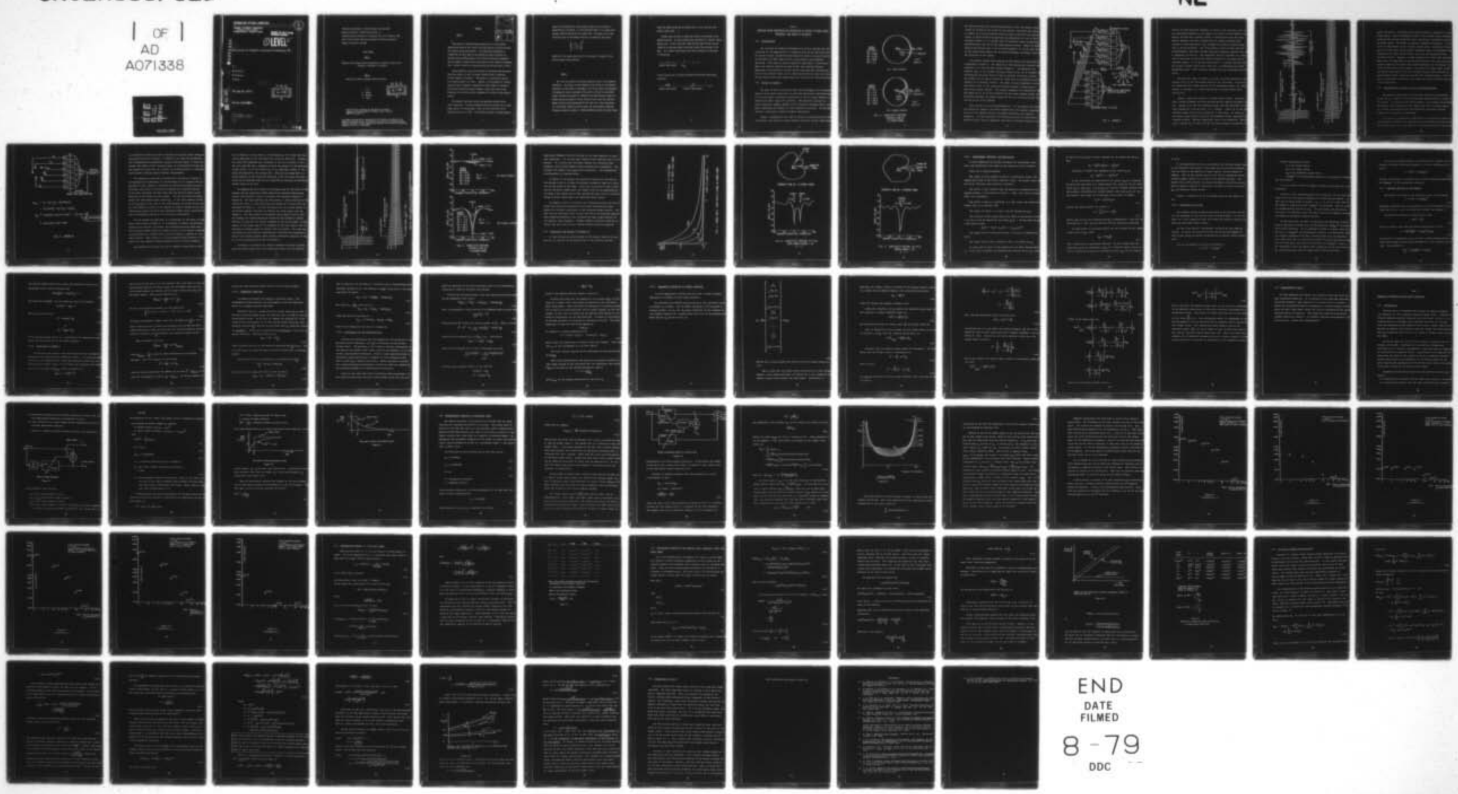
AD-A071 338 STANFORD UNIV CALIF INFORMATION SYSTEMS LAB F/G 17/4
RESEARCH ON ADAPTIVE ANTENNA TECHNIQUES II.(U)
MAR 78 B WIDROW, R CHESTEK, T SAXE N00019-77-C-0194

UNCLASSIFIED NL

| OF |
AD
A071338



END
DATE
FILMED
8-79
DDC

INFORMATION SYSTEMS LABORATORY

STANFORD ELECTRONICS LABORATORIES
DEPARTMENT OF ELECTRICAL ENGINEERING
STANFORD UNIVERSITY · STANFORD, CA 94305

(1) LEVEL II

AD A071338

(6)

Research on Adaptive Antenna Techniques II.

by

(10) B. Widrow
R. Chestek
T. Saxe

(12) 78p.

(11) March 14, 1978

(9) FINAL REPORT.

Prepared under
Contract N00019-77-C-0194 (15)

406 720 LB

79 07 13 039

Principle Investigator: Bernard Widrow (415) 497-4949

Name of Contractor: Stanford University

Effective Date of Contract: 15 February 1977 to 14 February 1978

Short Title of Work: Research on Adaptive Antenna Techniques

Amount of Contract: $54,426

FINAL REPORT

PART 1

Adaptive Array Processing for Separation of Inputs by Power Level, Frequency, and Angle of Incidence

PART 2

Analysis of Adaptive Weight Noise Covariance

by

B. Widrow

R. Chestek

T. Saxe

This work was supported by the Naval Air Systems Command of the Department of Defense under contract N00019-77-C-0194

SUMMARY

PART 1

The concept of adaptive power separation for single channel applications based on the "leaky" LMS algorithm has now been extended for adaptive array applications. A new algorithm designated as "scheme 6A" has been shown to be potentially highly effective for nulling strong jamming signals in an aircraft receiving environment. The new scheme is substantially less noisy and more simple to implement than "scheme 6" which was analysed during our previous year's effort.

Using scheme 6A, one could establish an omnidirectional quiescent receiving pattern in order to accept incoming signals regardless of their angles of incidence. This pattern is sustained as long as all incoming signals are "weak". However, in the presence of "strong" incoming signals (jammers), regardless of their angles of incidence, the quiescent receiving pattern changes as deep nulls form in the directions of the strong signals (jammers) as a result of the adaptive process.

The adaptive algorithm sustains the quiescent pattern with a "soft" constraint. A conflict develops with the incidence of a strong input, which, if strong enough, causes the soft constraint to be violated and a null to form. The separation between a strong (jammer)

input to be rejected and a weak (signal) input to be accepted is determined by a parameter γ in the algorithm which, in an operational system, could be controlled by a panel knob. The power gain of the receiving array in the jammer direction is reduced by the factor:

$$\left[\frac{1}{1 + \frac{\sigma_s^2}{\gamma} N}\right]^2$$

where σ_s^2 is the jammer power and N is the number of weights in the entire antenna array processor.

PART 2

All real-time adaptive processes experience noise in the adaptive parameters. The amount of noise depends on the nature of the adaptive algorithm, on the number of parameters, and on the speed of convergence. A fundamental study of parameter noise and its effects on the output signal has been undertaken for stochastic and deterministic inputs to weight-controlled adaptive filters driven by the leaky LMS algorithm. Weight noise has been determined for the case of an input consisting of a sinusoidal signal plus white noise, for wide ranges of SNR and frequency relative to Nyquist. Broad operating regions have been found where the output power due to weight noise is less than 5% of the output signal power.

Weight noise also has a significant effect on stability of the adaptive process. An exact analysis was performed of a special single-weight case. It was found that leaky LMS (and LMS) filters may not stabilize in the mean-square sense even though they converge in the mean. It is known from previous work that convergence in the mean is insured by:

$$\frac{1}{\text{(total input power)}} > \frac{1}{\lambda_{max}} > \mu > 0$$

A new criterion for μ has been found which guarantees mean-square stability:

$$\frac{0.288}{\text{(total input power)}} = \frac{\sqrt{3}}{6\,\text{(total input power)}} > \mu > 0$$

PART 1

ADAPTIVE ARRAY PROCESSING FOR SEPARATION OF INPUTS BY POWER LEVEL, FREQUENCY, AND ANGLE OF INCIDENCE

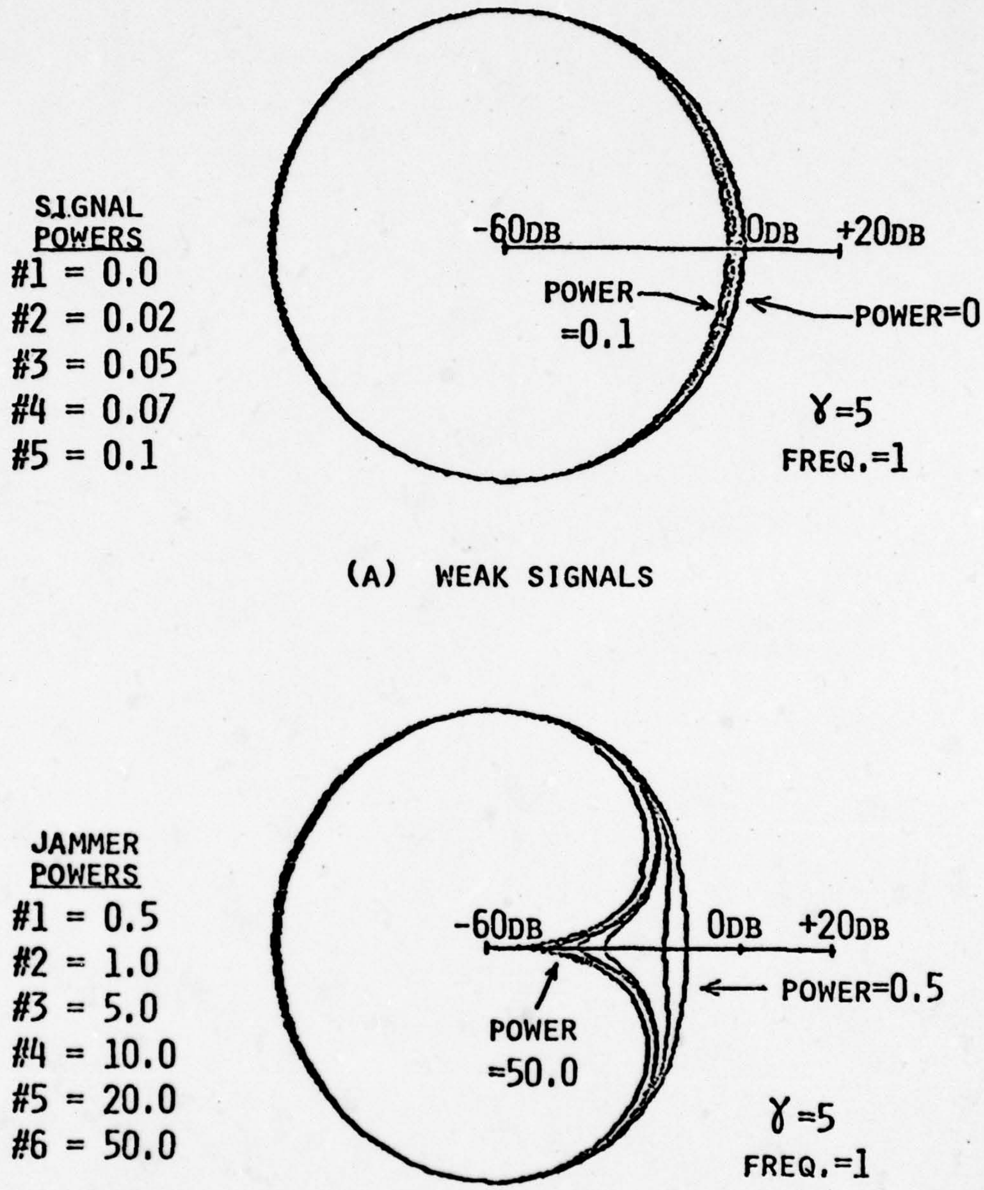
1-A. Introduction

Our research on adaptive techniques for antijam systems has concentrated on the application of a modified form of the LMS algorithm to antenna arrays, permitting discrimination against received signals on the basis of their power levels, using spatial and frequency filtering. This type of antijam system is applicable to situations where the desired signals are much weaker than jamming signals. Knowledge of direction of arrival and of other specific characteristics of the desired signal is not required beforehand since the adaptation process uses only signal power as the basis for discrimination.

1-B. Review of Scheme 6

The goal of this activity has been the development of an adaptive antenna array whose sensitivity is high and essentially omnidirectional to weak inputs, and whose sensitivity is low to strong signals regardless of their angles of incidence. The objective has been to cause high power jamming signals to be severely attenuated while lower power communication signals are only slightly attenuated. We thereby realize a substantial signal-to-jammer improvement.

Figure 1 demonstrates the type of antenna reception patterns that are desired (and which have been achieved with the current algorithm).



(B) STRONG JAMMERS

FIG. I. DIRECTIVITY PATTERNS
SCHEMES 6 AND 6A
6 ELEMENT ARRAY

The patterns shown are directivity patterns, where the antenna array itself is located at the origin of the pattern, and where the signal or jammer is arriving directly from the right (along the axis shown). Figure 1a shows directivity patterns corresponding to a set of weak incoming signals. It is evident that the patterns remain essentially omnidirectional for all of the signal powers listed. Figure 1b shows directivity patterns for a set of strong jammers. It can be seen that as the jammer power grows, the array places deeper and deeper nulls in the direction of the jammer.

An adaptive system which behaves in the manner just described is diagrammed in Figure 2. The behavior of this system was examined and reported in [1] and has been called "Scheme 6". We summarize its behavior here. A six element circular antenna array (as an example) is processed by six slave filters (TF_i) to produce the array's output. The weights of the slave filters are taken from a corresponding set of adaptive or "training" filters. The input to each independent training filter (TF_i) is the associated antenna element's signal plus a white noise "pilot signal". The impulse responses are adjusted by the LMS algorithm so as to best minimize (in the least squares sense) the difference between the summed outputs of the six training filters and the sum of the pilot signals.

This results in the following behavior: in the absence of incoming signals or jammers, the pilot signals force the impulse responses of all of the adaptive filters to become zero-delay unit impulse responses. In this situation, the antenna's pattern is essentially omnidirectional (due to symmetry). Now when an external signal is re-

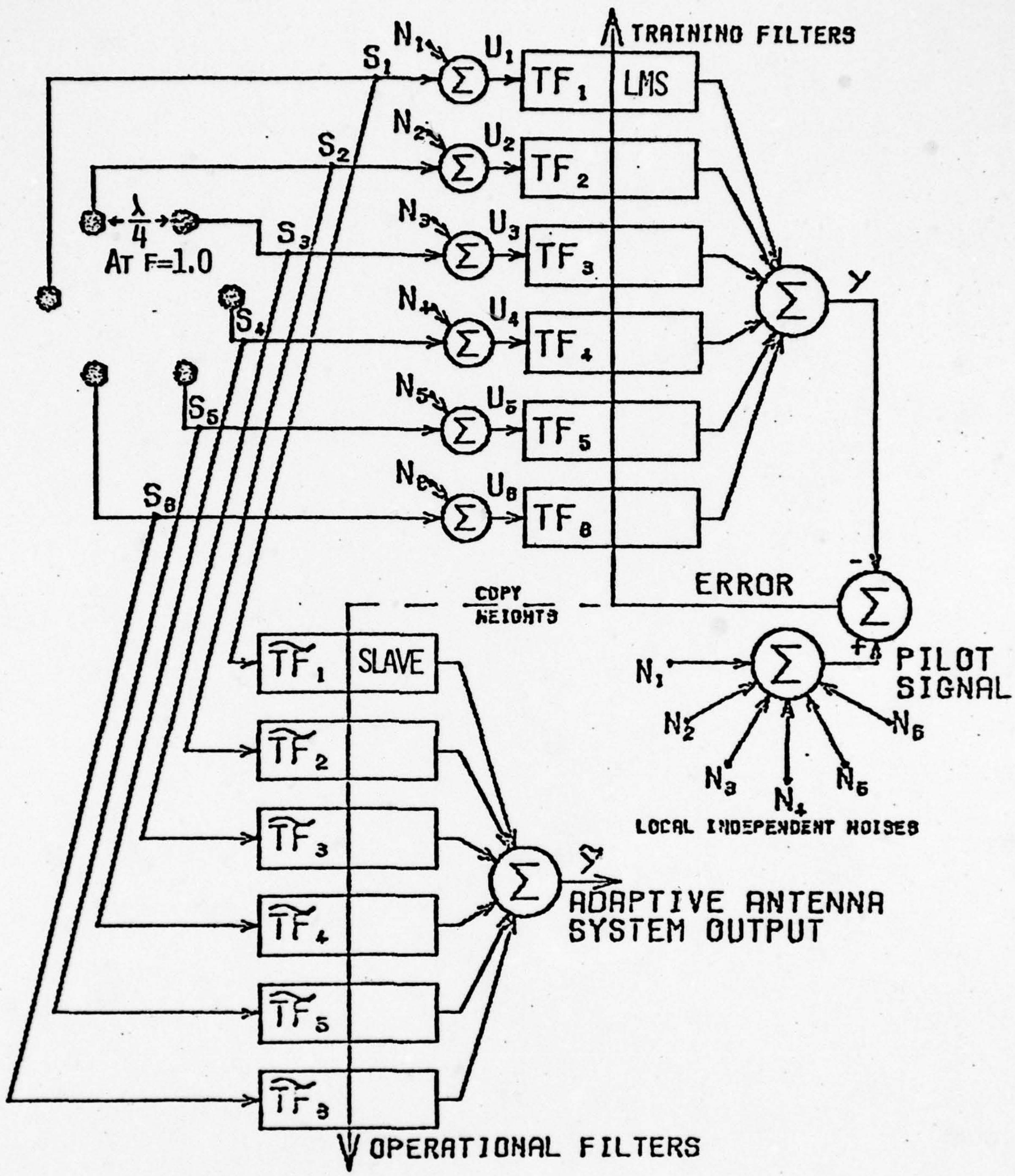


FIG. 2. SCHEME 6

ceived, the LMS algorithm attempts to reject it by creating a null in the sensitivity pattern in the signal's direction at the signal's frequency, because this incoming signal is uncorrelated with the pilot signal. However, the pilot signals and the adaptive process have created a "soft constraint" that attempts to maintain uniform reception in every direction. The array pattern cannot be omni and null simultaneously. Therefore the LMS algorithm computes a pattern that is a compromise of the two. The compromise solution achieved is a function of the power of the pilot signals relative to the power of the received signal. As we can see from Figure 1, a strong received signal causes a deeper notch than a weak one, accomplishing the stated beamforming goal. This intuitive argument has been confirmed analytically.

The patterns shown in Figure 1 are the receiving patterns of the antenna array at the convergence point of the algorithm. In reality, the algorithm never stops exactly at the convergence point, but moves around it slightly in a random fashion due to weight vector noise [2],[3],[4].

