped approach, we replace the known source signals in (11) by their estimates $\{y_i\}$. Therefore, the proposed scheme is described by the NxM equations:

$$y_{nm}(t) = y_n(t - \tau_m(\theta_n)) = z_m(t) - \sum_{(i=1)_{i\neq n}}^{N} y_i(t - \tau_m(\theta_i)) \;\; ; m = 1, ..., M : n = 1...., N \tag{13}$$

$$y_n(t) = \frac{1}{M} \sum_{m=1}^{M} y_{nm}(t + \tau_m(\theta_i)) \tag{14}$$

By transforming (13) to the frequency domain we get

$$Y_n(\omega) = \frac{1}{M}\sum_{m=1}^{M} e^{j\omega\tau_m(\theta_n)}[Z_m(\omega) - \sum_{(i=1)_{i\neq n}}^{N} Y_i(\omega)e^{-j\omega\tau_i(\theta_n)}] \;\; ; m = 1,...,M; n = 1,...,N \tag{15}$$

In a matrix form this equation is equivalent to

$$A^*(\omega,\underline{\theta})A(\omega,\underline{\theta})\underline{Y}(\omega) = A^*(\omega,\underline{\theta})\underline{Z}(\omega) \tag{16}$$

where $A(\omega,\underline{\theta})$ is given by (7) and (8). Therefore, the frequency domain representation of the output vector $\underline{y}(t) = (y_1(t),...,y_N(t))^T$ is exactly the same as $\hat{S}(\omega_k)$ of (9). i.e., the system described by (12) is a realization of the least squares estimator of the source signals. In Fig. 2 we present a block diagram of this system for the special case of $N = M = 2$.

3 Bootstrapped Separation of Sources at Unknown Locations

If the sources location is unknown, then $\underline{\theta} = (\theta_1, ..., \theta_N)^T$ is unknown, and $\tau_m(\theta_i)$ in (9) is replaced by its estimate, $\hat{\tau}_{mi} = \tau_m(\hat{\theta}_i)$, $m = 1, ..., M$, $i = 1, ..., N$. In that case, the frequency domain representation of the output vector is:

$$\underline{Y}(\omega_k) = [A^*(\omega_k, \underline{\hat{\theta}}) A(\omega_k, \underline{\hat{\theta}})]^{-1} A^*(\omega_k, \underline{\hat{\theta}}) \underline{Z}(\omega_k) = [\hat{A}^* \hat{A}]^{-1} \hat{A}^* \underline{Z} \qquad (17)$$

This is no-longer the least square estimate of $\underline{S}(\omega_k)$, but only an approximation. In the sequel, we propose an adaptive algorithm by which an estimate of the delays D_i will be found. We show that, in case of no additive noise, the algorithm converges to the least squares solution of (9). We demonstrate our results for the special case where $N = M = 2$. However, generalization of the algorithm for any $M \geq N \geq 2$ is straight forward. Consider the system of Fig. 3. It can be shown that for $\tau_1(\theta_1) = -\tau_2(\theta_1) = D_1$, $\tau_1(\theta_2) = -\tau_2(\theta_2) = D_2$, the 2x2 transfer function matrix, $H(\omega)$, which relates the two outputs, $y_1(t)$ and $y_2(t)$ to the two inputs $z_1(t)$ and $z_2(t)$ is exactly the same as those of the system of Figs. 1 and 2. That is, the system of Fig. 3 is another, alternative implementation of a least squares separator with $H(\omega)$ given by (10). If initially $\tau_1 \neq D_1$ and/or $\tau_2 \neq D_2$, then we intend to adapt τ_1 and τ_2 so that in the steady state they reach these optimal values. We notice that the scheme of Fig. 3 is similar to the bootstrapped "power-power" separator of [2-7] which is applied to the separate weighted sum of two uncorrelated signals. There, the delays τ_1 and τ_2 are replaced by complex weights, say W_1 and W_2. In the frequency domain our unknown delays are represented by $e^{j\omega\tau_1}$ and $e^{-j\omega\tau_2}$ and the input signals are weighted sums of the uncorrelated frequency domain signals $S_1(\omega)$ and $S_2(\omega)$. Therefore, the frequency domain representation of our problem is equivalent to the time domain separation problem of [2-10] and a similar approach can be considered. It was shown there that without noise ($n_1 = n_2 = 0$), the power of the two out-

