Adaptive Antenna Systems

B. WIDROW, MEMBER, IEEE, P. E. MANTEY, MEMBER, IEEE, L. J. GRIFFITHS, STUDENT MEMBER, IEEE, AND B. B. GOODE, STUDENT MEMBER, IEEE

Abstract—**A system consisting of an antenna array and an adaptive processor can perform filtering in both the space and the frequency domains, thus reducing the sensitivity of the signal-receiving system to interfering directional noise sources.**

Variable weights of a signal processor can be automatically adjusted by a simple adaptive technique based on the least-mean-squares (LMS) algorithm. During the adaptive process an injected pilot signal simulates a received signal from a desired "look" direction. This allows the array to be "trained" so that its directivity pattern has a main lobe in the previously specified look direction. At the same time, the array processing system can reject any incident noises, whose directions of propagation are different from the desired look direction, by forming appropriate nulls in the antenna directivity pattern. The array adapts itself to form a main lobe, with its direction and bandwidth determined by the pilot signal, and to reject signals or noises occurring outside the main lobe as well as possible in the minimum mean-square error sense.

Several examples illustrate the convergence of the LMS adaptation procedure toward the corresponding Wiener optimum solutions. Rates of adaptation and misadjustments of the solutions are predicted theoretically and checked experimentally. Substantial reductions in noise reception are demonstrated in computer-simulated experiments. The techniques described are applicable to signal-receiving arrays for use over a wide range of frequencies.

Introduction

THE SENSITIVITY of a signal-receiving array to interfering noise sources can be reduced by suitable processing of the outputs of the individual array elements. The combination of array and processing acts as a filter in both space and frequency. This paper describes a method of applying the techniques of adaptive filtering[1] to the design of a receiving antenna system which can extract directional signals from the medium with minimum distortion due to noise. This system will be called an *adaptive array*. The adaptation process is based on minimization of mean-square error by the LMS algorithm.[2]-[4] The system operates with knowledge of the direction of arrival and spectrum of the signal, but with no knowledge of the noise field. The adaptive array promises to be useful whenever there is interference that possesses some degree of spatial correlation; such conditions manifest themselves over the entire spectrum, from seismic to radar frequencies.

Manuscript received May 29, 1967; revised September 5, 1967.

B. Widrow and L. J. Griffiths are with the Department of Electrical Engineering, Stanford University, Stanford, Calif.

P. E. Mantey was formerly with the Department of Electrical Engineering, Stanford University. He is presently with the Control and Dynamical Systems Group, IBM Research Laboratories, San Jose, Calif.

B. B. Goode is with the Department of Electrical Engineering, Stanford University, Stanford, Calif., and the Navy Electronics Laboratory, San Diego, Calif.

The term "adaptive antenna" has previously been used by Van Atta[5] and others[6] to describe a self-phasing antenna system which reradiates a signal in the direction from which it was received. This type of system is called adaptive because it performs without any prior knowledge of the direction in which it is to transmit. For clarity, such a system might be called an adaptive *transmitting* array; whereas the system described in this paper might be called an adaptive *receiving* array.

The term "adaptive filter" has been used by Jakowatz, Shuey, and White[7] to describe a system which extracts an unknown signal from noise, where the signal waveform recurs frequently at random intervals. Davisson[8] has described a method for estimating an unknown signal waveform in the presence of white noise of unknown variance. Glaser[9] has described an adaptive system suitable for the detection of a pulse signal of fixed but unknown waveform.

Previous work on array signal processing directly related to the present paper was done by Bryn, Mermoz, and Shor. The problem of detecting Gaussian signals in additive Gaussian noise fields was studied by Bryn,[10] who showed that, assuming K antenna elements in the array, the Bayes optimum detector could be implemented by either K^2 linear filters followed by "conventional" beam-forming for each possible signal direction, or by K linear filters for each possible signal direction. In either case, the measurement and inversion of a $2K$ by $2K$ correlation matrix was required at a large number of frequencies in the band of the signal. Mermoz[11] proposed a similar scheme for narrowband known signals, using the signal-to-noise ratio as a performance criterion. Shor[12] also used a signal-to-noise-ratio criterion to detect narrowband pulse signals. He proposed that the sensors be switched off when the signal was known to be absent, and a pilot signal injected as if it were a noise-free signal impinging on the array from a specified direction. The need for specific matrix inversion was circumvented by calculating the gradient of the ratio between the output power due to pilot signal and the output power due to noise, and using the method of steepest descent. At the same time, the number of correlation measurements required was reduced, by Shor's procedure, to $4K$ at each step in the adjustment of the processor. Both Mermoz and Shor have suggested the possibility of real-time adaptation.