Figure 3 shows the effects of weight vector noise. We assume that a single sinusoid is being received by the computer-simulated array. Figure 3a is the time waveform output (the "RF waveform") of the array system for a strong input, i.e. a jammer. Figure 3b is the time waveform output of the array system for a weak input. In each case the figure shows several cycles of the sinusoids before adaptation is allowed to begin. Without adaptation, there is no rejection. Referring to Figure 3a, we note that for the strong jammer, attenuation is

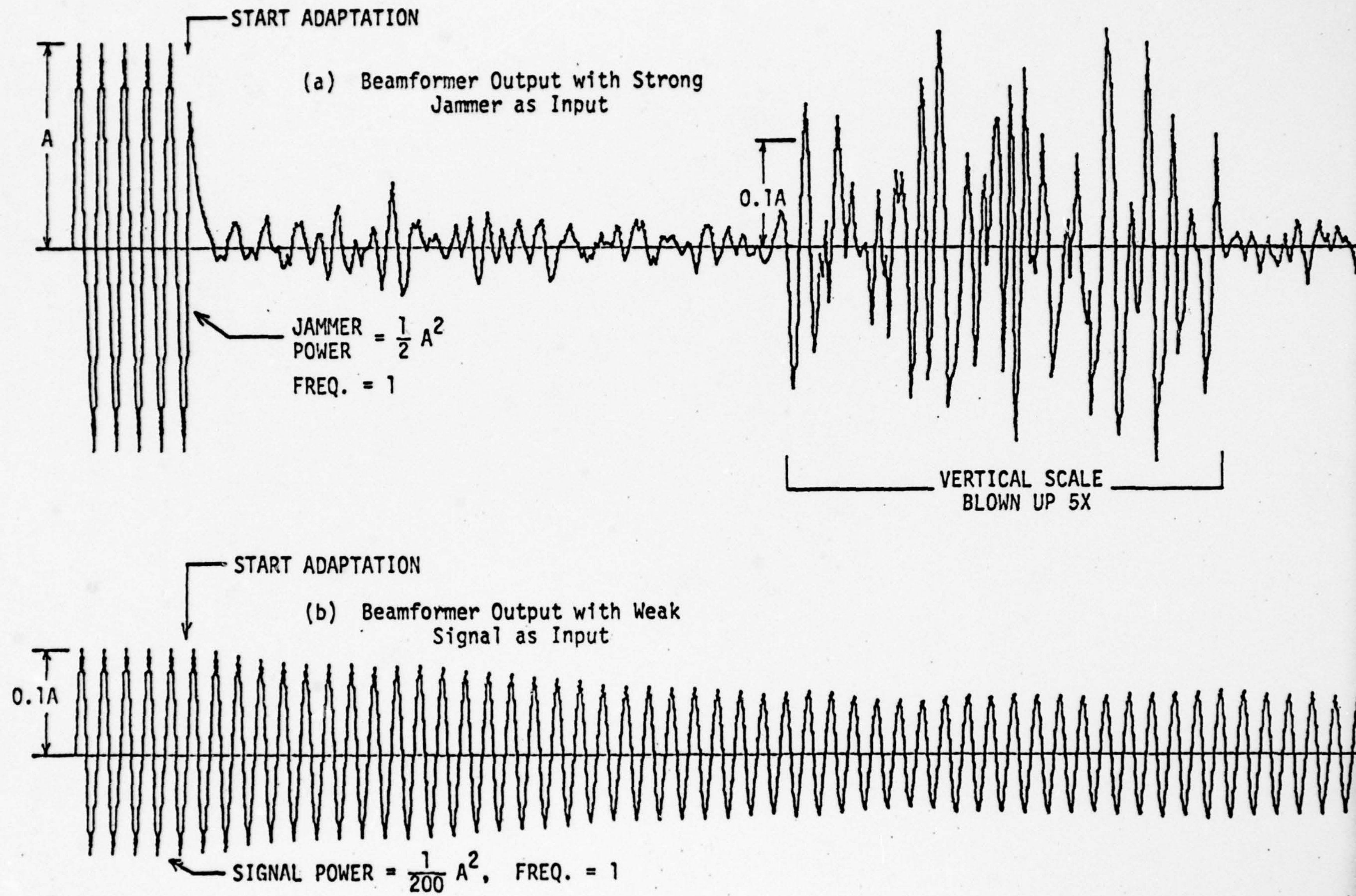


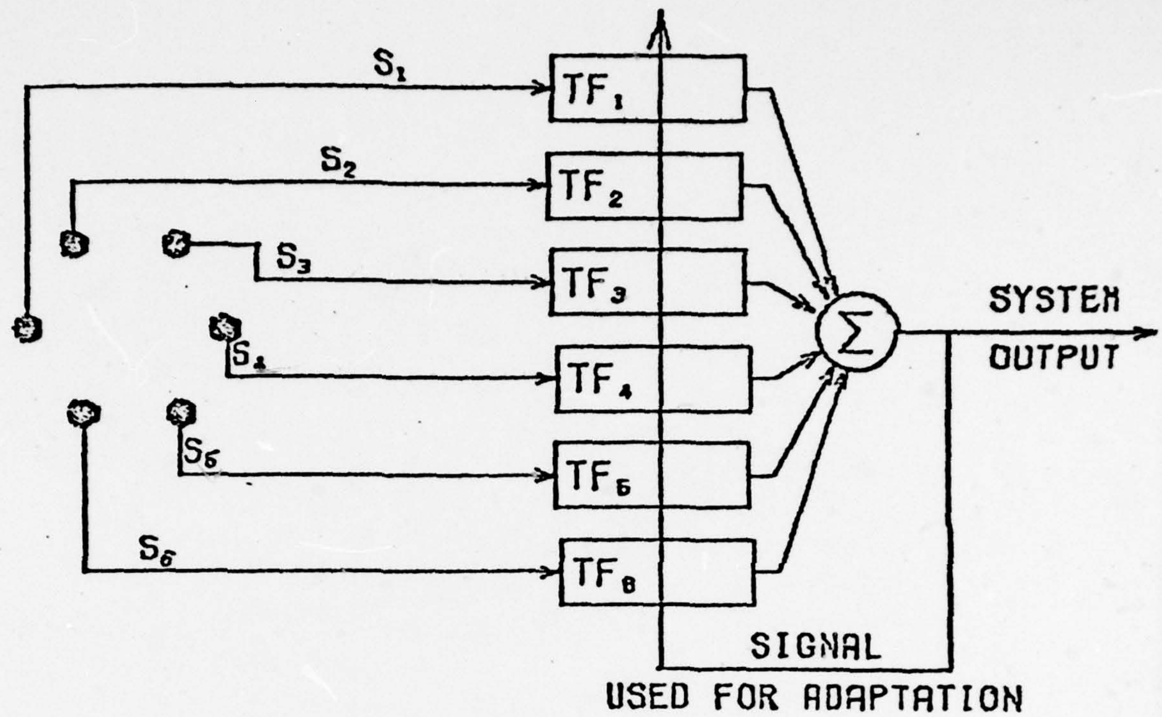
Figure 3: Output Waveforms of the Scheme 6 Adaptive Beamformer with 6 Element Circular Array

rapid, but noisy. A portion of the time waveform is plotted on a magnified scale. After convergence, noise in the weight vector randomly modulates the strong sinusoidal input, producing a substantial amount of output noise. The jammer is thus not completely rejected. Its effects are manifested in the residual output noise. Referring to Figure 3b, we note that the low power sinusoidal input is attenuated more slowly and to a much lesser extent, but that its essential characteristics remain. However, a close examination of this time waveform shows that its amplitude is varying somewhat over time. The array processing algorithm has added some random modulation to the signal. Because of the noise in the weights, the amount of output power resulting (after convergence) from the incoming strong signal exceeds the output signal power resulting from the weak incoming signal. This is not a satisfactory situation, even though the jammer has been substantially attenuated.

1-C. Introduction to Scheme 6A and its Characteristics

The behavior described in the previous section is excessively noisy and is a result of the LMS algorithm being driven by a noise pilot signal. To alleviate the noise problem, a new system was proposed and tested. This system does not require the use of a pilot signal. It is much less noisy and has the additional advantage of being simpler in its hardware requirements.

The system, and adaptation algorithm, are presented in Figure 4 and labeled as "Scheme 6A." Using the same six element circular array for illustration, the array signals are processed by only a single set



$$W_{J+1} = W_J + 2\mu\epsilon_J X_J + 2\mu\gamma(W_Q - W_J)$$

$$= (1-2\mu\gamma)W_J + 2\mu\epsilon_J X_J + 2\mu\gamma W_Q$$

$W_Q \triangleq$ "QUIESCENT WEIGHT VECTOR" $= [1\ 0\ 0 \ldots 0]^T$ (W_Q FOR SCHEME 6)

$\triangleq$ EQUIVALENT PILOT POWER

FIG. 4. SCHEME 6A

of transversal filters in order to produce the system output (Scheme 6 required two sets of filters). In Scheme 6 the required performance was accomplished by introducing a unique training signal. In this new scheme, the effect of the training signal is accomplished directly by the adaptation algorithm. As a result, no training signal is required, significantly reducing system hardware requirements.

The adaptation algorithm of Scheme 6A is presented in Figure 4. By examining this algorithm, its relationship to the standard LMS algorithm is clear. However, the Scheme 6A algorithm includes one more "driving" term which is necessary to produce the required behavior. To use this algorithm, one must know what the desired weight vector would be in the absence of any signals. We call this desired weight vector the "quiescent weight vector" $\underline{W}_Q$. If for each adaptive filter we set $\underline{W}_Q$ equal to the mean of the corresponding converged weight vector that Scheme 6 attains in the absence of inputs, Scheme 6A will automatically produce the same mean converged weight vector solutions as will be produced by Scheme 6, in all signal environments.

For the Scheme 6A algorithm, γ is equivalent to the power of each pilot signal used in Scheme 6. It is possible to choose any quiescent weight vector desired. To produce the receiving patterns shown here, the quiescent weight vector was chosen to have a unit impulse response at zero delay for a single filter, and zero impulse response for all other filters. This results in a truly omnidirectional pattern in the absence of any signals (assuming omnidirectional antenna elements).

By examining the second form of the Scheme 6A algorithm as written in Figure 4, we may notice a relationship to the "leaky LMS algorithm" described in [1]. The first two terms are identical. However, the leaky LMS algorithm has a tendency to drive the weights to zero. The new algorithm has a tendency to drive the weights to the quiescent weight vector. It is the third term in our algorithm, absent in the leaky LMS algorithm, that causes this. Thus the new algorithm is a generalization of the leaky LMS algorithm, since the leaky LMS algorithm can be obtained from the new algorithm by choosing the quiescent weight vector to be zero.

To demonstrate that Scheme 6A eliminates most of the noise in the weights and the associated random modulation problems demonstrated earlier with Scheme 6, Figure 5 presents time waveform outputs of Scheme 6A. The same signals used with Scheme 6, which resulted in the outputs of Figure 3, have been used with Scheme 6A in generating the waveforms of Figure 5. We note once again that the strong jammer is attenuated rapidly, and does not display the noisy output seen for Scheme 6. The weak signal is attenuated only slightly, and it does not display the random modulation distortion that Scheme 6 induced. Finally, we note that the output response to the strong jammer is much weaker at the output than is the response to the weak signal. The improvement in output jammer to signal ratio is evident. An array system using the new Scheme 6A algorithm and receiving both information signals and jamming signals would exhibit performance superior to that of an array system based on the old Scheme 6.

In figure 6 we present the antenna reception patterns attained when the weight vector is at the convergence point for the adaptive

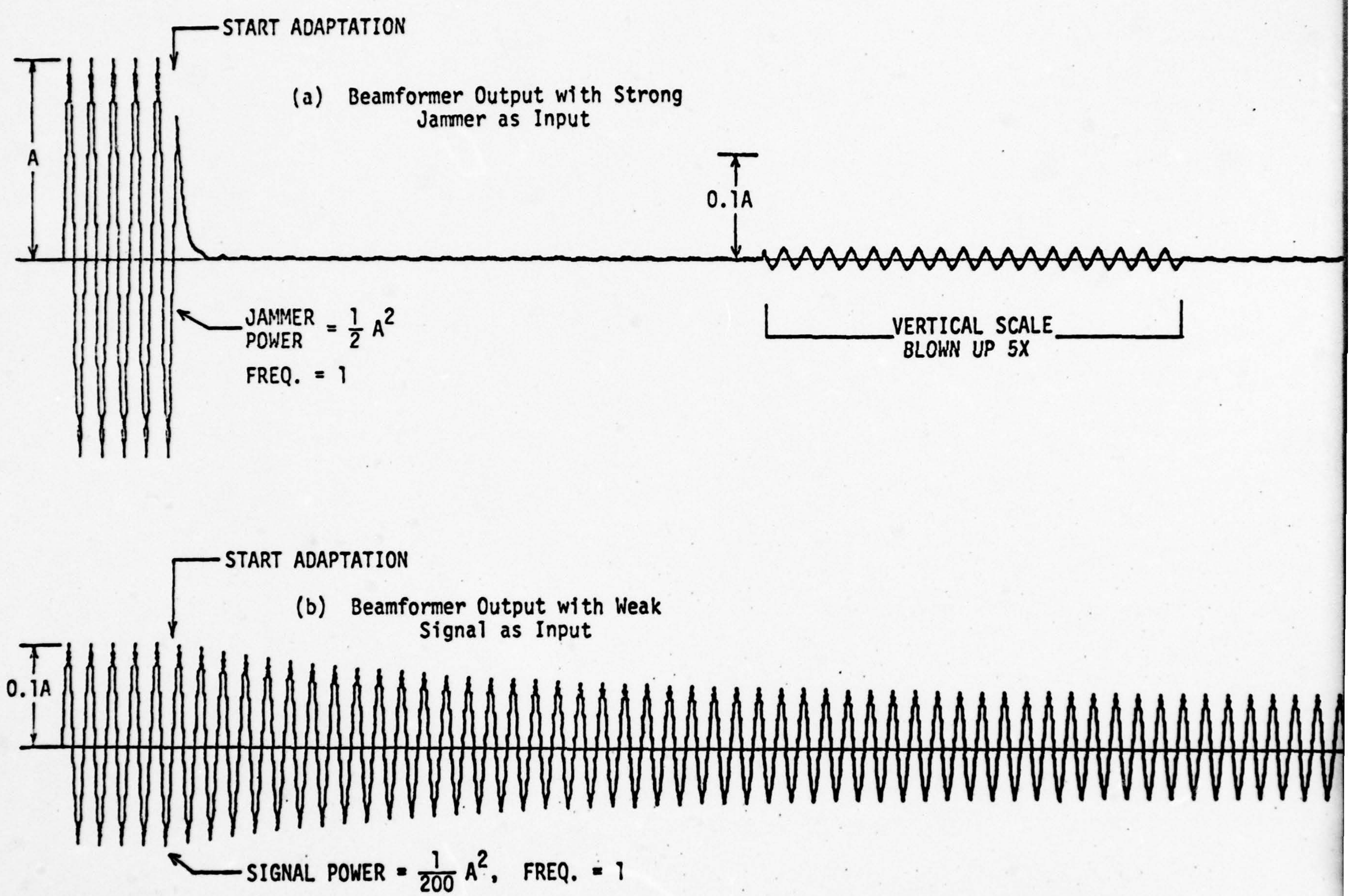


Figure 5: Output waveforms of the Scheme 6A Adaptive Beamformer with 6 Element Circular Array

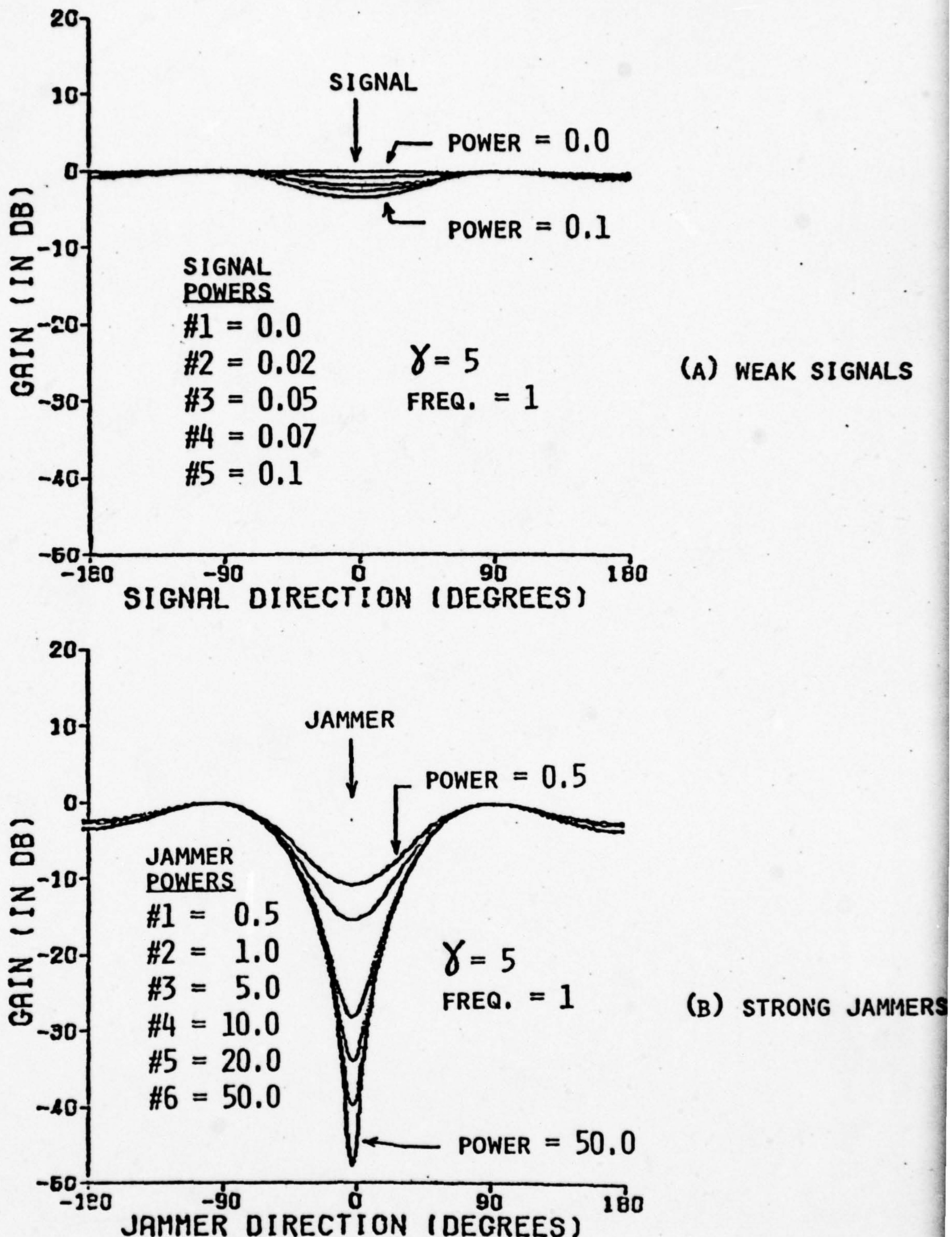


FIG. 6. DIRECTIVITY PATTERNS
SCHEMES 6 AND 6A
6 ELEMENT ARRAY

algorithms (Schemes 6 and 6A converge to the same expected, or average, solutions -- it is only their dynamic noise behavior that is different). It is clear that low power signals are attenuated only slightly, while strong jammers receive significant attenuation. The stronger the jammer, the greater the attenuation. The mathematical relationships are reported below.

In Figure 7, we present the theoretical curves for the gain of the antenna array to a single directional sinusoidal input as a function of the power of the input. Since the array gain is also a function of the pilot signal power for Scheme 6 and of the equivalent pilot power γ for Scheme 6A, we have plotted the gain curve for various values of pilot signal power (or equivalent pilot power).