put signals is minimal if and only if the controlled weights are equal to the unknown model parameters. It was also shown that for this configuration, the optimization criterion of minimum power is equivalent to the criterion of zero correlation between the two outputs [3-4]. By analogy, in our problem the equivalent optimization criterion should be minimum power (or zero correlation) in the frequency domain. However, since power and correlation are preserved when transforming from time to frequency and vice versa (Parseval), this criterion can be applied to our problem either in the frequency domain. or in the time domain. Inspirated by a possible hardware implementation using voltage controlled delay lines, we prefer the time domain approach. That is, we suggest to control the unknown delays τ_1 and τ_2 by an adaptive algorithm which seeks for the minimum power of both outputs, simultaneously (or, for the minimum power of their cross correlation signal). Notice, however, that this control procedure cannot be employed unless some information that distinguishes the signals to be separated is available. Mathematically, it can be shown that all possible optimization criteria (zero correlation [13], minimum power) yield the same, or linearly dependent. control equations. Notice, however, that to control both τ_1 and τ_2 one needs two independent equations. This difficulty can also be predicted by considering the separation problem as a multi-input multi-output identification problem, where it is well known that weighted sums cannot be separated if nothing is known about their components. In our application, it is also well known that the resolution capacity of an array of M sensors is bounded by the number of sensors (i.e., $M \geq N$) if absolutely nothing is known about the signals. Conversely, if prior information is available, it can be used to discriminate between the two signals in the control loops. Thus, as suggested in [3,4], such a dependency problem can be handled by introducing a "discriminator" which uses the distinguishing information to emphasize $s_1(t)$ in one of the control loops and another one which emphasizes $s_2(t)$ in the other loop. This

[13] Lately the de-correlation approach is also used in [11]

procedure yields two independent control equations, which guarantees convergence of the adaptive algorithm to the desired solution $\tau_1 = D_1$ and $\tau_2 = D_2$. The possible control loops are depicted in Fig. 4.

4 Conclusions

For separation of signals radiated from point sources, we propose the system of Fig. 3. where the delays are controlled by any of the algorithms of Fig. 4 (there are 4 different combinations, for two unknown delays). We have shown that if the delays are adapted to the unknown model parameters exactly, then the outputs of the proposed bootstrapped system are the least-squares estimates of the source signals. In a further, on-going study we investigate the adaptive algorithm with more details. we suggest alternative implementations of the separation configuration and we study the effect of the additive noise on the performance of the proposed separator.

References

[1] J.M. Mendel, "Lessons in Digital Estimation Theory", *Prentice Hall, 1987*

[2] Y. Bar-Ness, J. Rokach: "Cross-Coupled Bootstrapped Interference Canceler," *Proc. of the 1981 SP-S International Symposium, pp. 292-295.*

[3] Y. Bar-Ness, J.W. Carlin, M.L. Steinberger: "Bootstrapping Adaptive Cross-Pol Canceller for Satellite Communication," *Proc. of the Int. Conf. on Comm., paper 4F.S, June 1982.*

[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] J. Carlin, Y. Bar-Ness et. al.: "An IF Cross-Pol Canceler for Microwave radio," *IEEE Trans. SAC-S , pp. 502-514. April 1987.*

[6] A. Dinc, Y. Bar-Ness "Performance Comparison of LMS, Diagonalizer and Bootstrapped Adaptive Cross-Pol Canceler Over Non-Dispersive Channel," *MILCOM '1990, paper 3.7*

[7] A. Dinc, Y. Bar-Ness: "Error Probability of Bootstrapped Blind Adaptive Cross-Pol Cancelers for Many GAM Signals Over Non-Dispersive Fading Channels," *ICASSP-92, March 1992.*