This paper presents a potentially simpler scheme for obtaining the desired array processing improvement in real time. The performance criterion used is minimum mean-square error. The statistics of the signal are assumed

to be known, but no prior knowledge or direct measurements of the noise field are required in this scheme. The adaptive array processor considered in the study may be automatically adjusted (adapted) according to a simple iterative algorithm, and the procedure does not directly involve the computation of any correlation coefficients or the inversion of matrices. The input signals are used only once, as they occur, in the adaptation process. There is no need to store past input data; but there is a need to store the processor adjustment values, i.e., the processor weighting coefficients ("weights"). Methods of adaptation are presented here, which may be implemented with either analog or digital adaptive circuits, or by digital-computer realization.

Directional and Spatial Filtering

An example of a linear-array receiving antenna is shown in Fig. 1(a) and (b). The antenna of Fig. 1(a) consists of seven isotropic elements spaced $\lambda_0/2$ apart along a straight line, where λ_0 is the wavelength of the center frequency f_0 of the array. The received signals are summed to produce an array output signal. The directivity pattern, i.e., the relative sensitivity of response to signals from various directions, is plotted in this figure in a plane over an angular range of $-\pi/2<\theta<\pi/2$ for frequency f_0. This pattern is symmetric about the vertical line $\theta=0$. The main lobe is centered at $\theta=0$. The largest-amplitude side lobe, at $\theta=24°$, has a maximum sensitivity which is 12.5 dB below the maximum main-lobe sensitivity. This pattern would be different if it were plotted at frequencies other than f_0.

The same array configuration is shown in Fig. 1(b); however, in this case the output of each element is delayed in time before being summed. The resulting directivity pattern now has its main lobe at an angle of ψ radians, where

$$\psi = \sin^{-1}\left(\frac{\lambda_0 \delta f_0}{d}\right) = \sin^{-1}\left(\frac{c\delta}{d}\right) \quad . \quad (1)$$

in which

f_0 = frequency of received signal
λ_0 = wavelength at frequency f_0
δ = time-delay difference between neighboring-element outputs
d = spacing between antenna elements
c = signal propagation velocity $=\lambda_0 f_0$.

The sensitivity is maximum at angle ψ because signals received from a plane wave source incident at this angle, and delayed as in Fig. 1(b), are in phase with one another and produce the maximum output signal. For the example illustrated in the figure, $d=\lambda_0/2$, $\delta=(0.12941/f_0)$, and therefore $\psi=\sin^{-1}(2\delta f_0)=15°$.

There are many possible configurations for phased arrays. Fig. 2(a) shows one such configuration where each of the antenna-element outputs is weighted by two weights in parallel, one being preceded by a time delay of a quarter of a cycle at frequency f_0 (i.e., a 90° phase shift), denoted by $1/(4f_0)$. The output signal is the sum of all the weighted signals, and since all weights are set to unit values, the directivity pattern at frequency f_0 is by symmetry the same as that of Fig. 1(a). For purposes of illustration, an interfering directional sinusoidal "noise" of frequency f_0 incident on the array is shown in Fig. 2(a), indicated by the dotted arrow. The angle of incidence (45.5°) of this noise is such that it would be received on one of the side lobes of the directivity pattern with a sensitivity only 17 dB less than that of the main lobe at $\theta=0°$.

If the weights are now set as indicated in Fig. 2(b), the directivity pattern at frequency f_0 becomes as shown in that figure. In this case, the main lobe is almost unchanged from that shown in Figs. 1(a) and 2(a), while the particular side lobe that previously intercepted the sinusoidal noise in Fig. 2(a) has been shifted so that a null is now placed in the direction of that noise. The sensitivity in the noise direction is 77 dB below the main lobe sensitivity, improving the noise rejection by 60 dB.