In figures 8 and 9, we present the receiving patterns for the antenna array when two sinusoids of very close frequency and equal powers are received (calculated at the converged weight vector). Except for the necessary nulls, we see that the array maintains approximate omnidirectionality, even including the space angle between the two jammers, where possible. Sharper angular resolution could be attained, but more than the six antenna elements would be required.

1-D. Definition and Analysis of Scheme 6A.

In this section we define Scheme 6A and analyze some of its properties, to confirm the statements made in the previous section.

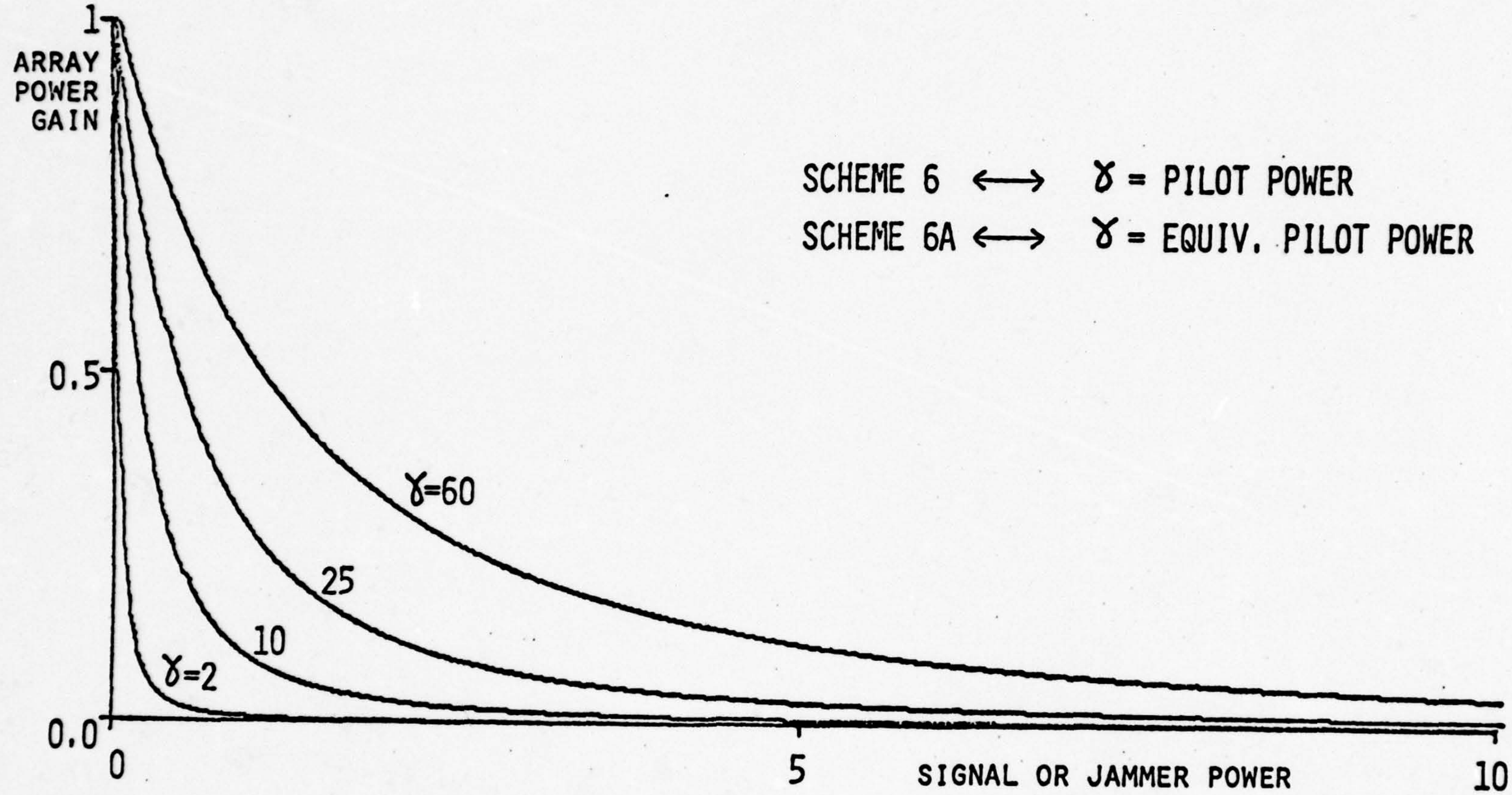
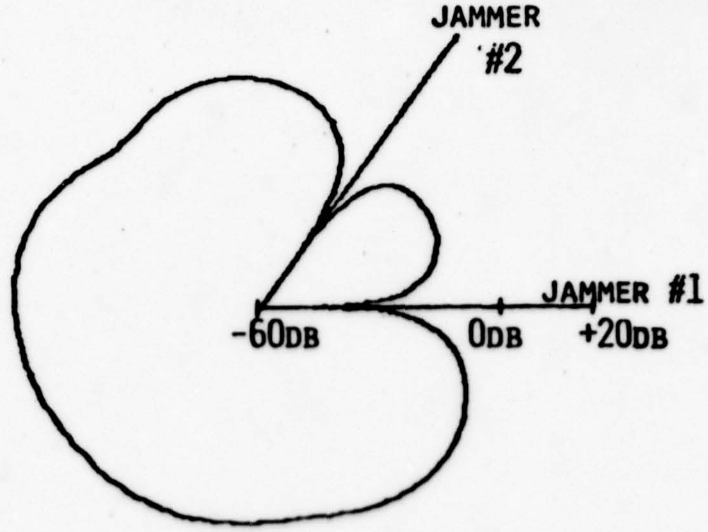


FIG. 7. SINGLE INPUT: ARRAY GAIN vs SIGNAL OR JAMMER POWER



SCHEMES 6 AND 6A, 6 ELEMENT ARRAY

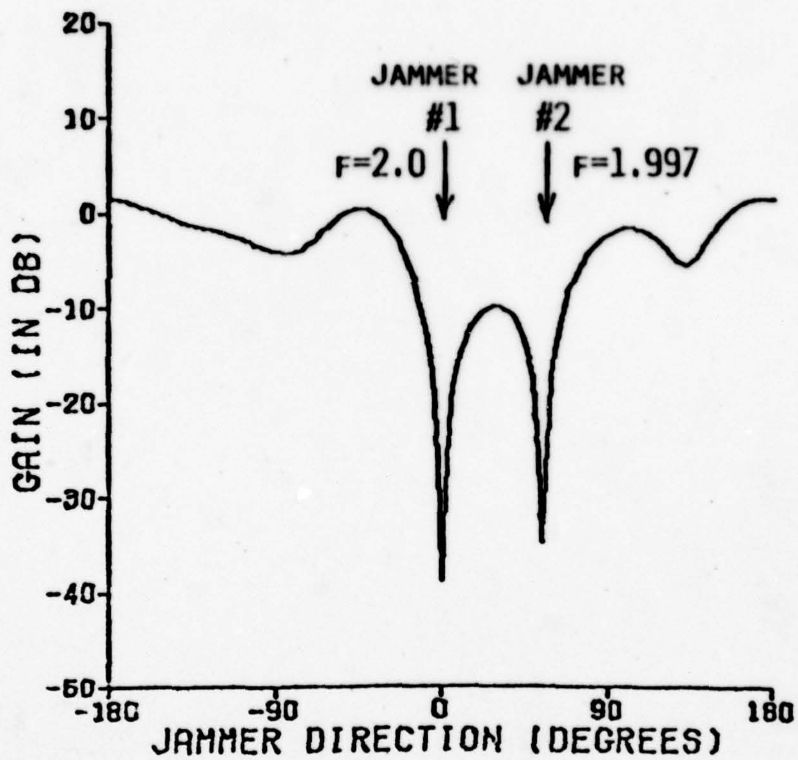
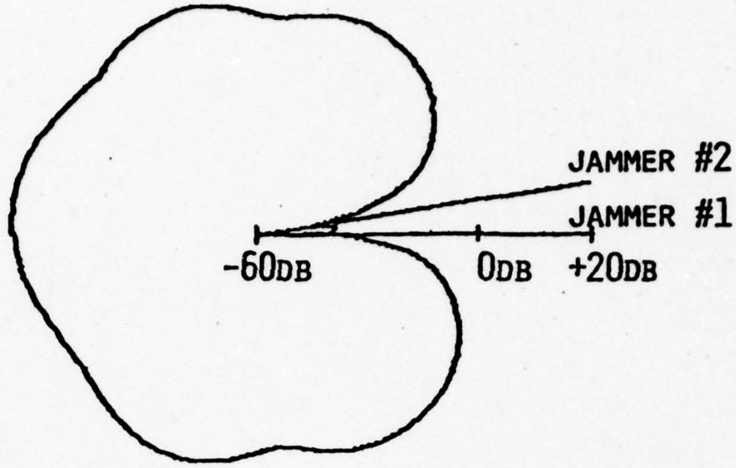


FIG. 8. DIRECTIVITY PATTERN, AT F=2.0
BOTH JAMMER POWERS = 20
$\gamma = 5$



SCHEMES 6 AND 6A, 6 ELEMENT ARRAY

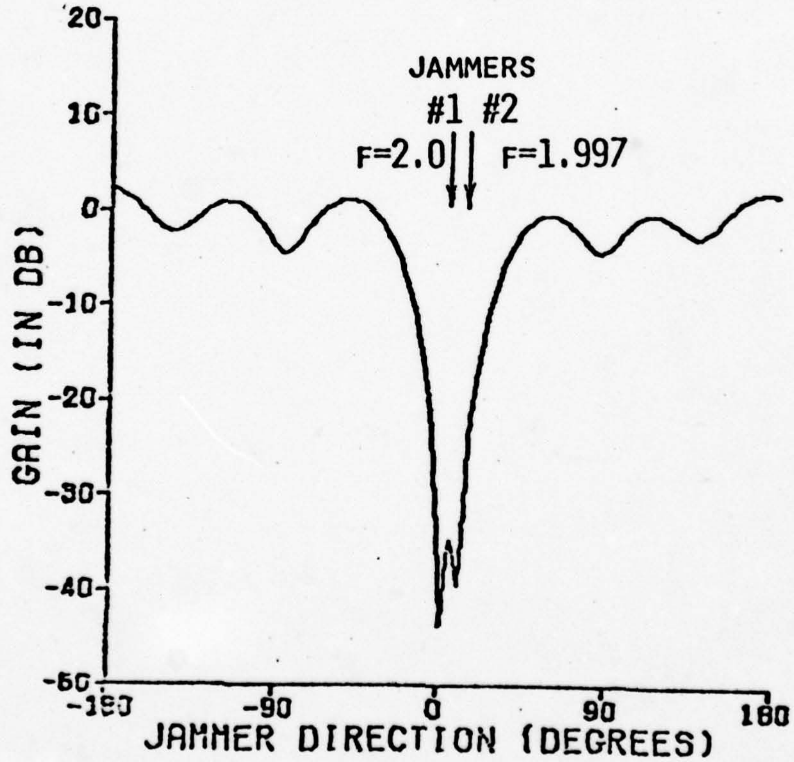


FIG. 9. DIRECTIVITY PATTERN, AT F=2.0
JAMMER POWERS = 20
$\gamma = 5$

1-D-1. Terminology, Notation, and Definitions

In this subsection we briefly introduce the terminology, notation, and definitions to be used in the remainder of this section.

There are k antenna elements.

The output of each element is fed to a transversal filter. The tapped delay line of the filter contains n taps. The weight vector of the filter therefore also contains n elements.

The letter i will indicate that the quantity is associated with antenna element i or transversal filter i. Thus i can take on values from 1 to k inclusive.

The letter j used as a subscript is a time index, and indicates a sample taken at a specific time.

The output of sensor i at time j will be denoted by $x_j(i)$.

The contents of the tapped delay line (TDL) of transversal filter i at time j will be denoted by the vector $\underline{X}_j(i)$. We see from the way a TDL operates that

$$\underline{X}_j(i) = [x_j(i)\ x_{j-1}(i)\ \ldots\ x_{j-k+1}(i)]^T$$

The output of the transversal filter i at time j is denoted by $y_j(i)$.

The output of the entire system at time j is denoted by y_j.

We will need to refer to the contents of all TDL's simultaneously. To do this, we define an augmented TDL contents vector $\underline{X}_j$, which is the set of all $\underline{X}_j(i)$ vectors "stacked up" to produce one vector. Thus

$$\underline{X}_j = [\underline{X}_j^T(1) \vdots \underline{X}_j^T(2) \vdots \ldots \vdots \underline{X}_j^T(k)]^T$$

Similarly, we define the augmented weight vector $\underline{W}_j$ as

$$\underline{W}_j = [\underline{W}_j^T(1) \vdots \ldots \vdots \underline{W}_j^T(k)]^T$$

We can now define the operation of the system: $\underline{X}_j(i)$ is obtained by doing the time shift of a tapped delay line, using as the new input value $x_j(i)$. Thus $\underline{X}_j(i)$ is just $\underline{X}_{j-1}(i)$ with all elements shifted down one position (discarding the bottom element) and using $x_j(i)$ for the top element. The output for transversal filter TF(i) is simply

$$y_j(i) = \underline{W}_j^T(i)\underline{X}_j(i) \tag{1-1}$$

and for the entire system the output is

$$y_j = \sum_{i=1}^{k} y_j(i) = \underline{W}_j^T\underline{X}_j \tag{1-2}$$

Notice that we have not discussed how $\underline{W}_j$ is determined -- this will be described in a later section dealing with the adaptation algorithm.

We next define a covariance matrix for the contents of the tapped delay lines as $\underline{R}_{xx}$, where

$$\underline{R}_{xx} = E[\underline{X}_j\underline{X}_j^T] \tag{1-3}$$

with E denoting the expectation operator. We will assume that the processes $x_j(i)$ are stationary so $\underline{R}_{xx}$ is a constant matrix. Since we will be referring to just one covariance matrix, $\underline{R}_{xx}$, we will denote it by $\underline{R}$.

In the application of this array scheme, one flexible aspect open to the designer's discretion is the shape of the directivity pattern that one wants in the absence of input signals, desired signals or jammers. This is the nominal pattern that will be notched by strong incoming jammer signals as a result of the adaptive process. The weight vector that provides the desired quiescent directivity pattern when no signals (desired or otherwise) are being received is denoted $\underline{W}_Q$, the quiescent weight vector.

Lastly, we define $\underline{W}^*$ to be the optimum value of the weight vector.

1-D-2. Performance Criterion

The adaptive system Scheme 6A presented in the previous section will be developed in this section from first principles. It will be shown in a later subsection that this system and Scheme 6 produce the same mean converged weight vectors.

We will first define a performance criterion for the adaptive system. In accord with this criterion, the current performance is used by the system to modify its parameters to improve future performance.

Let the performance criterion be defined as:

$$J = E[y^2] + ||\underline{W}-\underline{W}_Q||^2 \tag{1-4}$$

Recall from section B that:

y is the system output

$\underline{W}$ is the weight vector

$\underline{W}_Q$ is the quiescent weight vector

$||\underline{W}-\underline{W}_Q||^2$ is the magnitude squared of the difference between the two vectors.

It is our goal to have the adaptive system find $\underline{W}$ such that J is minimized.

To gain a clear understanding of this criterion, let us first examine it in the case when no signals are received by the system. In this case no input means no output, i.e. $y = 0$. So it is the task of the system to find $\underline{W}$ such that $||\underline{W}-\underline{W}_Q||^2$ is minimized. This is clearly accomplished by setting $\underline{W} = \underline{W}_Q$. In other words, the performance criterion J is minimized in the no signal case by the adaptive process attaining the quiescent reception pattern.

In the case where an input is available, the following tradeoff occurs: by making a change in $\underline{W}$, the magnitude of $||\underline{W}-\underline{W}_Q||^2$ increases, while $E[y^2]$ decreases. If the decrease in $E[y^2]$ is greater than the increase in $||\underline{W}-\underline{W}_Q||^2$, the performance criterion J is decreased. What the adaptive system will do then, is find $\underline{W}$ such that any decrease that would occur due to a decrease in $E[y^2]$ is exactly balanced by any increase that would occur due to an increase in $||\underline{W}-\underline{W}_Q||^2$. What we see happening is that the system will attempt to have $\underline{W}$ stay close to $\underline{W}_Q$, only moving away when the output power $E[y^2]$ grows large.

To allow the system designer to influence this tradeoff, we introduce a designer controlled parameter γ into the performance criterion:

$$J = E[y^2] + \gamma||\underline{W}-\underline{W}_Q||^2 \tag{1-5}$$

As such, the designer can control the relative tradeoff between $E[y^2]$ and $||\underline{W}-\underline{W}_Q||^2$ in the performance criterion.

1-D-3. Optimum solution to the problem

The problem as stated in the previous section is to minimize

$$J = E[y^2] + \gamma||\underline{W}-\underline{W}_Q||^2 \tag{1-6}$$

by proper selection of $\underline{W}$. Using gradients, we can determine the optimum value of $\underline{W}$ -- that is, the $\underline{W}$ which yields the minimum J.

We begin by rewriting J, using the fact that

$$||\underline{V}||^2 = \underline{V}^T\underline{V} \tag{1-7}$$

(for $\underline{V}$ a vector), and using the system equation for y (1-2):

$$J = E[\underline{W}^T\underline{X}_j\underline{X}_j^T\underline{W}] + \gamma(\underline{W}-\underline{W}_Q)^T(\underline{W}-\underline{W}_Q) \tag{1-8}$$

We see that J is quadratic in $\underline{W}$, so that a unique minimum exists. We find this minimum by setting the gradient of J with respect to $\underline{W}$ to zero. The gradient is

$$\nabla_{\underline{W}} J = 2E[\underline{X}_j\underline{X}_j^T]\underline{W} + 2\gamma(\underline{W}-\underline{W}_Q) \tag{1-9}$$

The optimal weight vector that causes the gradient to be zero is designated as $\underline{W}^*$, and is obtained from

$$2E[\underline{X}_j\underline{X}_j^T]\underline{W}^* + 2\gamma(\underline{W}^*-\underline{W}_Q) = 0 \tag{1-10}$$

Recalling that $E[\underline{X}_j\underline{X}_j^T] = \underline{R}$, and gathering terms in $\underline{W}^*$ yields:

$$(\underline{R}+\gamma\underline{I})\underline{W}^* - \gamma\underline{W}_Q = 0 \tag{1-11}$$

Now we can solve for $\underline{W}^*$:

$$\underline{W}^* = \gamma(\underline{R}+\gamma\underline{I})^{-1}\underline{W}_Q \tag{1-12}$$

or

$$\underline{W}^* = (\frac{1}{\gamma}\underline{R}+\underline{I})^{-1}\underline{W}_Q \tag{1-13}$$

So the optimum weight vector is a function of the quiescent weight vector and the covariance of the inputs signals.