[8] C. Jutten, J. Herault: "Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetic Architecture," *Signal Processing, Vol. 24. No. 1, pp.1-10, July 1991.*

[9] P. Common, C. Jutten, J. Herault: "Blind Separation of Sources, Part II: Problem Statement," *Signal Processing. Vol. 24, No. 1, pp. 11-20, July 1991.*

[10] E. Sorouchyati, C. Jutten, J. Herault: "Blind Separation of Sources, Part III: Stability Analysis," *Signal Processing. Vol. 24, No. 1, pp.21-29, July 1991.*

[11] E. Weinstein, M. Feder, O.V. Oppenhiem: "Multi-Channel Signal Separation Based on Decorrelation," submitted for publication in the *IEEE Trans. on Signal Processing.*

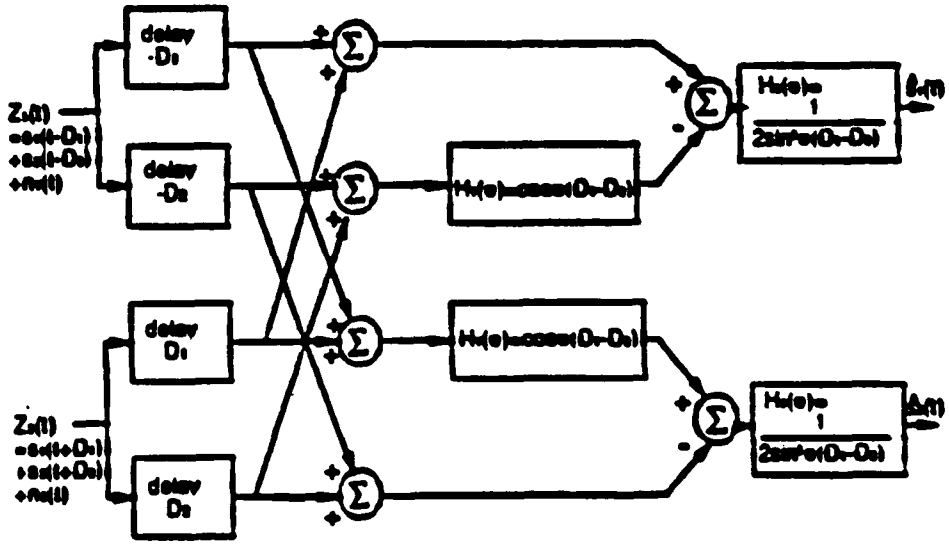


Fig. 1 Direct Implementation of the Least-Squares Operator.

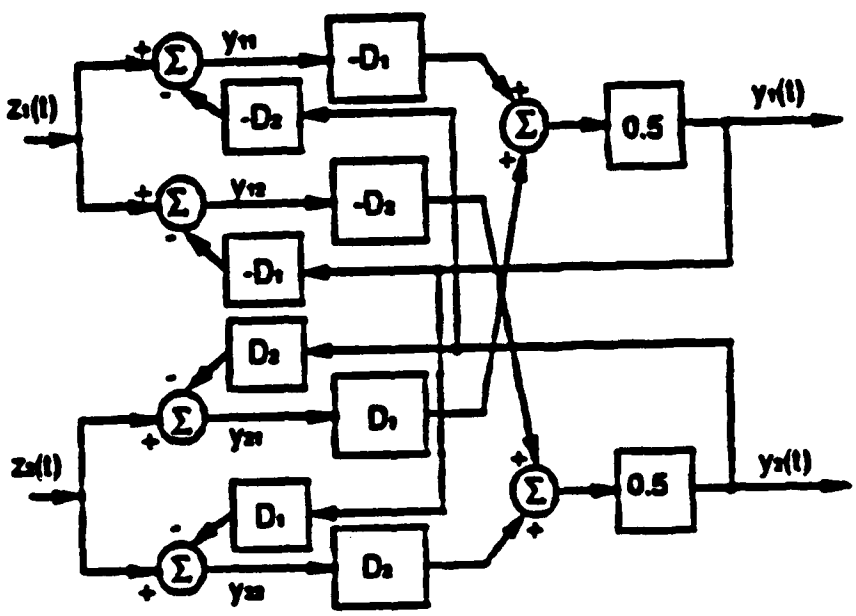


Fig. 2 Bootstrapped Implementation of the Least-Squares Separator.