A simple example follows which illustrates the existence and calculation of a set of weights which will cause a signal from a desired direction to be accepted while a "noise" from a different direction is rejected. Such an example is illustrated in Fig. 3. Let the signal arriving from the desired direction $\theta=0°$ be called the "pilot" signal $p(t)=P \sin \omega_0 t$, where $\omega_0 \triangleq 2\pi f_0$, and let the other signal, the noise, be chosen as $n(t)=N \sin \omega_0 t$, incident to the receiving array at an angle $\theta=\pi/6$ radians. Both the pilot signal and the noise signal are assumed for this example to be at exactly the same frequency f_0. At a point in space midway between the antenna array elements, the signal and the noise are assumed to be in phase. In the example shown, there are two identical omnidirectional array elements, spaced $\lambda_0/2$ apart. The signals received by each element are fed to two variable weights, one weight being preceded by a quarter-wave time delay of $1/(4f_0)$. The four weighted signals are then summed to form the array output.

The problem of obtaining a set of weights to accept $p(t)$ and reject $n(t)$ can now be studied. Note that with any set of nonzero weights, the output is of the form $A \sin(\omega_0 t+\phi)$, and a number of solutions exist which will make the output be $p(t)$. However, the output of the array must be independent of the amplitude and phase of the noise signal if the array is to be regarded as rejecting the noise. Satisfaction of this constraint leads to a unique set of weights determined as follows.

The array output due to the pilot signal is

$$P[(w_1 + w_3)\sin \omega_0 t + (w_2 + w_4)\sin(\omega_0 t - \pi/2)]. \quad (2)$$

For this output to be equal to the desired output of $p(t)=P \sin \omega_0 t$ (which is the pilot signal itself), it is necessary that

$$\left.\begin{aligned} w_1 + w_3 &= 1 \\ w_2 + w_4 &= 0 \end{aligned}\right\}. \quad (3)$$

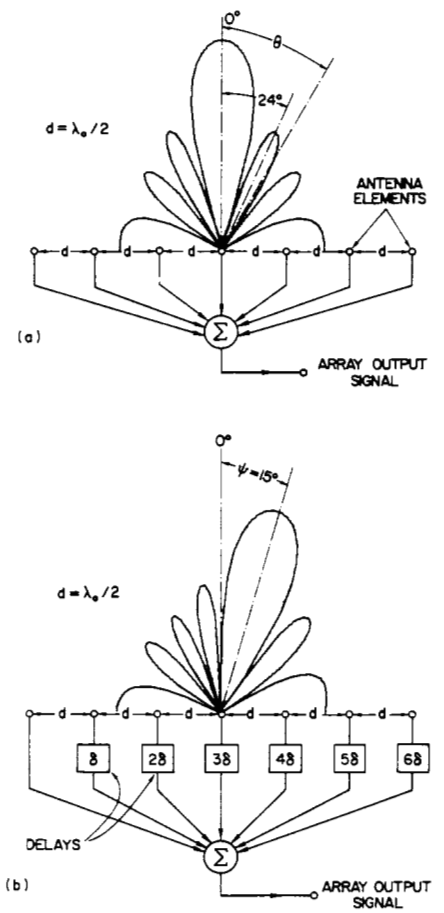

Fig. 1. Directivity pattern for a linear array. (a) Simple array. (b) Delays added.