1-D-4. Comparison to Scheme 6

In reference [1], Scheme 6 was also referred to as an Adaptive Beamformer with Injected Noise (ABWIN for short). It is now our goal to show that the ABWIN and the scheme just presented yield the same converged mean weight vector. Let us denote the converged mean weight vector for the ABWIN as $\underline{W}^*_{ABWIN}$. Then [1] showed that

$$\underline{W}^*_{ABWIN} = (\underline{R}+\sigma_n^2\underline{I})^{-1}\underline{P} \tag{1-14}$$

where the $\underline{R}$ is the same as we have defined, σ_n^2 is the power of the injected noise used for the pilot signal for the ABWIN, and $\underline{P}$ is the correlation between the contents of the tapped delay lines $[\underline{X}_j]$ and the pilot signal. This can be rewritten as:

$$\underline{W}^*_{ABWIN} = \left(\frac{1}{\sigma_n^2}\underline{R} + \underline{I}\right)^{-1} \frac{1}{\sigma_n^2}\underline{P} \tag{1-15}$$

In [1], in the earlier sections, it was shown that

$$\frac{1}{\sigma_n^2}\underline{P} = [1\ 0\ \cdots\ 0|1\ 0\ \cdots\ 0|\cdots|1\ 0\ \cdots\ 0]^T \tag{1-16}$$

and that the quiescent weight vector for the ABWIN was $\frac{1}{\sigma_n^2}\underline{P}$. In a later section of [1], a method was proposed for altering the pilot signal formation so that any value for the cross-correlation vector $\underline{P}$ could be attained, and $\frac{1}{\sigma_n^2}\underline{P}$ would be the quiescent weight vector.

Thus we showed in [1] that

$$\underline{W}^*_{ABWIN} = \left(\frac{1}{\sigma_n^2}\underline{R} + \underline{I}\right)^{-1}\underline{W}_{Q_{ABWIN}} \tag{1-17}$$

where $\underline{W}_{Q_{ABWIN}} = \frac{1}{\sigma_n^2}\underline{P}$ could be chosen beforehand by the system designer. Now if we compare our new system

$$\underline{W}^* = \left(\frac{1}{\gamma}\underline{R} + \underline{I}\right)^{-1}\underline{W}_Q \tag{1-18}$$

with the above relation for the ABWIN, we see that $\underline{W}^* = \underline{W}^*_{ABWIN}$ if we make the assignments $\gamma = \sigma_n^2$ and $\underline{W}_Q = \underline{W}^*_{ABWIN}$. So the new system can obtain the same solution weight vectors as the original ABWIN.

1-D-5. Adaptation Algorithm

It remains to present the adaptive algorithm itself. The development follows directly the development of the LMS algorithm, which is a steepest descent algorithm.

The basic idea is: assume we have a weight vector $\underline{W}_j$ at time j. We want to find a weight vector for time j+1 that is closer to the optimum weight vector. To do this, we compute the gradient of J with respect to $\underline{W}$, and evaluate it at the current weight vector $\underline{W}_j$. This gradient (denoted $\nabla_{\underline{W}} J$) defines the direction that $\underline{W}_j$ should be altered to increase J. Since we are interested in minimizing J, we go the opposite direction. Mathematically, we set

$$\underline{W}_{j+1} = \underline{W}_j - \mu \nabla_{\underline{W}} J \tag{1-19}$$

where μ governs how far in the direction specified by $\nabla_{\underline{W}} J$ we go. (If μ is too large, we could overshoot our goal so much that J increases again!)

From section 1-D-3 we already have an expression for $\nabla_{\underline{W}} J$:

$$\nabla_{\underline{W}} J = 2\underline{R}\ \underline{W}_j + 2\gamma(\underline{W}_j - \underline{W}_Q) \tag{1-20}$$

Using this in the expression above (1-19) we obtain

$$\underline{W}_{j+1} = \underline{W}_j - 2\mu[\underline{R}\ \underline{W}_j + \gamma(\underline{W}_j - \underline{W}_Q)] \tag{1-21}$$

Now in general we do not know $\underline{R}$. Instead we use an instantaneous but unbiased estimate of $\underline{R}$. The estimate is $\underline{X}_j\underline{X}_j^T$. Using this in the above expression we obtain

$$\underline{W}_{j+1} = \underline{W}_j - 2\mu\underline{X}_j\underline{X}_j^T\underline{W}_j - 2\mu\gamma(\underline{W}_j-\underline{W}_Q) \tag{1-22}$$

Now since $y_j = \underline{X}_j^T\underline{W}_j$ (1-2) we have

$$\underline{W}_{j+1} = \underline{W}_j - 2\mu y_j\underline{X}_j - 2\mu\gamma(\underline{W}_j-\underline{W}_Q) \tag{1-23}$$

which may also be rewritten as

$$\underline{W}_{j+1} = (1-2\mu\gamma)\underline{W}_j - 2\mu y_j\underline{X}_j + 2\mu\gamma\underline{W}_Q \tag{1-24}$$

This is the adaptation rule used in Scheme 6A.

1-D-6. Convergence of the Adaptation Rule

We must now demonstrate that the adaptation rule presented in the previous section converges, and that it converges to the optimum weight vector. The quantity we will study to indicate convergence is the mean of the weight vector. Other criteria are possible, as in stochastic approximation techniques. However, many algorithms based on stochastic approximation have a tendency to "turn themselves off" after a time span, ignoring later data. While this may be suitable for a truly stationary environment, it does not allow any capability for following changes in a nonstationary environment.

First we will show that with a stationary input environment, if the algorithm converges, the mean of the weight vector has only one point to converge to--the only stationary point of the performance criterion J, which is therefore the optimum.

To find the stationary point, take the expectation of both sides of the adaptation rule (1-23):

$$E[\underline{W}_{j+1}] = E[\underline{W}_j] - 2\mu E[y_j \underline{X}_j] - 2\mu\gamma E[\underline{W}_j - \underline{W}_Q] \tag{1-25}$$

Now, at convergence, (which has yet to be demonstrated) we would have

$$\lim_{j \to \infty} E[\underline{W}_j] = \underline{W}^* \tag{1-26}$$

Recalling that $\underline{W}^*$ denotes the optimal weight vector. Thus we write:

$$\underline{W}^* = \underline{W}^* - 2\mu \lim_{j \to \infty} E[\underline{X}_j \underline{X}_j^T \underline{W}_j] - 2\mu\gamma(\underline{W}^* - \underline{W}_Q) \tag{1-27}$$

where we have reexpanded y_j as $\underline{X}_j^T \underline{W}_j$ (1-2). Continuing,

$$2\mu \underline{R}\, \underline{W}^* = -2\mu\gamma \underline{W}^* + 2\mu\gamma \underline{W}_Q \tag{1-28}$$

where we have assumed that, with $\underline{W}_j$ converging as $j \to \infty$,

$$\lim_{j \to \infty} E[\underline{X}_j \underline{X}_j^T \underline{W}_j] = \lim_{j \to \infty} \{E[\underline{X}_j \underline{X}_j^T] E[\underline{W}_j]\}$$
$$= \underline{R}\, \underline{W}^* \tag{1-29}$$

Finally, after grouping terms, we can find $\underline{W}^*$:

$$(\underline{R} + \gamma \underline{I})\underline{W}^* = \gamma \underline{W}_Q$$
$$\underline{W}^* = \gamma(\underline{R} + \gamma \underline{I})^{-1} \underline{W}_Q$$

$$= \left(\frac{1}{\gamma}\underline{R}+\underline{I}\right)^{-1}\underline{W}_Q$$

1-30

which is the optimum solution found in section D.

We must next show that the adaptation rule causes $E[\underline{W}_j]$ to converge to a single value, which must necessarily be the stationary point found above. Since the proof of convergence is directly parallel to that of the LMS algorithm, and is rather lengthy, we will not produce it here, but will point out the modifications needed to the proof of convergence for the LMS algorithm. (A proof of convergence of the LMS algorithm is contained in references 5 and 6). The major difference is that the term which appears as

$$[I+2k_s\Lambda]$$

in reference 6, must be modified to be

$$[I + 2k_s\gamma I + 2k_s\Lambda] = [(1+2k_s\gamma)I + 2k_s\Lambda]$$

1-31

Notice that the eigenvectors in matrix Q have not changed. (Also note that k_s in [6] corresponds to μ in this report).

The other change required is the replacement of the term $\phi(k,d)$ by $2k_s\gamma\underline{W}_Q$.

With these modifications, the proof of convergence follows the same steps leading to the conclusion that the adaptation rule causes $E[\underline{W}_j]$ to converge to the optimum solution so long as

$$0 < \mu < \frac{1}{\gamma+\lambda_{max}}$$

1-32

where λ_{max} is the maximum eigenvalue of the matrix $\underline{R}$.

1-D-7. Response of Scheme 6A to single sinusoids

In this subsection we analyze the gain that a single sinusoid impinging on a Scheme 6A array would encounter.

We represent the sinusoid being received at the individual antenna element as a phasor. Let us denote the phase of the sinusoid at antenna element i as ϕ_i. Let the phase difference of the sinusoid at two adjacent elements of a tapped delay line be θ. If the sinusoid has power σ_s^2 then $\underline{X}_j$ may be written as:

$$\underline{X}_j = \sqrt{2}\sigma_s \begin{bmatrix} e^{j\phi_1} \\ e^{j(\phi_1-\theta)} \\ e^{j(\phi_1-2\theta)} \\ \vdots \\ e^{j(\phi_1-(n-1)\theta)} \\ \hline e^{j\phi_2} \\ e^{j(\phi_2-\theta)} \\ \vdots \\ \hline \vdots \\ \hline e^{j\phi_k} \\ \vdots \\ e^{j(\phi_k-(n-1)\theta)} \end{bmatrix} = \sqrt{2}\sigma_s\underline{V} \tag{1-33}$$

Notice that we have defined the vector $\underline{V}$ to be the above column matrix.

Next we note that the weight vector $\underline{W}$ consists of n real weights. However, using complex notation, we replace the n real weights by n/2 complex weights which produce the same output. Henceforth, we represent all weight vectors as vectors of n/2 complex weights, and $\underline{X}$ as a vector of n/2 complex samples, with autocorrelation matrix

$$\underline{R}_{xx} = \sigma_s^2 \underline{V}\ \underline{V}^+ \tag{1-34}$$

(where $\underline{V}^+$ denotes the complex conjugate of $\underline{V}$).

Using this notation, we can express the expected output power of the array for a single sinusoidal input as:

$$E[y^2] = \sigma_s^2 \underline{W}^+ \underline{V}\ \underline{V}^+ \underline{W} \tag{1-35}$$

for any steering vector $\underline{V}$, signal power σ_s^2, and weight vector $\underline{W}$.

Now, if adaptation is not allowed and the weight vector is set to an initial value $\underline{W}_Q$, then the expected output power is:

$$E[y^2]\Big|_{\underline{W}=\underline{W}_Q} = \sigma_s^2 \underline{W}_Q^+ \underline{V}\ \underline{V}^+ \underline{W}_Q \tag{1-36}$$

Consider next the expected output power at convergence. We have shown that the weight vector at convergence is:

$$\underline{W}^* = \left(\frac{\underline{R}}{\gamma} + \underline{I}\right)^{-1} \underline{W}_Q \tag{1-37}$$

Thus, we have

$$\underline{W}^* = \left(\frac{\sigma_s^2}{\gamma} \underline{V}\ \underline{V}^+ + \underline{I}\right)^{-1} \underline{W}_Q \tag{1-38}$$

By applying the matrix inversion lemma [Theorem 5.22, reference 7] we can obtain

$$\left(\frac{\sigma_s^2}{\gamma}\underline{V}\ \underline{V}^+ + \underline{I}\right)^{-1} = \underline{I}^{-1} - \frac{\underline{I}^{-1}\ \frac{\sigma_s^2}{\gamma}\underline{V}\ \underline{V}^+\underline{I}^{-1}}{1 + \frac{\sigma_s^2}{\gamma}\underline{V}^+\underline{I}^{-1}\underline{V}}$$

$$= \underline{I} - \frac{\frac{\sigma_s^2}{\gamma}\underline{V}\ \underline{V}^+}{1 + \frac{\sigma_s^2}{\gamma}\underline{V}^+\underline{V}}$$

1-39

Now, from the definition of $\underline{V}$, we can see that

$$\underline{V}^+\underline{V} = ||\underline{V}||^2 = \frac{nk}{2}$$

1-40

(recalling that k is the number of antenna elements, and n/2 is the number of taps in each tapped delay line (complex samples)). Let us define N = nk/2. Then we can use these formulas to give us a converged weight vector of

$$\underline{W}^* = \left[\underline{I} - \frac{\frac{\sigma_s^2}{\gamma}\underline{V}\ \underline{V}^+}{1 + \frac{\sigma_s^2}{\gamma}\ N}\right]\underline{W}_Q$$

1-41

Now we may compute the expected power output at convergence (using $\underline{W}^*$ in 1-35):

$$E[y^2]\Big|_{\underline{W}=\underline{W}^*} = \sigma_s^2\underline{W}^{*+}\underline{V}\ \underline{V}^+\underline{W}^*$$

$$= \sigma_s^2 \underline{W}_Q^+ \left[\underline{I} - \frac{\frac{\sigma_s^2}{\gamma} \underline{V}\ \underline{V}^+}{1 + \frac{\sigma_s^2}{\gamma} N} \right] \underline{V}\ \underline{V}^+ \left[\underline{I} - \frac{\frac{\sigma_s^2}{\gamma} \underline{V}\ \underline{V}^+}{1 + \frac{\sigma_s^2}{\gamma} N} \right] \underline{W}_Q$$

$$= \sigma_s^2 \underline{W}_Q^+ \left[\underline{V}\ \underline{V}^+ - 2 \frac{\frac{\sigma_s^2}{\gamma} \underline{V}\ \underline{V}^+ \underline{V}\ \underline{V}^+}{1 + \frac{\sigma_s^2}{\gamma} N} + \frac{\frac{\sigma_s^4}{\gamma^2} \underline{V}\ \underline{V}^+ \underline{V}\ \underline{V}^+ \underline{V}\ \underline{V}^+}{(1 + \frac{\sigma_s^2}{\gamma} N)^2} \right] \underline{W}_Q \tag{1-42}$$

Using (1-40) again gives us:

$$E[y^2]_{\underline{W}=\underline{W}^*} = \sigma_s^2 \underline{W}_Q^+ \left[\underline{V}\ \underline{V}^+ - 2 \frac{\frac{\sigma_s^2}{\gamma} N}{1 + \frac{\sigma_s^2}{\gamma} N} \underline{V}\ \underline{V}^+ + \frac{\frac{\sigma_s^4}{\gamma^2} N^2}{(1 + \frac{\sigma_s^2}{\gamma} N)^2} \underline{V}\ \underline{V}^+ \right] \underline{W}_Q$$

$$= \sigma_s^2 \underline{W}_Q^+ \underline{V} \left[1 - 2 \frac{\frac{\sigma_s^2}{\gamma} N}{1 + \frac{\sigma_s^2}{\gamma} N} + \frac{\frac{\sigma_s^4}{\gamma^2} N^2}{(1 + \frac{\sigma_s^2}{\gamma} N)^2} \right] \underline{V}^+ \underline{W}_Q$$

$$= \sigma_s^2 \left[1 - \frac{\frac{\sigma_s^2}{\gamma} N}{1 + \frac{\sigma_s^2}{\gamma} N} \right]^2 \underline{W}_Q^+ \underline{V}\ \underline{V}^+ \underline{W}_Q$$

$$= \sigma_s^2 \left[\frac{1}{1 + \frac{\sigma_s^2}{\gamma} N} \right]^2 \underline{W}_Q^+ \underline{V}\ \underline{V}^+ \underline{W}_Q \tag{1-43}$$

This can be rewritten (using 1-36) as:

$$E[y^2]\Big|_{\underline{W}=\underline{W}^*} = \left[\frac{1}{1 + \frac{\sigma_s^2}{\gamma} N}\right]^2 E[y^2]\Big|_{\underline{W}=\underline{W}_Q} \qquad 1\text{-}44$$

Thus we see that the expected output power of the array at convergence is the expected output power of the array initially, multiplied by an attenuation factor. This factor is always less than (or at most equal to) one since $\frac{\sigma_s^2}{\gamma} N \geq 0$. This attenuation factor is independent of the antenna geometry, arrival direction of the sinusoid, and of the quiescent weight vector. The attenuation factor depends only on the sinusoid's power (σ_s^2), the total number of taps in the array filters (N), and the equivalent pilot noise power (γ). Thus, if we know the response of the quiescent unadapted array to a single sinusoid, we can easily calculate its response at convergence of the adaptive process.

1-E. Conclusions of Part 1

We have proposed and studied a new adaptive algorithm which has been designated Scheme 6A. It is related to the leaky LMS algorithm previously studied. When applied to an antenna array, it yields a method of antijamming based on attenuation of received signals on the basis of their input power levels. No a priori knowledge of the signal characteristics is required. This algorithm replaces a previously studied algorithm, Scheme 6. It exhibits improved noise behavior and requires less hardware for implementation.

PART 2

ANALYSIS OF ADAPTIVE WEIGHT NOISE COVARIANCE

2-A. Introduction

Previous work [1] introduced the concept of using an adaptive line enhancer [8, 9] to separate signals by power level. Although the original proposal required injected noise and 'slave' filters, a refined version was developed that eliminated the need for injected noise by replacing the LMS adaptive filter with a 'leaky' LMS filter. Because an adaptive line enhancer using the 'leaky' LMS algorithm can separate signals by power level it is called an adaptive power separator (APS).

During the past year we have been engaged in analysing the performance of the APS. Performance in the mean has already been described in [1] ; however, performance in the mean does not completely characterise an adaptive filter. Specifically it is important to know about the noise in the weights since, by a modulation process, weight noise causes noise components in the filter output [9, 10] which degrades the performance of the APS. The first step is to characterise the variance of the noise in the weights.

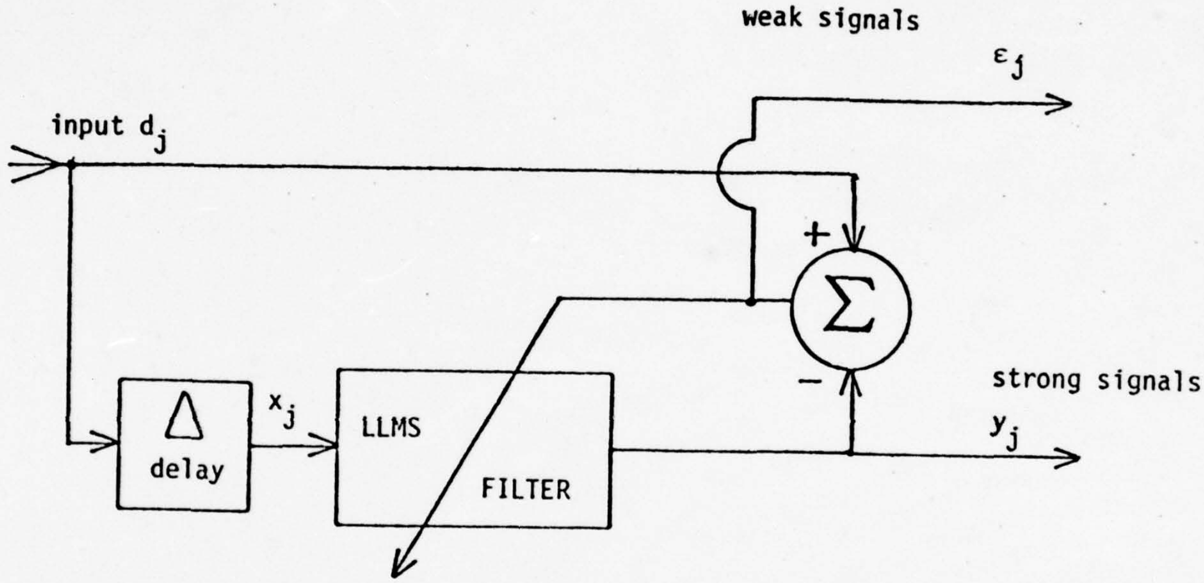
Analysis of weight noise covariance has proceeded in four main phases:

1) a deterministic analysis with the input consisting of a sinusoid
2) a statistical analysis with the input consisting of a sinusiod

3) statistical analysis with the input consisting of only noise and then approximate extensions to sinusoids in noise

4) exact analysis of a single weight APS for comparison with the previous approximate analyses.