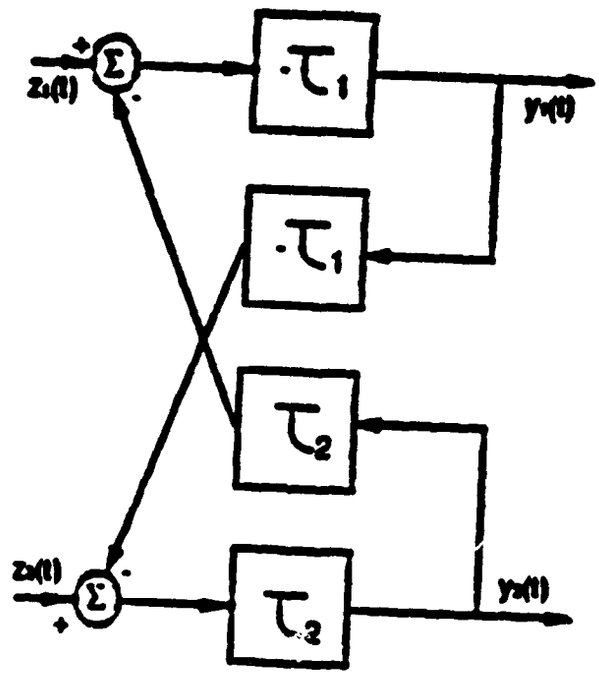


Fig. 3 Alternative Implementation of the Bootsrapped Least-Squares Seperator.

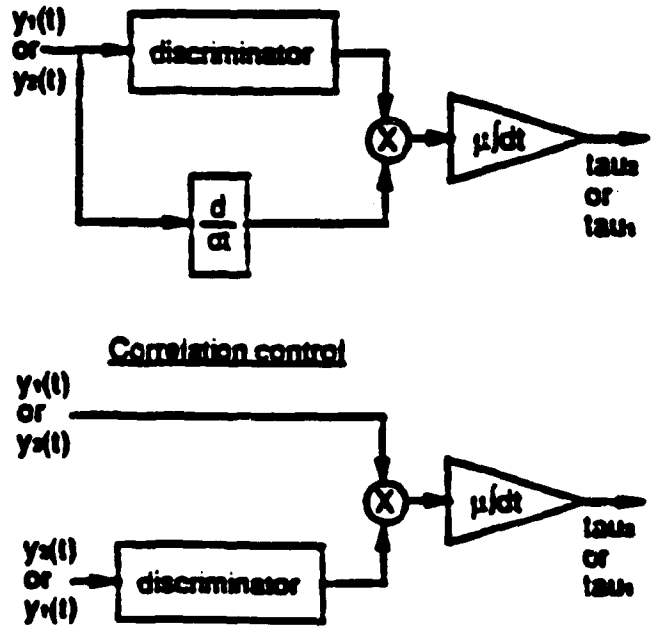


Fig. 4 Block Diagram of the Delay Control Algorithms.

*U.S. GOVERNMENT PRINTING OFFICE: 1993-710-093-60075

MISSION

OF

ROME LABORATORY

Rome Laboratory plans and executes an interdisciplinary program in research, development, test, and technology transition in support of Air Force Command, Control, Communications and Intelligence (C^3I) activities for all Air Force platforms. It also executes selected acquisition programs in several areas of expertise. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C^3I systems. In addition, Rome Laboratory's technology supports other AFSC Product Divisions, the Air Force user community, and other DOD and non-DOD agencies. Rome Laboratory maintains technical competence and research programs in areas including, but not limited to, communications, command and control, battle management, intelligence information processing, computational sciences and software producibility, wide area surveillance/sensors, signal processing, solid state sciences, photonics, electromagnetic technology, superconductivity, and electronic reliability/maintainability and testability.