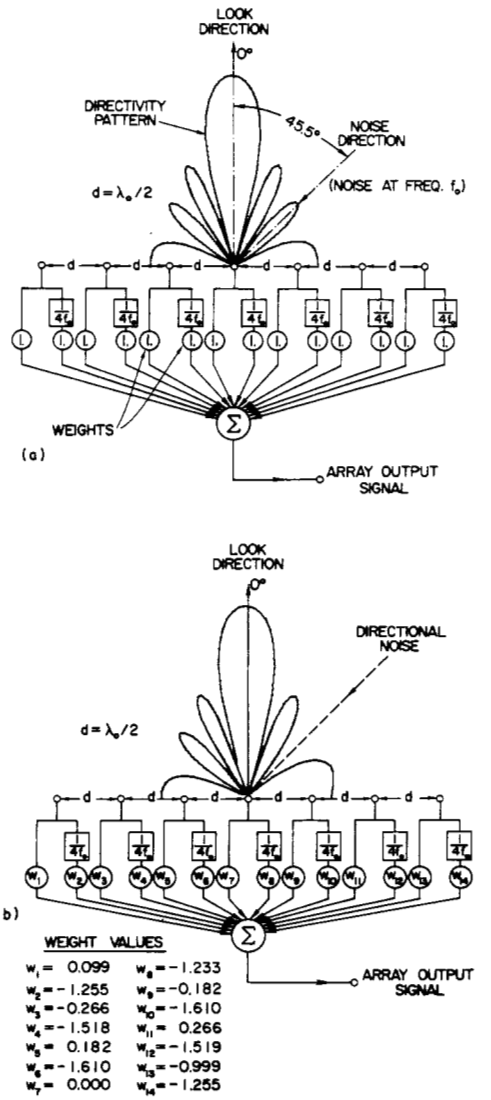

Fig. 2. Directivity pattern of linear array. (a) With equal weighting. (b) With weighting for noise elimination.

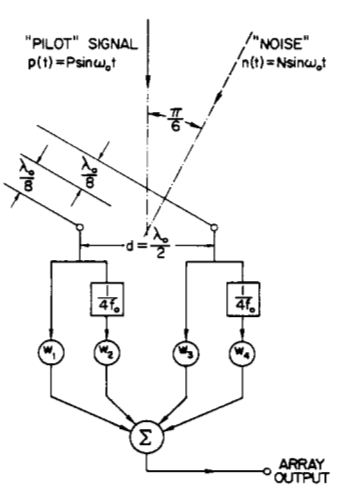

Fig. 3. Array configuration for noise elimination example.

With respect to the midpoint between the antenna elements, the relative time delays of the noise at the two antenna elements are $\pm[1/(4f_0)]\sin\pi/6 = \pm 1/(8f_0) = \pm\lambda_0/(8c)$, which corresponds to phase shifts of $\pm\pi/4$ at frequency f_0. The array output due to the incident noise at $\theta=\pi/6$ is then

$$N\left[w_1 \sin\left(\omega_0 t - \frac{\pi}{4}\right) + w_2 \sin\left(\omega_0 t - \frac{3\pi}{4}\right) + w_3 \sin\left(\omega_0 t + \frac{\pi}{4}\right) + w_4 \sin\left(\omega_0 t - \frac{\pi}{4}\right)\right]. \quad (4)$$

For this response to equal zero, it is necessary that

$$\left.\begin{aligned} w_1 + w_4 &= 0 \\ w_2 - w_3 &= 0 \end{aligned}\right\}. \quad (5)$$

Thus the set of weights that satisfies the signal and noise response requirements can be found by solving (3) and (5) simultaneously. The solution is

$$w_1 = \tfrac{1}{2},\ w_2 = \tfrac{1}{2},\ w_3 = \tfrac{1}{2},\ w_4 = -\tfrac{1}{2}. \quad (6)$$

With these weights, the array will have the desired properties in that it will accept a signal from the desired direction, while rejecting a noise, even a noise which is at the same frequency f_0 as the signal, because the noise comes from a different direction than does the signal.

The foregoing method of calculating the weights is more illustrative than practical. This method is usable when there are only a small number of directional noise sources, when the noises are monochromatic, and when the directions of the noises are known *a priori*. A practical processor should not require detailed information about the number and the nature of the noises. The adaptive processor described in the following meets this requirement. It recursively solves a sequence of simultaneous equations, which are generally overspecified, and it finds solutions which minimize the mean-square error between the pilot signal and the total array output.

Configurations of Adaptive Arrays

Before discussing methods of adaptive filtering and signal processing to be used in the adaptive array, various spatial and electrical configurations of antenna arrays will be considered. An adaptive array configuration for processing narrowband signals is shown in Fig. 4. Each individual antenna element is shown connected to a variable weight and to a quarter-period time delay whose output is in turn connected to another variable weight. The weighted signals are summed, as shown in the figure. The signal, assumed to be either monochromatic or narrowband, is received by the antenna element and is thus weighted by a complex gain factor $Ae^{j\phi}$. Any phase angle $\phi = -\tan^{-1}(w_2/w_1)$ can be chosen by setting the two weight values, and the magnitude of this complex gain factor $A = \sqrt{w_1^2 + w_2^2}$ can take on a wide range of values limited only by the range limitations of the two individual weights. The latter can assume a continuum of both positive and negative values.