Figure 2-1 depicts the APS and helps to define the quantities



ADAPTIVE POWER SEPARATOR

Figure 2-1

used throughout this section:

d_j is the system input at time j

$x_j = d_{j-\Delta}$ is the filter input at time j

y_j is the filter output at time j

ϵ_j is the 'error' signal at time j. This signal is used to modify, or update, the filter weights according to the 'leaky' LMS algorithm.

The operation of the 'leaky' LMS (LLMS) filter is defined as follows:

If L= length of filter (number of weights)

W= weight vector $=[w_1,w_2,\ldots,w_L]^T$

S= state vector of filter $=[x_j,x_{j-1},\ldots,x_{j-L+1}]^T$

then

$$y_j=W^TS = \sum_{i=1}^{L} w_i x_{j-i+1} \tag{2-1}$$

$$\epsilon_j = d_j-y_j \tag{2-2}$$

$$W_{j+1} = \nu W_j+2\mu\epsilon j S_j \tag{2-3}$$

μ = a constant controlling rate of adaption

ν = the 'leak' factor (generally less than 1)

$= 1-2\mu\gamma$ (2-4)

γ = the equivalent injected noise power. That is, the effect of the leak is the same as adding noise of power γ to the input and then using a conventional LMS filter. (Note: LMS is a special case of LLMS where $\nu=1$)

To characterise the average performance of the APS we have to introduce several new quantities which define the composition of the input signal d_j:

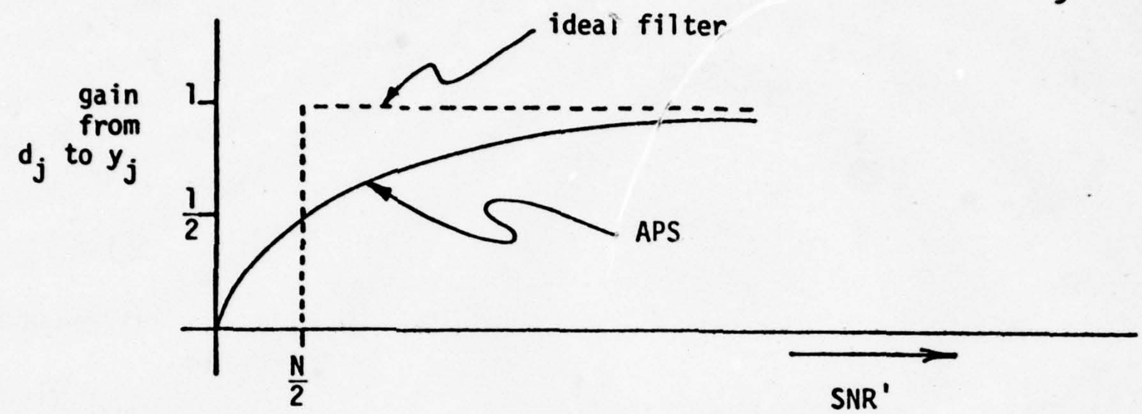
σ^2 = power of input noise

$\sigma'^2 = \sigma^2 + \gamma$ = effective power of input noise

p = power of input sinusoid

$SNR' = \frac{p}{\sigma'^2}$ = effective signal to noise ratio

With these definitions we can draw a gain curve from input, d_j, to th



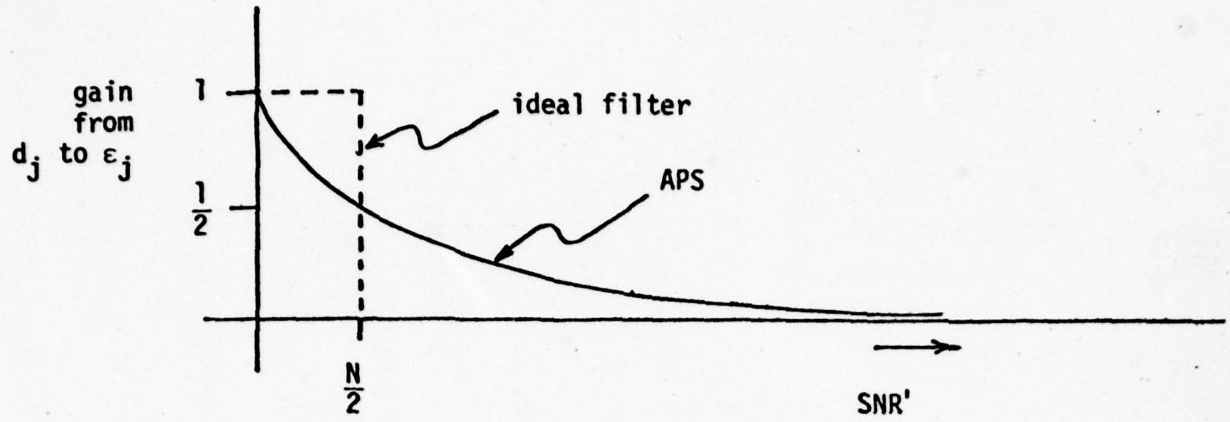
APS used to reject low powered inputs

Figure 2-2

filter output, y_j, versus SNR' (see figure 2-2). Alternatively we ca draw the gain curve from the input, d_j, to the error output, ϵ_j, versus SNR' (see figure 2-3).

Thus to discriminate against weak signals we use the y output, and to discriminate against strong signals we use the ϵ output. Not that SNR' can be varied by selection of γ since:

$$SNR' = \frac{p}{\sigma^2 + \frac{1-\gamma}{2\mu}}$$



APS used to reject high powered inputs

Figure 2-3

2-B. Deterministic Analysis of Sinusoidal Input

One case of interest is the behavior of the APS when the input consists of a sinusoid and little or no noise. This might occur in practice if an APS were used in a jam resistant mobile communication system and the base station were close to the mobile receiver, or the mobile receiver were very close to a powerful narrowband jammer. By assumption the sinusoidal component of the input is either the desired signal (say narrowband FM or AM) or a narrowband jammer -- the difference is power level.

For this case we let the delay, Δ, be one unit and so:

$$d_j = a \cos[\Theta j] \tag{2-6}$$

$$x_j = a \cos[\Theta j - \Theta] \tag{2-7}$$

$$\Theta = wT \tag{2-8}$$

w = frequency of sinusoid

T = sampling interval

From previous analysis of the mean performance [1] we know that the output will be approximately:

$$y_j = b \cos[\Theta j] \tag{2-9}$$

Substituting for d_j and y_j in equation 2-2 yields:

$$\epsilon_j = (a-b)\cos[\theta j] \tag{2-10}$$

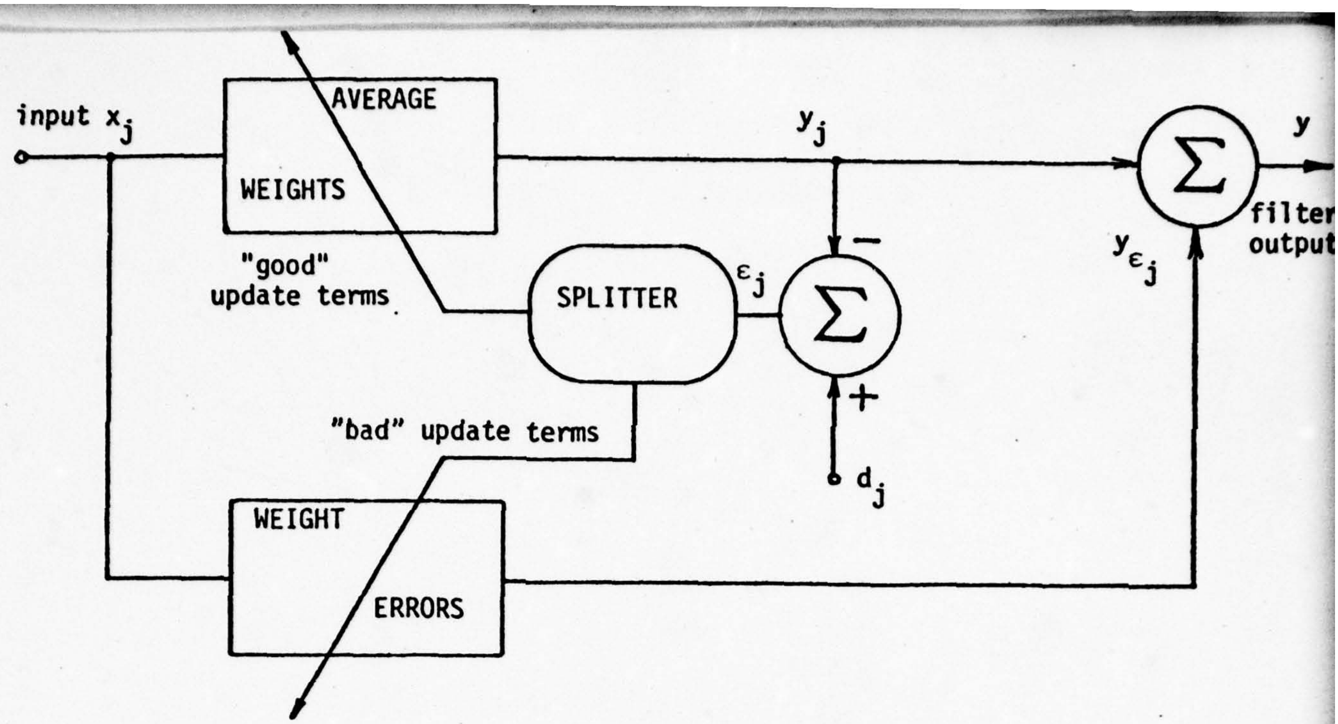
Using this we compute:

$$(\epsilon_j S_j)_i = \frac{a-b}{2}\{\cos[\theta i]+\cos[2\theta j-\theta i]\} \tag{2-11}$$

Notice that the first term of equation 2-11 is not a function of time j, only of weight index i. The second term is a function of time and weight index. In previous analysis of the APS, when the mean performance was desired, the second term was ignored on the basis that it "averages out" over a period. Hence only the first term contributes to the mean weight vector. Indeed the mean value of the weight vector can be found by using only the first term of equation 2-11. We see that the second term causes time changes in the weight vector and therefore is undesirable.

We can model the LLMS filter, which is the heart of the APS, as an "ideal" filter whose weights are fixed at the average weight vector, in parallel with a "perturbation" or "error" filter (Figure 2-4) whose weights have zero mean and vary according to the fluctuations about the average of the actual adapting weights.

The "good" update term is $\frac{a-b}{2}\cos[\theta i]$ and the "bad" term is $\frac{a-b}{2}\cos[2\theta j-\theta i]$. Both "good" and "bad" update terms are present in a filter using the LLMS algorithm. The "good" term drives the filter towards the minimum mean square error solution; the "bad" term is at twice the input frequency and causes the weights to churn, which is



Weight perturbation model of an LLMS filter

Figure 2-4

nonproductive and causes output distortion. To determine the weight variations of the weight error filter we apply the "bad" update term to the LLMS update scheme (equation 2-3).

To make the problem solvable, note that equation 2-3 can be Z-transformed to give:

$$W_{j+1} = \nu W_j + 2\mu\epsilon_j S_j$$

$$(Z-\nu)W(Z) = 2\mu\epsilon(Z)S(Z)$$

$$\frac{W(Z)}{\epsilon(Z)S(Z)} = \frac{2\mu}{Z-\nu} \tag{2-12}$$

Since the input to the system defined by equation 2-12 is a sinusoid, we know that the output will be a sinusoid of the same frequency. The complex gain of the difference equation at this frequency is:

$$g_{2\theta} = \frac{2\mu}{e^{j2\theta} - V}$$

The amplitude of the weights, W_{ϵ}, of the weight error filter will be:

$$\frac{a-b}{2} |g_{2\theta}| \tag{2-13}$$

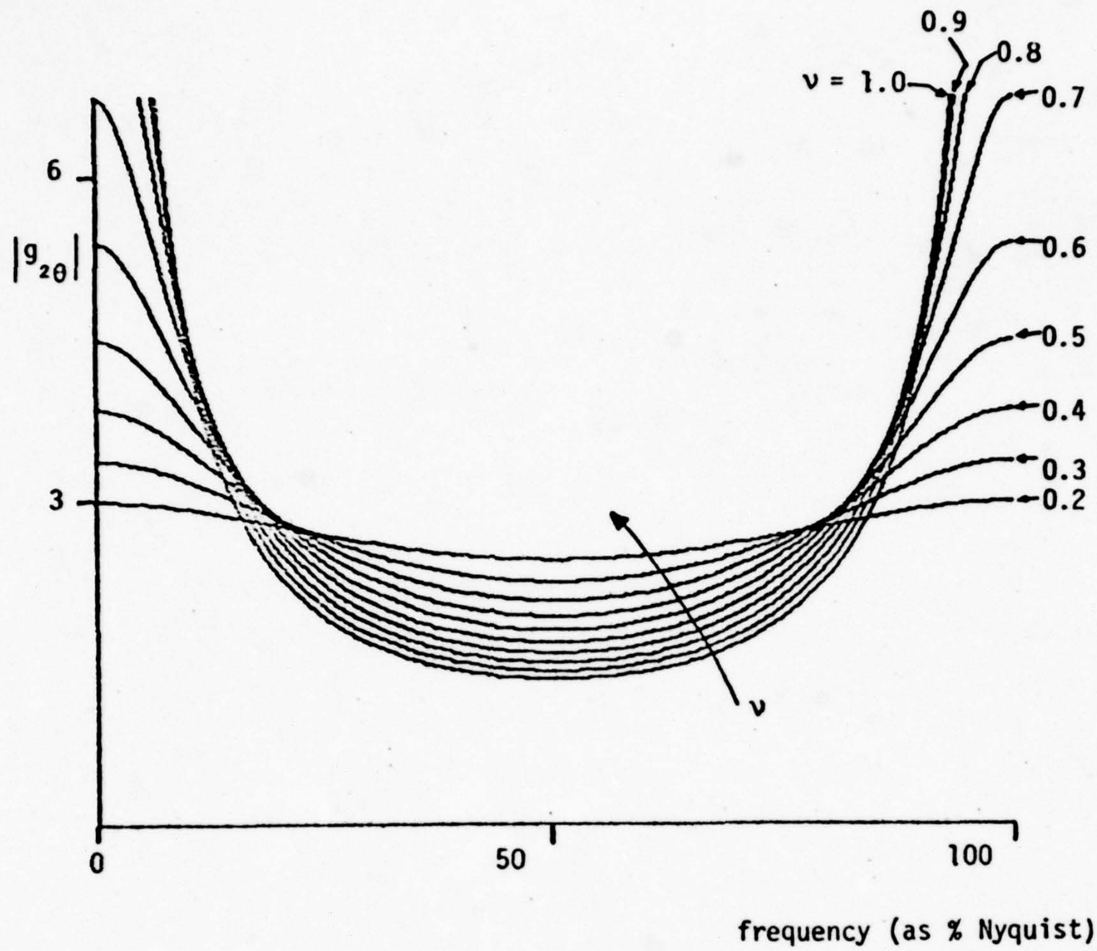
Figure 2-5 shows $|g_{2\theta}|$ vs. θ for a variety of V's. Using equation 2-1 we can compute, in the time domain, the output of the weight error filter as:

$$\begin{aligned} y_{\epsilon j} &= \sum_{i=1}^{L} w_{\epsilon i} x_{j-i-1} \\ &= \sum_{i=1}^{L} \{\frac{a-b}{2} |g_{2\theta}| \text{COS}[2\theta j - \theta i + \phi]\}\{a\text{COS}[\theta j - \theta i]\} \\ &= \frac{a(a-b)}{2} |g_{2\theta}| \sum_{i=1}^{L} \frac{1}{2}\{\text{COS}[\theta j - \phi] + \text{COS}[3\theta j - 2\theta i + \phi]\} \\ &= \frac{a(a-b)}{4} |g_{2\theta}| L \cos[\theta j - \phi] + \frac{a(a-b)}{4} |g_{2\theta}| \sum_{i=1}^{L} \cos[3\theta j - 2\theta i + \phi] \end{aligned} \tag{2-14}$$

Where $\phi = \text{ARG}(g_{2\theta}) \doteq \tan^{-1}\left[\frac{\text{Imaginary}(g)}{\text{Real}(g)}\right]$.

The first term of $y_{\epsilon j}$ is at the same frequency as the desired output and has amplitude $\text{amp}_1 = \frac{a(a-b)}{4} |g_{2\theta}| L$. We note that $b \leq a$, so $a-b \leq a$ which implies $\text{amp}_1 \leq \frac{a^2 L}{4} |g_{2\theta}|$, but this can be made arbitrarily small by selection of μ which controls $|g_{2\theta}|$ unless θ is very small. If we restrict our attention to a reasonable range of frequencies, say 20-80% of Nyquist (see also [8] where Treichler makes this same assumption in analysing the mean performance), then:

$$\text{amp}_1 \leq \frac{a^2 L}{4} |g_{2\theta}| \leq \mu \frac{3a^2 L}{4} \qquad \forall V \tag{2-15}$$



Gain vs. Frequency
as a function of leak

Figure 2-5

The second term of 2-14 will cause an output at three times the original frequency. While the expression for the amplitude is rather complicated we can easily bound it:

$$\sum_{i=1}^{L} \cos[3\theta j-2\theta i+\phi] \leq L$$

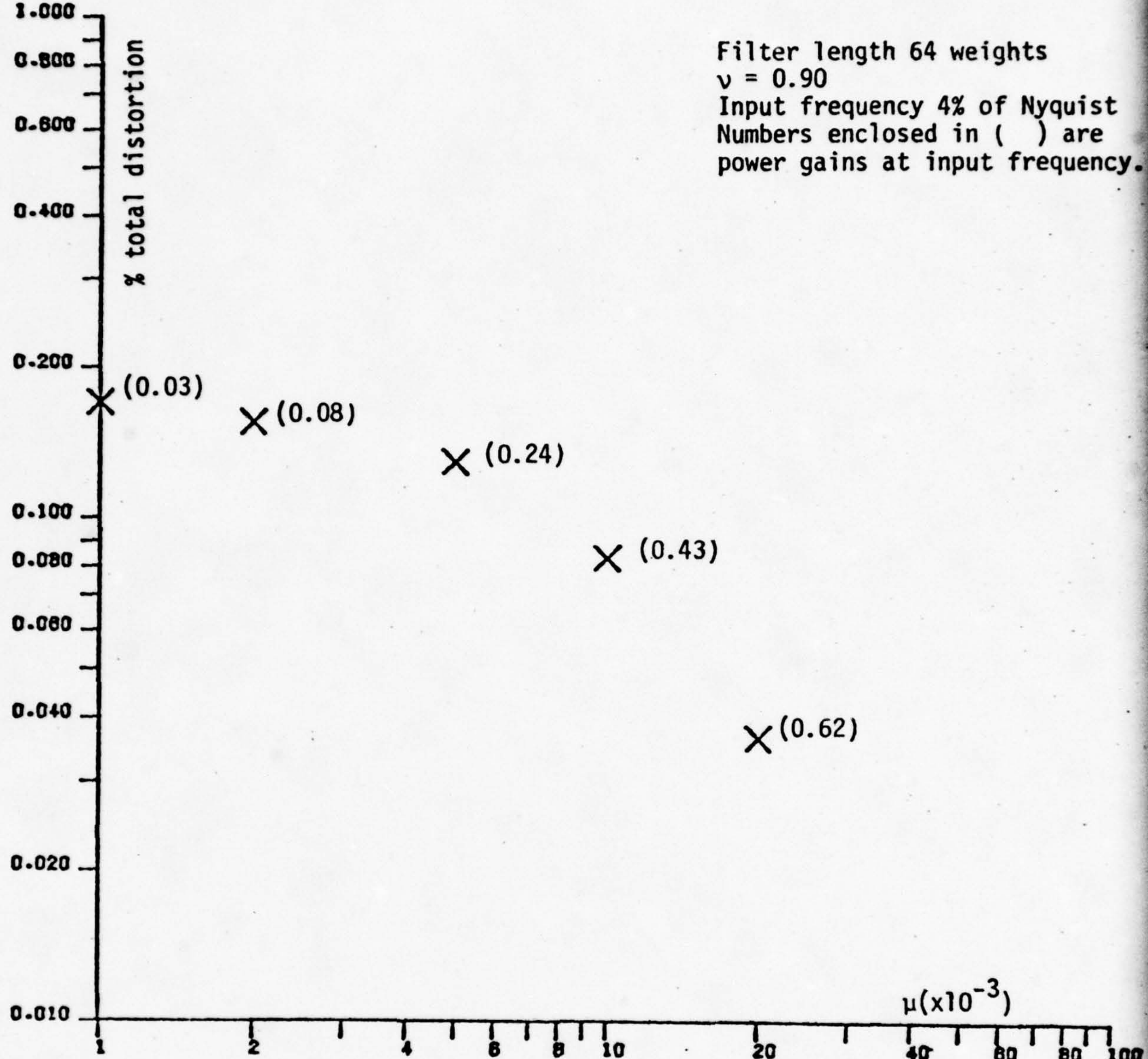
from which we see that the amplitude of this third harmonic distortion is also bounded by equation 2-15.