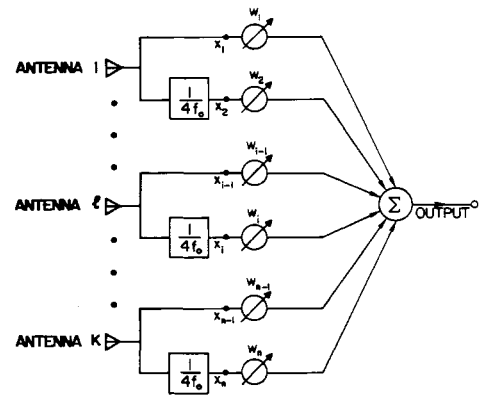

Fig. 4. Adaptive array configuration for receiving narrowband signals.

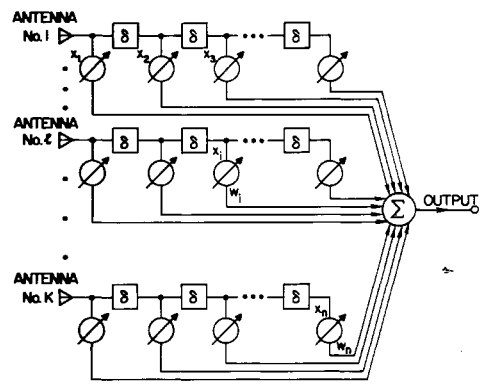

Fig. 5. Adaptive array configuration for receiving broadband signals.

Thus the two weights and the $1/(4f_0)$ time delay provide completely adjustable linear processing for narrowband signals received by each individual antenna element.

The full array of Fig. 4 represents a completely general way of combining the antenna-element signals in an adjustable linear structure when the received signals and noises are narrowband. It should be realized that the same generality (for narrowband signals) can be achieved even when the time delays do not result in a phase shift of exactly $\pi/2$ at the center frequency f_0. Keeping the phase shifts close to $\pi/2$ is desirable for keeping required weight values small, but is not necessary in principle.

When one is interested in receiving signals over a wide band of frequencies, each of the phase shifters in Fig. 4 can be replaced by a tapped-delay-line network as shown in Fig. 5. This tapped delay line permits adjustment of gain and phase as desired at a number of frequencies over the band of interest. If the tap spacing is sufficiently close, this network approximates the ideal filter which would allow complete control of the gain and phase at each frequency in the passband.

Adaptive Signal Processors

Once the form of network connected to each antenna element has been chosen, as shown for example in Fig. 4 or Fig. 5, the next step is to develop an adaptation procedure which can be used to adjust automatically the multiplying weights to achieve the desired spatial and frequency filtering.

The procedure should produce a given array gain in the specified look direction while simultaneously nulling out interfering noise sources.

Fig. 6 shows an adaptive signal-processing element. If this element were combined with an output-signal quantizer, it would then comprise an adaptive threshold logic unit. Such an element has been called an "Adaline"[13] or a threshold logic unit (TLU).[14] Applications of the adaptive threshold element have been made in pattern-recognition systems and in experimental adaptive control systems.[2],[3],[14]–[17]

In Fig. 6 the input signals $x_1(t), \cdots, x_i(t), \cdots, x_n(t)$ are the same signals that are applied to the multiplying weights $w_1, \cdots, w_i, \cdots, w_n$ shown in Fig. 4 or Fig. 5. The heavy lines show the paths of signal flow; the lighter lines show functions related to weight-changing or adaptation processes.

The output signal $s(t)$ in Fig. 6 is the weighted sum

$$s(t) = \sum_{i=1}^{n} x_i(t) w_i \tag{7}$$

where n is the number of weights; or, using vector notation

$$s(t) = \boldsymbol{W}^T \boldsymbol{X}(t) \tag{8}$$

where $\boldsymbol{W}^T$ is the transpose of the weight vector

$$\boldsymbol{W} \triangleq \begin{bmatrix} w_1 \\ \cdot \\ \cdot \\ \cdot \\ w_i \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix}$$

and the signal-input vector is

$$\boldsymbol{X}(t) \triangleq \begin{bmatrix} x_1(t) \\ \cdot \\ \cdot \\ \cdot \\ x_i(t) \\ \cdot \\ \cdot \\ \cdot \\ x_n(t) \end{bmatrix}.$$

For digital systems, the input signals are in discrete-time sampled-data form and the output is written

$$s(j) = \boldsymbol{W}^T \boldsymbol{X}(j) \tag{9}$$

where the index j indicates the jth sampling instant.

In order that adaptation take place, a "desired response" signal, $d(t)$ when continuous or $d(j)$ when sampled, must be supplied to the adaptive element. A method for obtaining this signal for adaptive antenna array processing will be discussed in a following section.

The difference between the desired response and the output response forms the error signal $\varepsilon(j)$:

$$\varepsilon(j) = d(j) - \boldsymbol{W}^T \boldsymbol{X}(j). \tag{10}$$

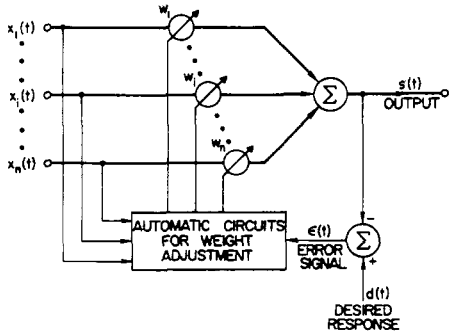

Fig. 6. Basic adaptive element.

This signal is used as a control signal for the "weight adjustment circuits" of Fig. 6.

Solving Simultaneous Equations

The purpose of the adaptation or weight-changing processes is to find a set of weights that will permit the output response of the adaptive element at each instant of time to be equal to or as close as possible to the desired response. For each input-signal vector $\boldsymbol{X}(j)$, the error $\varepsilon(j)$ of (10) should be made as small as possible.

Consider the finite set of linear simultaneous equations

$$\begin{cases} \boldsymbol{W}^T \boldsymbol{X}(1) = d(1) \\ \boldsymbol{W}^T \boldsymbol{X}(2) = d(2) \\ \quad \vdots \quad \vdots \\ \boldsymbol{W}^T \boldsymbol{X}(j) = d(j) \\ \quad \vdots \quad \vdots \\ \boldsymbol{W}^T \boldsymbol{X}(N) = d(N) \end{cases} \tag{11}$$

where N is the total number of input-signal vectors; each vector is a measurement of an underlying n-dimensional random process. There are N equations, corresponding to N instants of time at which the output response values are of concern; there are n "unknowns," the n weight values which form the components of $\boldsymbol{W}$. The set of equations (11) will usually be overspecified and inconsistent, since in the present application, with an ample supply of input data, it is usual that $N \gg n$. [These equations did have a solution in the simple example represented in Fig. 3. The solution is given in (6). Although the simultaneous equations (3) in that example appear to be different from (11), they are really the same, since those in (3) are in a specialized form for the case when all inputs are deterministic sinusoids which can be easily specified over all time in terms of amplitudes, phases, and frequencies.]

When N is very large compared to n, one is generally interested in obtaining a solution of a set of N equations [each equation in the form of (10)] which minimizes the sum of the squares of the errors. That is, a set of weights $\boldsymbol{W}$ is found to minimize

$$\sum_{j=1}^{N} \varepsilon^2(j). \tag{12}$$

When the input signals can be regarded as stationary stochastic variables, one is usually interested in finding a set of weights to minimize mean-square error. The quantity of interest is then the expected value of the square of the error, i.e., the mean-square error, given by

$$E[\varepsilon^2(j)] \triangleq \overline{\varepsilon^2}. \tag{13}$$

The set of weights that minimizes mean-square error can be calculated by squaring both sides of (10) which yields