Looking at the filter model (figure 2-4), we notice that the output of the weight error filter, which is part of the total filter output, was not subtracted from the desired signal d_j. If y_ϵ were subtracted from d_j as well as y_i, it would cause a new component in the error signal, ϵ_j, equal to $-y_\epsilon$. To account for this, another weight error filter should be added. This filter is updated using $W_{j+1} = V\, W_j - 2\mu y_\epsilon S_j$ and causes a new output $y_{\epsilon 2}$ which will have terms at the fundamental, third harmonic and fifth harmonic. These outputs contribute to ϵ_j and so on. However, note that y_ϵ can be made arbitrarily small, and $y_{\epsilon 2} \leq \frac{a y_\epsilon}{4}\{|g_{2\theta}|+|g_{4\theta}|\} \leq \frac{a^3 L}{16}\{|g_{2\theta}|+|g_{4\theta}\}$. The weakness of this analysis technique becomes apparent here. Previously we restricted $|g_{2\theta}|$ by selecting a reasonable range of values for θ. Similiarly we could select a range of θ to restrict $|g_{4\theta}|$. But, it is evident that $y_{\epsilon 2}$ will give rise to other error terms which will cascade to create an infinite number of error terms of all different frequencies. Some of these error terms must surely have very large $|g|$ since they will be near some multiple of the sampling frequency. However, experimental results indicate that if θ is between 20% and 80% of Nyquist frequency, then the only output component of any significance is the third harmonic. Hence for practical purposes only the first "weight error" filter needs to be analysed.

Computer simulations were performed to verify these analytic predictions. The frequency of the input sinusoid was set at 4% of Nyquist. The results are graphed in figures 2-6 through 2-11. Each plot shows percent power not at the fundamental versus μ. Thus each plot relates total distortion power to μ. From 2-6 we see that in all cases the total distortion power was less than 0.2% of the output power at the fundamental frequency. From 2-7 we get the same shape of curve but now the total distortion is less than 0.06% of the desired output power. In 2-8 the fundamental frequency has been increased to 8% of Nyquist. The error power is substantially lower than for the previous cases, as is to be expected.

In an attempt to stress the APS to cause significant distortion, three simulations were run in which the fundamental frequency was 1% of Nyquist as opposed to being in the 20-80% range. The results are plotted in figures 2-9,2-10 and 2-11. Note that the greatest observed distortion power is 4% of the power of the desired signal.

A final point to consider is at what frequencies the distortion power occurred. Analytic efforts indicate that the 3^{rd} harmonic should be dominant. In fact, the simulations supported this. Even when the fundamental frequency was 1% of Nyquist, over 98% of the distortion power was in the 3^{rd} harmonic.



NOTE: Rate of convergence is proportional to μ

Figure 2-6

% distortion vs. μ

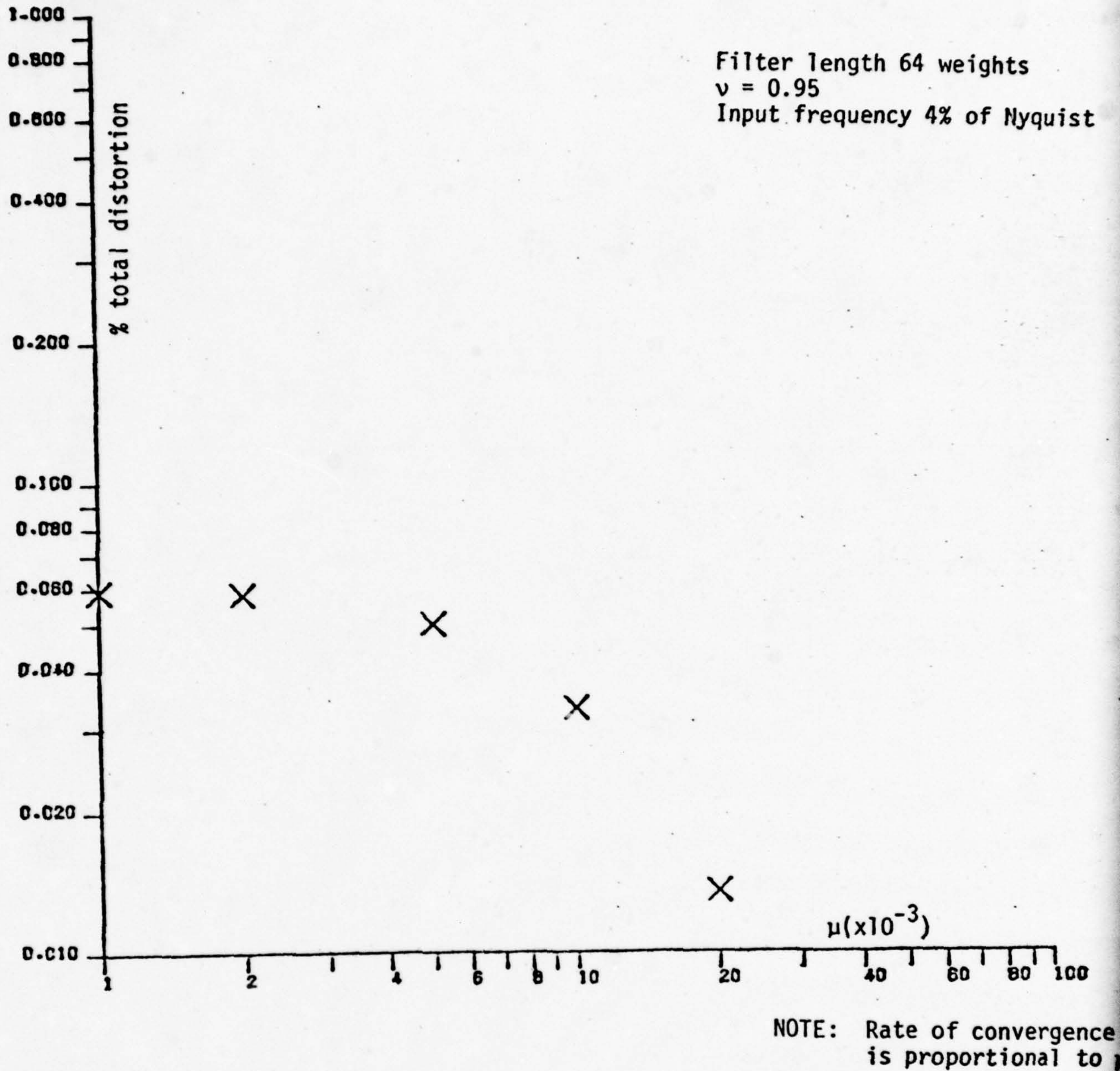
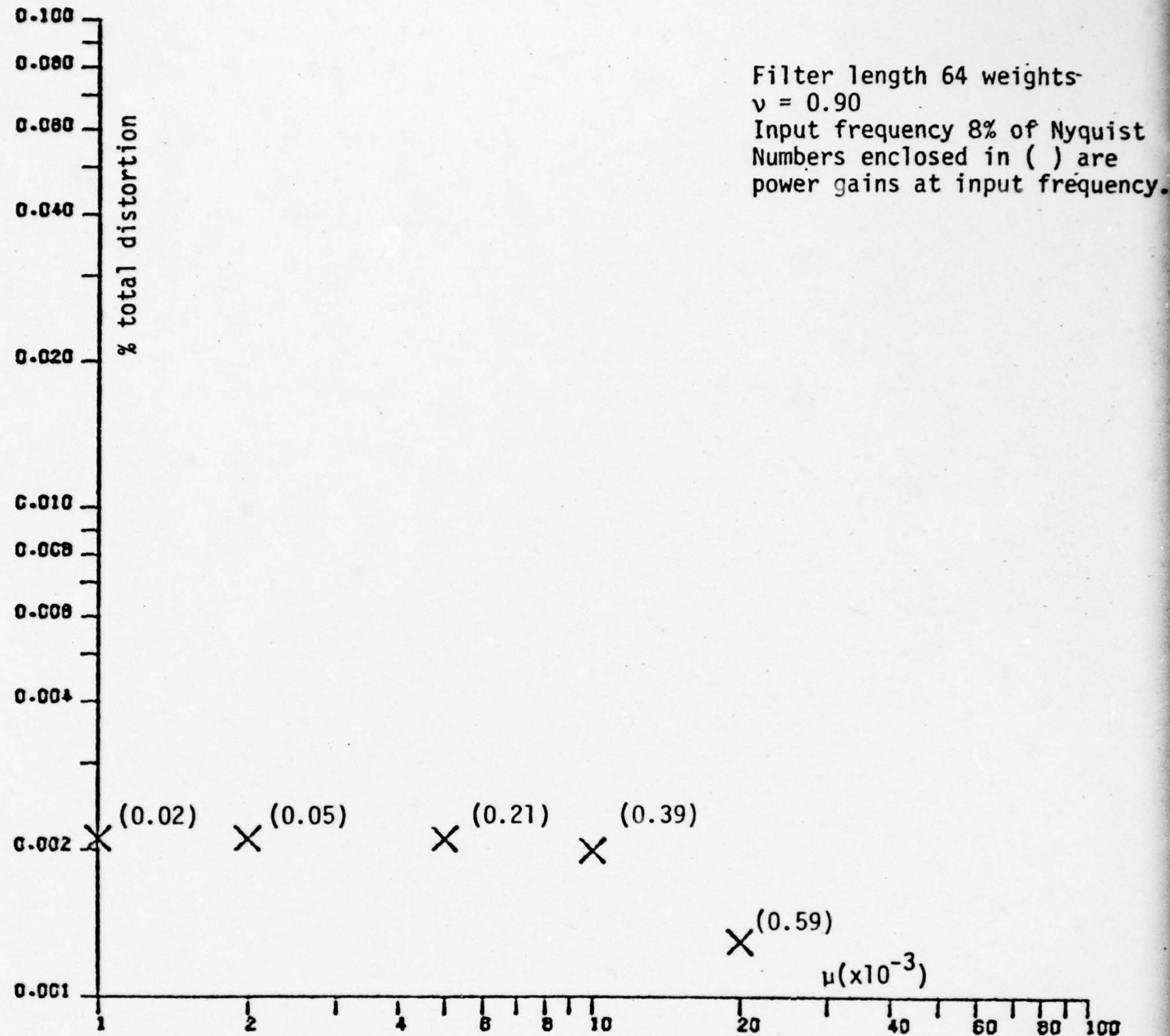


Figure 2-7

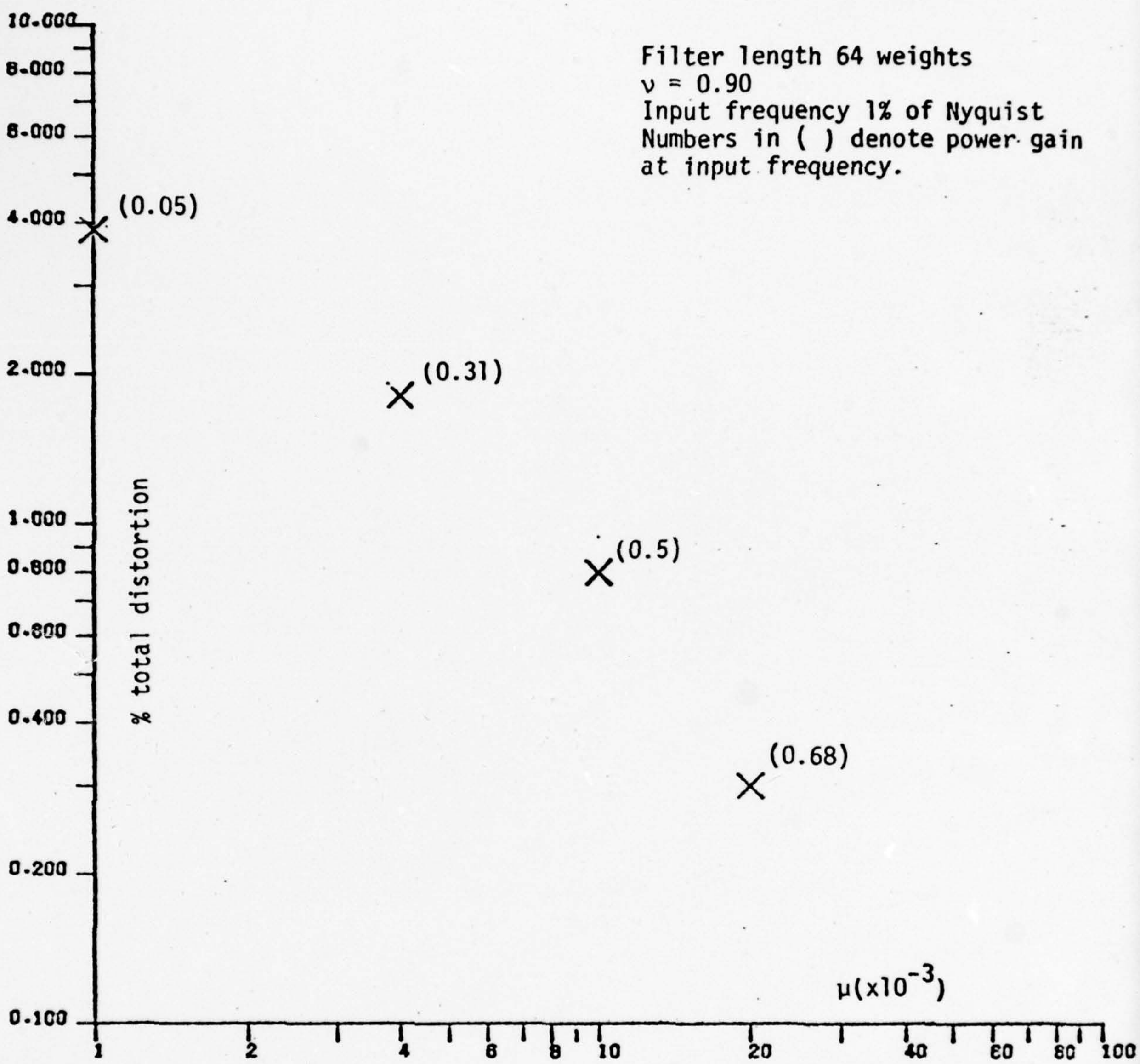
% distortion vs. μ



NOTE: Rate of convergence is proportional to μ.

Figure 2-8

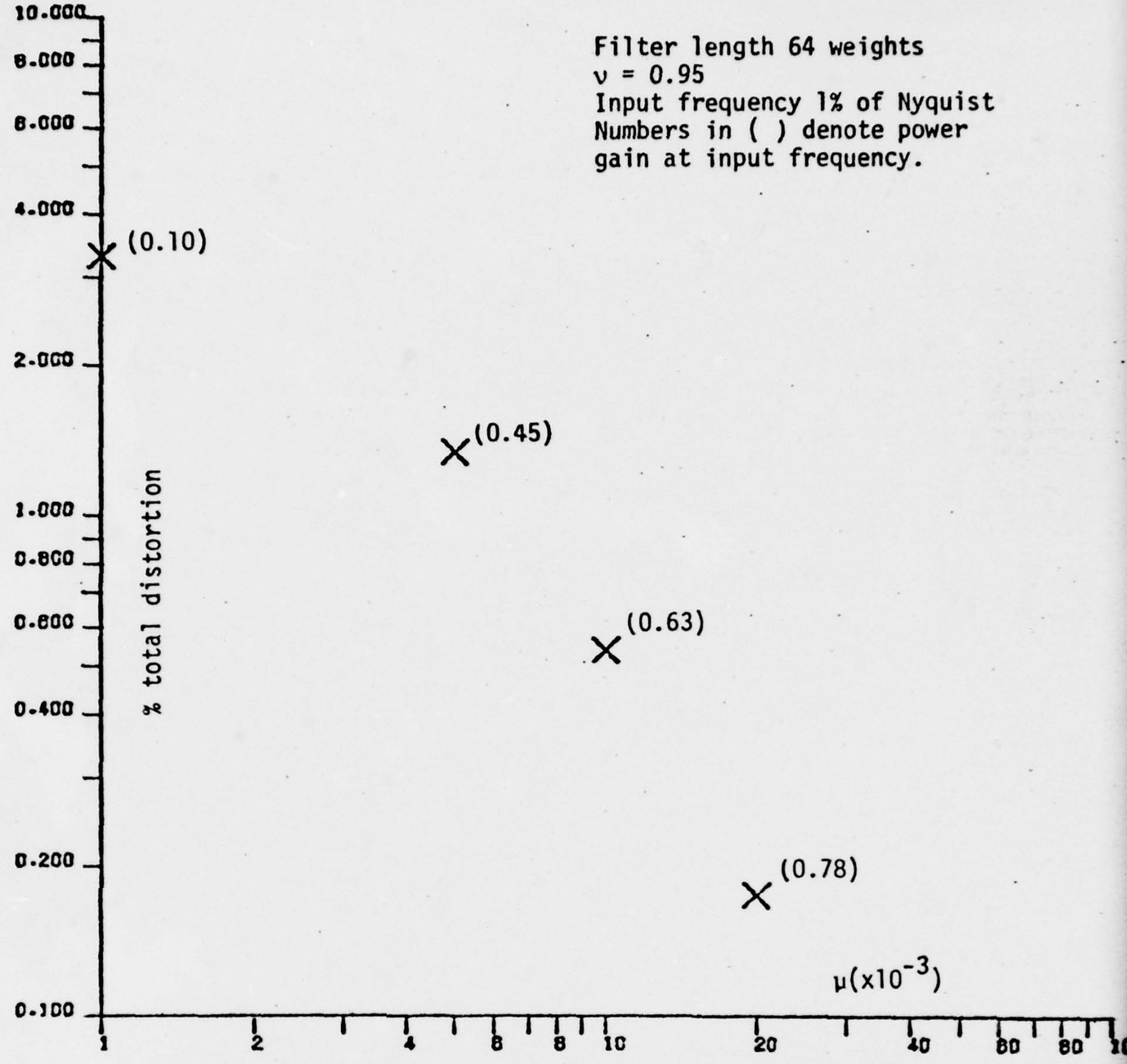
% distortion vs. μ



NOTE: Rate of convergence is proportional to μ.