$$\varepsilon^2(j) = d^2(j) + \boldsymbol{W}^T\boldsymbol{X}(j)\boldsymbol{X}(j)^T\boldsymbol{W} - 2d(j)\boldsymbol{W}^T\boldsymbol{X}(j) \tag{14}$$

and then taking the expected value of both sides of (14)

$$\begin{aligned} E[\varepsilon^2(j)] &= E[d^2 + \boldsymbol{W}^T\boldsymbol{X}(j)\boldsymbol{X}^T(j)\boldsymbol{W} - 2\boldsymbol{W}^T d(j)\boldsymbol{X}(j)] \\ &= E[d^2] + \boldsymbol{W}^T\boldsymbol{\Phi}(x, x)\boldsymbol{W} - 2\boldsymbol{W}^T\boldsymbol{\Phi}(x, d) \end{aligned} \tag{15}$$

where

$$\boldsymbol{\Phi}(x, x) \triangleq E[\boldsymbol{X}(j)\boldsymbol{X}^T(j)] \triangleq E\begin{bmatrix} x_1x_1 & x_1x_2 \cdots x_1x_n \\ x_2x_1 & \cdots x_2x_n \\ \vdots & \\ x_nx_1 & x_nx_n \end{bmatrix} \tag{16}$$

and

$$\boldsymbol{\Phi}(x, d) \triangleq E[\boldsymbol{X}(j)d(j)] \triangleq E\begin{bmatrix} x_1d \\ x_2d \\ \vdots \\ x_id \\ \vdots \\ x_nd \end{bmatrix}. \tag{17}$$

The symmetric matrix $\boldsymbol{\Phi}(x, x)$ is a matrix of cross correlations and autocorrelations of the input signals to the adaptive element, and the column matrix $\boldsymbol{\Phi}(x, d)$ is the set of cross correlations betweeen the n input signals and the desired response signal.

The mean-square error defined in (15) is a quadratic function of the weight values. The components of the gradient of the mean-square-error function are the partial derivatives of the mean-square error with respect to the weight values. Differentiating (15) with respect to $\boldsymbol{W}$ yields the gradient $\nabla E[\varepsilon^2]$, a linear function of the weights,

$$\nabla E[\varepsilon^2] = 2\boldsymbol{\Phi}(x, x)\boldsymbol{W} - 2\boldsymbol{\Phi}(x, d). \tag{18}$$

When the choice of the weights is optimized, the gradient is zero. Then

$$\begin{aligned} \boldsymbol{\Phi}(x, x)\boldsymbol{W}_{\text{LMS}} &= \boldsymbol{\Phi}(x, d) \\ \boldsymbol{W}_{\text{LMS}} &= \boldsymbol{\Phi}^{-1}(x, x)\boldsymbol{\Phi}(x, d). \end{aligned} \tag{19}$$

The optimum weight vector $\boldsymbol{W}_{\text{LMS}}$ is the one that gives the least mean-square error. Equation (19) is the Wiener-Hopf equation, and is the equation for the multichannel least-squares filter used by Burg[18] and Claerbout[19] in the processing of digital seismic array data.

One way of finding the optimum set of weight values is to solve (19). This solution is generally straightforward, but presents serious computational problems when the number of weights n is large and when data rates are high. In addition to the necessity of inverting an $n \times n$ matrix, this method may require as many as $n(n+1)/2$ autocorrelation and cross-correlation measurements to obtain the elements of $\boldsymbol{\Phi}(x, x)$. Furthermore, this process generally needs to be continually repeated in most practical situations where the input signal statistics change slowly. No perfect solution of (19) is possible in practice because of the fact that an infinite statistical sample would be required to estimate perfectly the elements of the correlation matrices.

Two methods for finding approximate solutions to (19) will be presented in the following. Their accuracy is limited by statistical sample size, since they find weight values based on finite-time measurements of input-data signals. These methods do not require explicit measurements of correlation functions or matrix inversion. They are based on gradient-search techniques applied to mean-square-error functions. One of these methods, the LMS algorithm, does not even require squaring, averaging, or differentiation in order to make use of gradients of mean-square-error functions. The second method, a relaxation method, will be discussed later.