Figure 2-9

% distortion vs. μ



NOTE: Rate of convergence is proportional to μ.

Figure 2-10

% distortion vs. μ

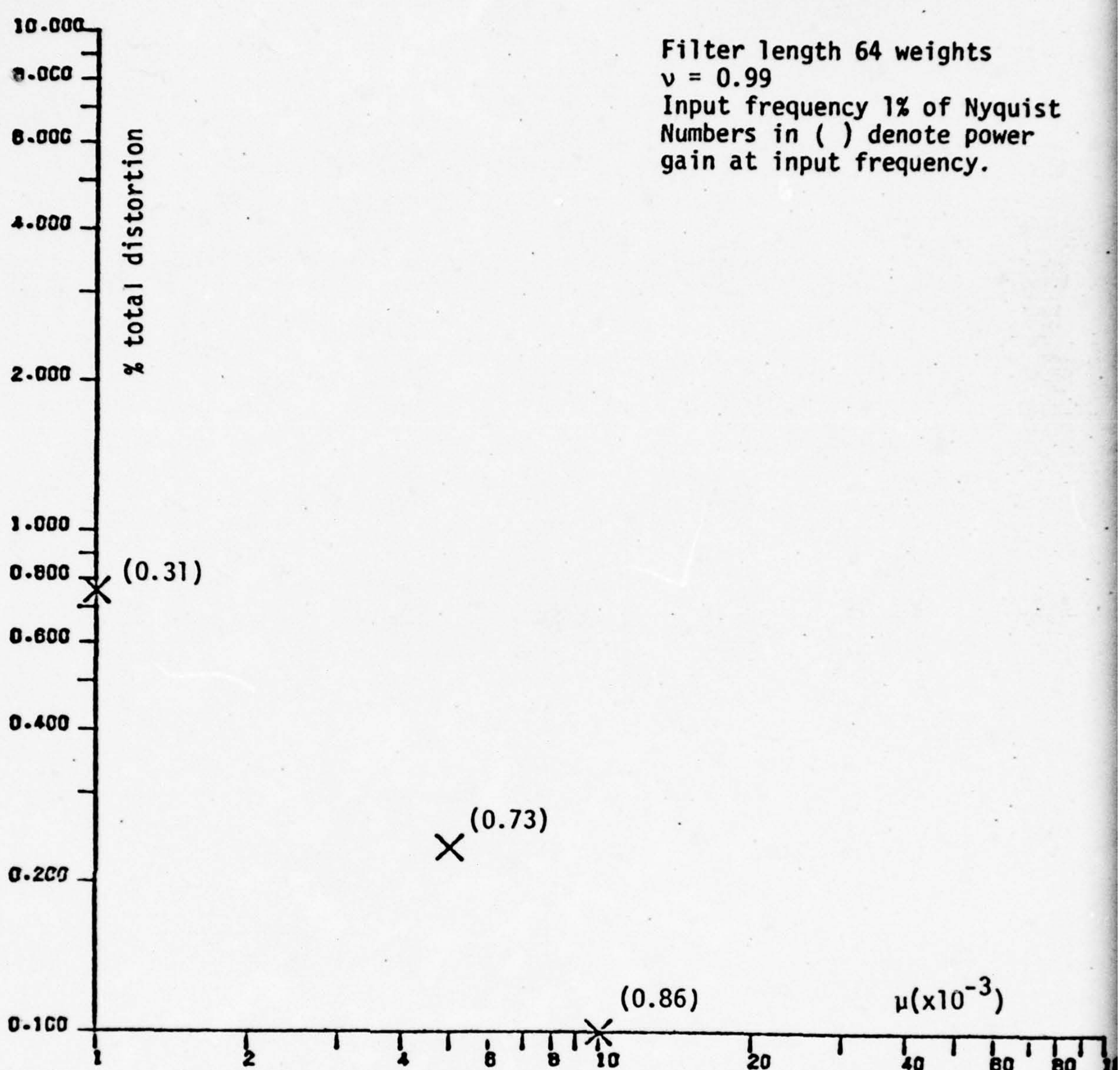


Figure 2-11

% distortion vs. μ

2-C. Statistical Analysis of Sinusoidal Input

From previous work [1] the general form of tha APS output is known. If we use equations 2-6,7 & 8 to define the input, then we know that the output will be approximately:

$$y_j = b\cos(\theta j) = \frac{a}{1+\frac{2}{L\ SNR'}}\cos[\theta j]$$

2-1

If we define ΔW_j as $W_j - E\{W\}$

2-2

and then notice that $(1-\gamma)E\{W\} = E\{2\mu\epsilon S\}$,

so by taking the expectation of 2-3 and 2-18 we get:

$$\Delta W = \gamma\Delta W_j + 2\mu\epsilon_j S_j - E\{2\mu\epsilon_j S_j\}$$

2-3

hence

$$\frac{\Delta W(Z)}{2\mu\epsilon(Z)S(Z)} = \frac{1}{z-\gamma}$$

but if we know $VAR\{2\mu\epsilon_j S_j\}$ then we know

$$VAR\{\Delta W_j\} = \left|\frac{1}{z-\gamma}\right|^2 VAR\{2\mu\epsilon_j S_j\}$$

2-4

$$VAR\{2\mu\epsilon_j S_j\} = VAR\{2\mu[a\cos(\theta j) - \frac{a}{1+\frac{2}{L\ SNR'}}\cos(\theta j)]S_j\}$$

$$= 4\mu^2 VAR\{\frac{a}{\frac{L\ SNR'}{2}+1}\cos(\theta j)S_j\}$$

$$(VAR\{2\mu\epsilon_j S_j\})_i = 4\mu^2\frac{a^2}{(1+\frac{L\ SNR'}{2})^2}\ VAR\{(\cos[\theta j])(a\cos[\theta j-\theta i])\}$$

$$= \frac{4\mu^2 a^2}{(1+\frac{L\ SNR'}{2})^2}\ \frac{a^2}{8} \quad = \frac{\mu^2 a^4}{2(1+\frac{L\ SNR'}{2})^2} \qquad 2\text{-}5$$

Thus

$$(VAR\{\Delta W_j\})_i = \left|\frac{1}{z-V}\right|^2 \frac{\mu^2 a^4}{2(1+\frac{L\ SNR'}{2})^2}$$

$$= \left|\frac{1}{e^{j2\Theta}-V}\right|^2 \frac{\mu^2 a^4}{2(1+\frac{L\ SNR'}{2})^2} \qquad 2\text{-}6$$

These formulae are in close agreement with experimental results as Table 2-1 shows. In the cases where w is 12.5% of Nyquist, there is a 15% error in determining $VAR\{2\mu\epsilon_j S_j\}$, although $VAR\{\Delta W\}$ is still given by equation 2-19 if we use the observed value for $VAR\{2\mu\epsilon_j S_j\}$.

To summarise we used a simple statistical analysis to determine the variance of the driving function, $2\mu\epsilon_j S_j$ (equation 2-21). Then we considered how this affected the weight vector (equation 2-19) and derived a relationship (equation 2-22) for the weight variance in terms of the input quantities. These equations (2-21 & 22) will give a good idea of the weight variance (note $VAR\{W_j\} = VAR\{\Delta W_j\}$) providing the input frequency is not to near DC. A reasonable range of input frequencies appears to be from 20% to 80% of Nyquist.

μ	ν	ω	CVAR	MVAR	% error
10^{-3}	10^{-3}	25.0	4.54×10^{-9}	4.55×10^{-9}	0.2
10^{-4}	10^{-4}	25.0	4.54×10^{-11}	4.55×10^{-11}	0.2
10^{-4}	10^{-2}	25.0	4.54×10^{-9}	1.38×10^{-8}	39.6
10^{-3}	10^{-3}	12.5	4.54×10^{-9}	5.25×10^{-9}	15.6
10^{-4}	10^{-4}	12.5	4.54×10^{-11}	5.22×10^{-11}	14.9
10^{-5}	10^{-5}	12.5	4.54×10^{-13}	1.39×10^{-12}	69.4*
10^{-3}	10^{-3}	20.0	4.54×10^{-9}	4.55×10^{-9}	0.9
10^{-3}	10^{-3}	15.0	4.54×10^{-9}	4.54×10^{-9}	0.0

*This value almost certainly incorrect due to numerical inaccuracies in the computer simulation.

ω is measured in % of Nyquist frequency

CVAR is the calculated variance

MVAR is the measured variance

$$\% \text{ error} = \frac{\text{MVAR-CVAR}}{\text{CVAR}} \times 100\%$$

Table 2-1

2-D. Statistical Aalysis of the Adaptive Power Separator (APS) with Noise Input

In a real communication environment the input to an APS might well consist of both sinusoidal components (either signal or narrowband jammers) and broadband noise (thermal noise or wideband jammers). Thus we wish to understand the behavior of the APS to a sinusoid in broadband noise. The first step is to characterise the weight vector variance when the input consists only of noise.

Note that:

$$VAR\{X\} = E\{XX^T\}-E\{X\}E\{X\}^T \tag{2-7}$$

also

$$d_j = n_j$$

$$x_j = n_{j-1}$$

where

n_j is from a white, zero-mean Gaussian process with variance σ^2. 2-8

From equations 2-2,3 & 4

$$W_{j+1} = (I-2\mu\gamma I-2\mu SS^T)W_j + 2\mu d_j S_j \tag{2-9}$$

If we assume $E\{SW^T\} = 0$, which is a common assumption (well supported by experience if μ is small enough), then we find:

$$E\{W_{j+1}\} = \overline{W} = (I-2\mu\gamma I-2\mu\sigma^2 I)\overline{W}_j + 0$$

2-10

$$\begin{aligned}
VAR\{W_{j+1}\} &= E\{W_{j+1}W_{j+1}^T\} - \overline{W}_{j+1}\overline{W}_{j+1}^T \\
&= (1-2\mu\gamma)^2 VAR\{W_j\}-4\mu(1-2\mu\gamma)\sigma^2 VAR\{W_j\}+4\mu^2\sigma^4 I \\
&\quad + 4\mu^2 E\{SS^T VAR\{W_j\}SS^T\} \\
&= (1-4\mu\gamma+4\mu^2\gamma^2-4\mu\sigma^2+8\mu^2\gamma\sigma^2)VAR\{W_j\}+4\mu^2\sigma^4 I+0
\end{aligned}$$

2-11

Where we have assumed:

$$\mu E\{SS^T VAR\{W_j\}SS^T\} << \sigma^2 VAR\{W_j\}$$

2-12

If we now assume steady state so that $VAR\{W_j\} = VAR\{W_{j+1}\}$, then:

$$\begin{aligned}
VAR\{W\} &= \frac{4\mu^2\sigma^4}{4\mu\sigma^2+4\mu\gamma+4\mu^2\gamma^2-8\mu^2\gamma\sigma^2} I \\
&= \frac{\mu\sigma^2}{1+\frac{\gamma}{\sigma^2}-\mu\frac{\gamma^2}{\sigma^2}-2\mu\gamma} I \\
&= \frac{\mu\sigma^2}{1+\frac{\gamma}{\sigma^2}} I
\end{aligned}$$

2-13

Providing $\mu\gamma << 1+\frac{\gamma}{\sigma^2}$ and $\mu\frac{\gamma^2}{\sigma^2} << 1+\frac{\gamma}{\sigma^2}$

$$\Rightarrow \mu << \frac{1}{\gamma}+\frac{1}{\sigma^2} \qquad \text{and} \qquad \mu << \frac{1}{\gamma}+\frac{\sigma^2}{\gamma^2}$$

2-14

Notice that for LMS $\gamma = 0$, and so $VAR\{W\} = \mu\sigma^2 I$ which corresponds to earlier analysis [9] of the LMS filter. Also note that the 'leaky' algorithm never increases the weight variance, in fact it actually reduces the variance. This supports the conjecture [1] that algorithmically simulated noise (via 'leaky' LMS) is more desirable than actual noise injection, even though the converged mean weight vector solution is the same in both cases.

In equation 2-28 we assumed that

$$\mu E\{SS^T VAR\{W\} SS^T\} << \sigma^2 VAR\{W\}$$

To check this assumption we note that:

$$E\{SS^T VAR\{W\} SS^T\} = \sigma^4 VAR\{W\} + \sigma^4 Diag[VAR\{W\}] + \sigma^4 Trace[VAR\{W\}] I \qquad \text{2-15}$$

where Diag[.] means the matrix consisting only of the diagonal elements of the operand.

Assuming that 2-29 is substantially correct leads to the following conclusion:

$$E\{SS^T VAR\{W\} SS^T\} = \frac{2\mu\sigma^6}{1+\frac{\gamma}{\sigma^2}} + \frac{L\mu\sigma^6}{1+\frac{\gamma}{\sigma^2}} I = \frac{\mu(L+2)\sigma^6}{1+\frac{\gamma}{\sigma^2}} I \qquad \text{2-16}$$

Therefore 2-28 requires:

$$\frac{\mu^2 (L+2)\sigma^6}{1+\frac{\gamma}{\sigma^2}} << \sigma^2 \frac{\mu\sigma^2}{1+\frac{\gamma}{\sigma^2}}$$

$$\text{which implies} \quad \mu << \frac{1}{\sigma^2} \tag{2-17}$$

This condition is often assumed in adaptive filtering and it implies slow, low-noise adaptation.

Extension to the case of a sinusoid in noise is unfortunately not obvious. Intuitively it is appealing to argue that the correct formula should be:

$$\mathrm{VAR}\{W\} = \frac{\mu\xi_{min}}{1+\frac{\gamma}{\xi_{min}}} I$$

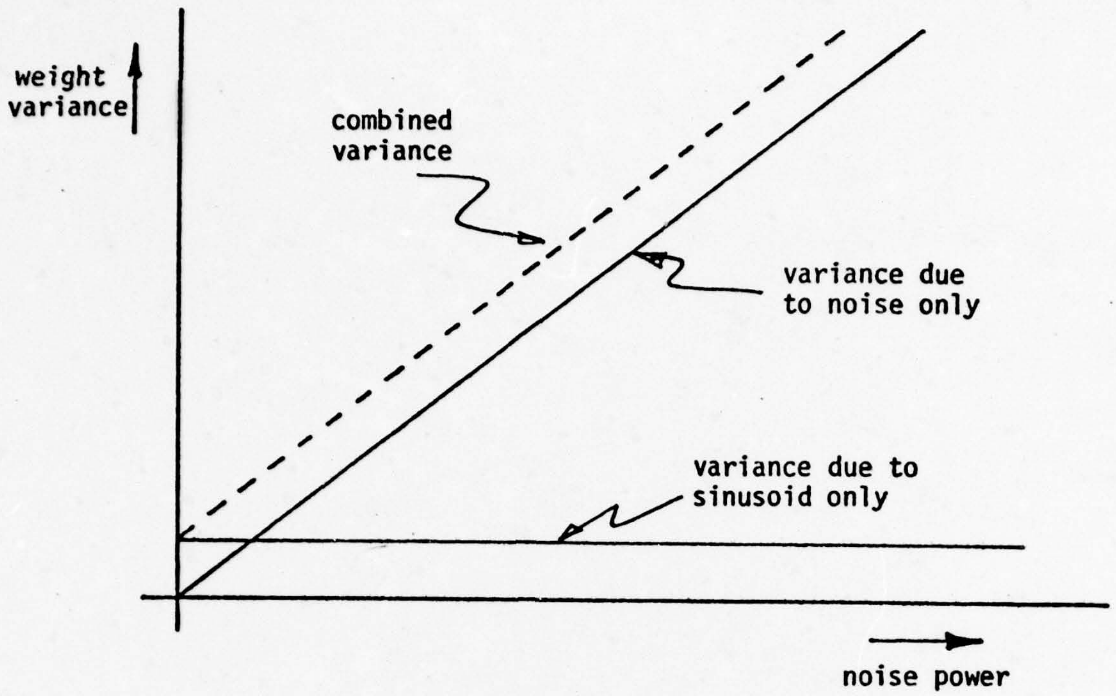
by analogy with the conventional LMS formula of:

$$\mathrm{VAR}\{W\} = \mu\xi_{min} I$$

Experiments indicate that this is not the case, in fact it is easy to see that this would not be true since in the no noise case the formula is substantially incorrect.

Another approximation assumes that the noise and sinusoid affect the weights independently, which results in the curve of figure 2-12.

This approach is not strictly correct either. However, if the weight variance due to noise alone is substantially greater than the variance due to the sinusoid alone, then the formula is indicative of the actual variance. This result has been checked experimentally, and some results are presented in Table 2-2. An important point to consider is that computing weight variance by using the formula:



Weight variance obtained by assuming independent effects of noise and sinusoid

Figure 2-12

$$\mathrm{VAR}\{W\} = \mu(\text{error-output-power})$$

or

$$\mathrm{VAR}\{W\} = \frac{\mu(\text{error-output-power})}{1+\dfrac{\gamma}{(\text{error-output-power})}}$$

(as was done in [10] for example) is apparently very conservative. The power due to sinusoidal components does not contribute to weight variance as much, proportionally, as power due to random components. We are currently working to quantify this effect.

noise power	$1-\gamma$	μ	weight variance	approx. #1	approx. #2
1.17×10^{-3}	10^{-3}	10^{-3}	3.27×10^{-8}	9.31×10^{-6}	1.37×10^{-9}
1.0	2×10^{-3}	5×10^{-4}	1.75×10^{-4}	3.27×10^{-4}	2.5×10^{-4}
1.0	10^{-3}	5×10^{-5}	5.67×10^{-6}	3.27×10^{-4}	2.5×10^{-5}
1.0	10^{-3}	5×10^{-4}	2.56×10^{-4}	3.27×10^{-4}	2.5×10^{-4}
1.0	10^{-3}	10^{-6}	5.81×10^{-9}	6.55×10^{-7}	5.0×10^{-7}
5×10^{-4}	10^{-3}	10^{-3}	1.29×10^{-8}	9.18×10^{-6}	2.5×10^{-10}

Frequency: 25% of Nyquist
Power of sinusoid: 1.0
Number of weights: 20

$$\text{approx. \#1 VAR} = \frac{\mu \xi^2_{min}}{1 + \dfrac{\gamma}{\xi^2_{min}}}$$

$$\text{approx. \#2 VAR} = \frac{\mu\, \sigma^2}{1 + \dfrac{\gamma}{\sigma^2}}$$

Table 2-2

Comparison of measured weight variance and two approximate expressions

2-E. The Single Weight Power Separator

Analysis of a single weight adaptive power-separator is of great interest for many reasons. Among them is that a complete analysis can be performed with no need for approximations. In this report the mean weight value and weight variance are analysed, which leads to some surprising results that further illuminate the performance of an LMS filter. The results show that previous approximate analysis of multi-weight filters by Widrow [9, 10] , Senne [11] , Brown [12] , and Davisson [13] to name a few, are accurate enough for useful results in most cases, with the greatest errors occurring during fast adaptation.

The configuration of the adaptive power-separator is a leaky-LMS filter in a line-enhancer as shown in figure 2-1. Note that in this case $W_j = w_j$ has a single element; and $S_j = x_j$, a single element. The input, $d_j = n_j$, is composed of a DC value, a, which is the 'signal'; and white, zero-mean Gaussian noise with variance σ^2. Physically this may be thought of as a degenerate (zero-frequency) sinusoid of power a^2 in noise of power σ^2.

By substituting Eq. 2-2 into Eq. 2-3 and then expanding Eq. 2-3, we find:

$$W_{N+1} = 2\mu n_N n_{N-1} + 2\mu \sum_{i=0}^{N-1} \left[n_i n_{i-1} \prod_{j=i+1}^{N} (\nu - 2\mu n_{j-1}^2) \right] + W_0 \prod_{j=0}^{N} (\nu - 2\mu n_{j-1}^2) \qquad \text{2-18}$$

Using this relationship we can derive the expected (or average) weight value as:

$$E\{W_{N+1}\} = E\{2\mu n_N n_{N-1}\} + 2\mu \sum_{i=0}^{N-1} E\{n_{i-1} n_i \prod_{j=i+1}^{N} (V - 2\mu n_{j-1}^2)\}$$
$$+ W_0 E\{\prod_{j=0}^{N} (V - 2\mu n_{j-1}^2)\}$$

2-19

Using the following properties of white, Gaussian noise with the defined parameters:

$$E\{n_i n_j\} = \begin{cases} \sigma^2 + a^2 & i=j \\ a^2 & i \neq j \end{cases}$$

$$E\{f(n_i) f(n_j)\} = E\{f(n_i)\} E\{f(n_j)\} \quad \text{if} \quad i \neq j$$

we get:

$$E\{W_{N+1}\} = 2\mu a^2 + 2\mu \sum_{i=0}^{N-1} [aE\{n_i (V - 2\mu n_i^2)\} \prod_{j=i+2}^{N} E\{V - 2\mu n_{j-1}^2\}]$$
$$+ W_0 \prod_{j=0}^{N} E\{V - 2\mu n_{j-1}^2\}$$

$$= 2\mu a^2 + 2\mu \sum_{i=0}^{N-1} [a(aV - 2\mu a^3 - 6\mu\sigma^2 a) \prod_{j=i+2}^{N} (V - 2\mu\sigma^2 - 2\mu a^2)]$$
$$+ W_0 \prod_{j=0}^{N} (V - 2\mu a^2 - 2\mu\sigma^2)$$

$$= 2\mu a^2 + 2\mu \sum_{i=0}^{N-1} [(a^2 V - 2\mu a^4 - 6\mu\sigma^2 a^2)(V - 2\mu\sigma^2 - 2\mu a^2)^{N-i-1}]$$
$$+ W_0 (V - 2\mu a^2 - 2\mu\sigma^2)^{N+1}$$

$$= 2\mu a^2 + 2\mu(a^2 V - 2\mu a^4 - 6\mu\sigma^2 a^2) \left[\frac{1 - (V - 2\mu\sigma^2 - 2\mu a^2)^N}{1 - V + 2\mu(\sigma^2 + a^2)}\right]$$

$$+ W_0(\gamma - 2\mu a^2 - 2\mu\sigma^2)^{N+1}$$

2-20

This equation (2-36) expresses the mean value of the weight as a function of the amount of data, or time, used in adaption. An interesting result is the final mean value of the weight. We see that if $|\gamma - 2\mu\sigma^2 - 2\mu a^2| < 1$ then the adaptive process is stable (convergent), and so we can find:

$$\lim_{N \to \infty} E\{W_N\} = 2\mu a^2 + \frac{2\mu(a^2\gamma - 2\mu a^4 - 6\mu\sigma^2 a^2)}{1-\gamma + 2\mu\sigma^2 + 2\mu a^2}$$
$$= \frac{a^2 - 4\mu\sigma^2 a^2}{a^2 + (\sigma^2 + \frac{1-\gamma}{2\mu})}$$

2-21

The Wiener solution (minimum mean-square-error) for the converged weight value can be shown to be:

$$W^* = \frac{a^2}{a^2 + \sigma^2}$$

2-22

The adaptive power separator based on the leaky LMS algorithm generates a converged solution similiar to a Wiener solution for a problem in which the variance of the noise is $\sigma^2 + \frac{1-\gamma}{2\mu}$. However, the adaptive solution given by equation 2-36 is not quite equal to the Wiener solution since there is an extra term of $\frac{-4\mu\sigma^2 a^2}{a^2 + (\sigma^2 + \frac{1-\gamma}{2\mu})}$. Senne [11] demonstrated by experiment that LMS filters do not converge, in general, to a Wiener solution if the inputs are correlated, and since the input has a DC component it is correlated over time. Still, we note that if $\mu << \frac{1}{4\sigma^2}$ the adaptive solution will be essentially the Wiener solution.

To reconsider the conditions under which we can find the limiting value of the weight, we note that $\mathcal{V}$ is chosen in the range $0 < \mathcal{V} < 1$ for normal use, and hence from the convergence conditions that:

$$|\mathcal{V} - 2\mu\sigma^2 - 2\mu a^2| < 1$$

$$\mu < \frac{1+\mathcal{V}}{2(\sigma^2 + a^2)} \qquad 2\text{-}23$$

will guarantee that the mean weight value converges. Note also that transients die geometrically as $(\mathcal{V} - 2\mu a^2 - 2\mu\sigma^2)^k$.

While the average, or expected, solution is very useful in understanding the behavior and utility of the filter (see for example previous final report [1]), it is not the complete story. In this section we shall analyse the variance of the single weight. From this information we can determine how much extra noise (misadjustment [1]) appears in the output due to the adaptive process. Also, we will find a new convergence criterion.

The first step is to find $E\{W_{N+1}^2\}$, the mean-square value of the weight. Knowing this, we can determine the variance of the weight by using the following formula:

$$VAR\{W_{N+1}\} = E\{W_{N+1}^2\} - (E\{W_{N+1}\})^2 \qquad 2\text{-}24$$

For $N \geq 3$ we can show that:

$$E\{W_{N+1}^2\} = 4\mu^2p^2 + 8\mu^2a^2t + 8\mu^2a^2q^2\left[\frac{1-s^{N-1}}{1-s}\right]$$
$$+ 4\mu W_0 aqs^N + 4\mu^2 pv\left[\frac{1-r^N}{1-r}\right] + 8\mu^2 aut\left[\frac{1-r^{N-1}}{1-r}\right]$$
$$+ \frac{8\mu^2aq^2u}{1-s}\left[\frac{1-r^{N-2}}{1-r} - \frac{s^{N-1}-sr^{N-2}}{s-r}\right] + 4\mu W_0 uqs^{N-1}$$
$$+ 4\mu W_0 qu\left[\frac{r^N-rs^{N-1}}{r-s}\right] + W_0^2 r^{N+1}$$

2-25

Where

$$p \doteq a^2+\sigma^2$$
$$q \doteq V\, a-2\mu(a^3+3a\sigma^2)$$
$$r \doteq V^2-4\mu V(a^2+\sigma^2) + 4\mu^2(a^4+6a^2\sigma^2+3\sigma^4)$$
$$s \doteq V - 2\mu(a^2+\sigma^2)$$
$$t \doteq V(a^2+\sigma^2) - 2\mu(a^4+6a^2\sigma^2+3\sigma^4)$$
$$u \doteq V^2a - 4\mu V(a^3+3a\sigma^2) + 4\mu^2(a^5+10a^3\sigma^2+15a\sigma^4)$$
$$v \doteq V^2(a^2+\sigma^2) - 4\mu V(a^4+6a^2\sigma^2+3\sigma^4)$$
$$+ 4\mu^2(a^6+15a^4\sigma^2+45a^2\sigma^4+15\sigma^6)$$

NOTE: this result was reported earlier in Quarterly Report #3 for this contract. However, due to an algebraic error the original equation, equation #8, is incorrect. Thus Eq. 2-41 and 2-42 of this report supplant Eq. 8 and 9 of Quarterly Report #3, "Exact Analysis of a Special Adaptive Power Separator".

This is a formidable expression! However, if $|s|<1$ and $|r|<1$, then $\lim_{N\to\infty} E\{W_N^2\} = E\{W^2\}$ will converge to:

$$E\{W^2\} = 4\mu^2p^2 + 8\mu^2a^2t + \frac{8\mu^2a^2q^2}{1-s} + \frac{4\mu^2pv}{1-r}$$

$$+ \frac{8\mu^2 aut}{1-r} + \frac{8\mu^2 aq^2 u}{(1-s)(1-r)}$$

2-26

Furthermore, if we take a to be zero and $\mathcal{V}$ to be 1, then

$$\mathrm{VAR}\{W\} = 4\mu^2\sigma^4 + \frac{4\mu^2\sigma^2(\sigma^2-12\mu\sigma^4+60\mu^2\sigma^6)}{4\mu\sigma^2-12\mu^2\sigma^4} - E\{W\}$$

$$= \mu\sigma^2\left[\frac{1-8\mu\sigma^2+48\mu^2\sigma^4}{1-3\mu\sigma^2}\right]$$

2-27

This case is that of a conventional LMS filter, and was analysed by Widrow in [5] The approximate analysis used by Widrow indicated that the variance of the weight should be $\mu\sigma^2$. This agrees very well with the exact analysis if we recall Widrow's stipulation that $\mu\sigma^2 << 1$, which corresponds to slow, low-noise adaptation.

Recall that convergence of $E\{W_N^2\}$ required that $|s|<1$ and $|r|<1$. The first condition becomes:

$$|s|<1 \quad \Rightarrow |\mathcal{V} - 2\mu(a^2+\sigma^2)| < 1$$

$$\Rightarrow \mu < \frac{1+\mathcal{V}}{2(a^2+\sigma^2)}$$

Which is the same condition as for convergence of the mean weight value. The second condition becomes:

$$|r|<1 \quad \Rightarrow |\mathcal{V}^2 - 4\mu\mathcal{V}p + 4\mu^2(a^4+6a^2\sigma^2+3\sigma^4)| < 1$$

$$\Rightarrow 0 > \mathcal{V}^2 - 1 - 4\mu\mathcal{V}p + 4\mu^2(a^4+6a^2\sigma^2+3\sigma^4)$$

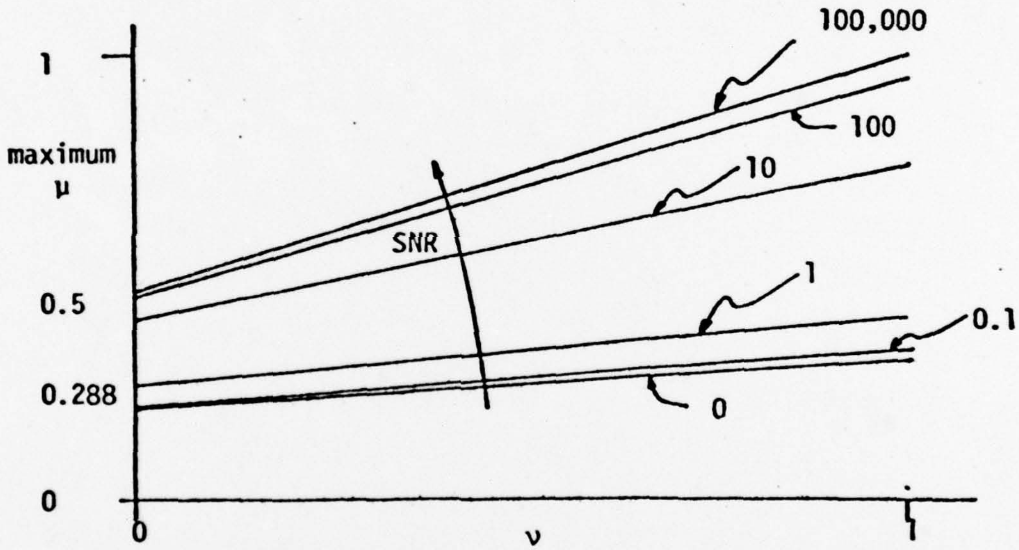
$$\Rightarrow \mu < \frac{\mathcal{V}(a^2+\sigma^2) + \sqrt{a^4+6a^2\sigma^2+3\sigma^4-4\mathcal{V}^2a^2\sigma^2-2\mathcal{V}^2\sigma^4}}{2(a^4+6a^2\sigma^2+3\sigma^4)}$$

2-28

if $SNR \doteq \frac{a^2}{\sigma^2}$

$$\mu < \frac{\nu(SNR+1) + \sqrt{SNR^2+6SNR+3-4\nu^2 SNR-2\nu^2}}{2(a^2 SNR+6a^2+3\sigma^2)} \qquad 2\text{-}29$$

Again, this is not a particularly simple expression. Figure 2-13 is a plot of the maximum possible μ vs ν for various SNR's (holding total input power = 1, maximum μ assuring convergence of mean and



Maximum μ for convergence of variance vs. ν, as a function of SNR (holding total input power at 1).

Figure 2-13

variance of the weight vector). From this we see that there are two limiting cases: a) high SNR and b) low SNR. In the high SNR case equation 2-44 simplifies to:

$$\mu < \frac{\nu + 1}{2 \times \text{total-input-power}}$$

2-30

which varies from $\frac{1}{2 \text{ x total-input-power}}$ to $\frac{1}{\text{input-power}}$ as ν varies from 0 to 1. In the low SNR case equation 2-44 simplifies to:

$$\mu < \frac{\nu + \sqrt{3-2\nu^2}}{6 \text{ x total-input-power}} \qquad 2\text{-}31$$

which varies from $\frac{\sqrt{3}}{6 \text{ x total-input-power}}$ to $\frac{1}{3 \text{ x total-input-power}}$ as ν varies from 0 to 1. The most stringent requirement on μ occurs when $\nu = 0$ (although in actual practice ν is rarely less than 0.99) and SNR is low. For this case $\mu < \frac{\sqrt{3}}{6 \text{ x total-input-power}} = \frac{0.288}{\text{input-power}}$. If μ is selected by this criterion, the variance will always converge and remain finite. Note that this condition is more stringent than the condition for guaranteed convergence of the mean which required that:

$$\mu < \frac{1}{\text{total-input-power}}$$

in the worst case. Also, note that the conditions for convergence in the mean presented in [9, 5] for an LMS filter ($\mu < \frac{1}{\text{input-power}}$) where $\nu = 1$ is not sufficient to guarantee convergence of the variance in low SNR regions. Of course, in normal practice μ is very much smaller than the bound so this problem does not arise. However, if a μ is selected so that the filter converges in the mean but not in variance, then we would expect the weight to oscillate randomly about the mean value with ever larger oscillations. This analysis of a single-weight filter corroborates Senne's earlier experimental work with multi-weight filters: that to guarantee convergence of the weight noise variance requires μ to be several times smaller than the value needed to assure convergence of the mean weight value.

2-F. Conclusions of Part 2

We have studied the second order effects of the Leaky LMS (LLMS) algorithm. The LLMS algorithm causes the weights to vary about the mean or expected solution. This will modulate the output of the filter, causing added undesirable noise components in the filter's output. For sinusoidal inputs with frequencies between 10% and 90% of Nyquist frequency we found that the distortion power was less than 0.2% of the desirable output power. For very low frequency inputs the distortion power increased, but even at 1% of Nyquist the distortion power was still less than 5%. Furthermore, most (98%) of the distortion was in the third harmonic.

For white noise inputs to a LLMS based Adaptive Power Separator (APS) we have derived an equation for the variance of the noise in the weight vector. This equation (Eq. 2-29) agrees with previous analysis [5, 9, 11] of weight vector noise in LMS filters, a special case of LLMS in which $\gamma = 1$. This analysis also confirms an ealier conjecture [1] that the LLMS algorithm would have less weight noise than an equivalent injected noise scheme.

Finally, an exact analysis of a special case (single-weight) of the LLMS driven APS was performed. This analysis confirms Senne's [11] observation that LMS does not converge to the Wiener solution if the input is correlated. However, the bias from the Wiener solution can be made arbitrarily small by decreasing μ. Also, we found that LLMS (and LMS) filters may not stabalize in a mean-square sense even though they converge in the mean. A new criterion for μ is presented which guarantees mean-square stability.

REFERENCES

1. B. Widrow, R. Chestek, J. R. Treichler, "Final Report on Adaptive Antenna Techniques", March 1977, Naval Air System Command Contract No. N00019-76-C-0250.

2. B. Widrow, J. M. McCool, M. G. Larimore, C. R. Johnson, Jr., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," Proc. IEEE, Vol. 64, No. 8, August 1976, pp. 1151-1162.

3. J. K. Kim and L. D. Davisson, "Adaptive Linear Estimation for Stationary M-dependent Processes," IEEE Trans. Information Theory, Vol. IT21, pp. 23-31, January 1975

4. L. E. Brennan, E. L. Pugh, and I. S. Reed, "Control-loop Noise in Adaptive Array Antennas," J. Acoust. Soc. Am., Vol. 34, pp. 289-297, March 1962

5. B. Widrow, "Adaptive Filters I: Fundamentals," Stanford Electronic Labs Technical Report No. 6764-6, December 1966

6. B. Widrow, "Adaptive Filters," from Aspects of Network and System Theory, edited by R. E. Kalman and N. DeClaris; Holt, Reinhart, and Winston, Inc., N.Y., 1970

   (Also reprinted in "Proceedings of the Adaptive Antenna Systems Workshop," March 11-13, 1974, Volume I; W. F. Gabriel, working chairman, published by Naval Research Laboratory, Washington, D. C. as NRL Report 7803, September 27, 1974)

7. B. Noble, Applied Linear Algebra, Prentice Hall, Inc., Englewood Cliffs, N.J., 1969

8. J. R. Treichler, The spectral line enhancer: the concept, an implementation and an application, Ph. D. dissertation, Stanford Univ., Stanford Calif., June. 1977.

9. B. Widrow et al., "Adaptive noise cancelling: Principles and Applications", Proc. of IEEE, vol 63 No. 12, pp 1692-1716, Dec. 1975.

10. S. T. Alexander, The ALE Output Covariance Function For a Sinusoid in Uncorrelated Noise, Technical report 162, Naval Ocean Systems Center, San Diego Calif., 1977.

11. K. Senne, Adaptive linear discrete-time estimation, Stanford Electronics Lab., Stanford Univ., Rep. SU-SEL-68-090, June 1968 (Ph. D. dissertation).

12. J. E. Brown, Adaptive Estimation in Nonstationary Environments, Stanford Electronics Lab., Stanford Univ., Rep. SU-SEL-70-056, Aug. 1970 (Ph. D. dissertation).

13. L. D. Davisson, " Steady-state error in adaptive mean-square minimisation", IEEE Transactions on Information Theory, vol IT-16 No. 4, pp 382-385, July 1970.