The LMS Algorithm

A number of weight-adjustment procedures or algorithms exist which minimize the mean-square error. Minimization is usually accomplished by gradient-search techniques. One method that has proven to be very useful is the LMS algorithm.[1]–[3],[17] This algorithm is based on the method of steepest descent. Changes in the weight vector are made along the direction of the estimated gradient vector. Accordingly,

$$\boldsymbol{W}(j + 1) = \boldsymbol{W}(j) + k_s\hat{\nabla}(j) \tag{20}$$

where

$\boldsymbol{W}(j) \triangleq$ weight vector before adaptation
$\boldsymbol{W}(j + 1) \triangleq$ weight vector after adaptation
$k_s \triangleq$ scalar constant controlling rate of convergence and stability ($k_s < 0$)
$\hat{\nabla}(j) \triangleq$ estimated gradient vector of $\overline{\varepsilon^2}$ with respect to $\boldsymbol{W}$.

One method for obtaining the estimated gradient of the mean-square-error function is to take the gradient of a single time sample of the squared error

$$\hat{\nabla}(j) = \nabla[\varepsilon^2(j)] = 2\varepsilon(j)\nabla[\varepsilon(j)].$$

From (10)

$$\begin{aligned} \nabla[\varepsilon(j)] &= \nabla[d(j) - \boldsymbol{W}^T(j)\boldsymbol{X}(j)] \\ &= -\boldsymbol{X}(j). \end{aligned}$$

Thus

$$\hat{\nabla}(j) = -2\varepsilon(j)\boldsymbol{X}(j). \tag{21}$$

The gradient estimate of (21) is unbiased, as will be shown by the following argument. For a given weight vector $\boldsymbol{W}(j)$,

the expected value of the gradient estimate is

$$E[\hat{\nabla}(j)] = -2E[\{d(j) - \boldsymbol{W}^T(j)\boldsymbol{X}(j)\}\boldsymbol{X}(j)] = -2[\boldsymbol{\Phi}(x, d) - \boldsymbol{W}^T(j)\boldsymbol{\Phi}(x, x)]. \quad (22)$$

Comparing (18) and (22), we see that

$$E[\hat{\nabla}(j)] = \nabla E[\varepsilon^2]$$

and therefore, for a given weight vector, the expected value of the estimate equals the true value.

Using the gradient estimation formula given in (21), the weight iteration rule (20) becomes

$$\boldsymbol{W}(j + 1) = \boldsymbol{W}(j) - 2k_s\varepsilon(j)\boldsymbol{X}(j) \quad (23)$$

and the next weight vector is obtained by adding to the present weight vector the input vector scaled by the value of the error.

The LMS algorithm is given by (23). It is directly usable as a weight-adaptation formula for digital systems. Fig. 7(a) shows a block-diagram representation of this equation in terms of one component w_i of the weight vector $\boldsymbol{W}$. An equivalent differential equation which can be used in analog implementation of continuous systems (see Fig. 7(b)) is given by

$$\frac{d}{dt}\boldsymbol{W}(t) = -2k_s\varepsilon(t)\boldsymbol{X}(t).$$

This equation can also be written as

$$\boldsymbol{W}(t) = -2k_s\int_0^t \varepsilon(\xi)\boldsymbol{X}(\xi)\,d\xi.$$

Fig. 8 shows how circuitry of the type indicated in Fig. 7(a) or (b) might be incorporated into the implementation of the basic adaptive element of Fig. 6.

Convergence of the Mean of the Weight Vector

For the purpose of the following discussion, we assume that the time between successive iterations of the LMS algorithm is sufficiently long so that the sample input vectors $\boldsymbol{X}(j)$ and $\boldsymbol{X}(j+1)$ are uncorrelated. This assumption is common in the field of stochastic approximation.[20]–[22]

Because the weight vector $\boldsymbol{W}(j)$ is a function *only* of the input vectors $\boldsymbol{X}(j-1)$, $\boldsymbol{X}(j-2)$, $\cdots$, $\boldsymbol{X}(0)$ [see (23)] and because the successive input vectors are uncorrelated, $\boldsymbol{W}(j)$ is independent of $\boldsymbol{X}(j)$. For stationary input processes meeting this condition, the expected value $E[\boldsymbol{W}(j)]$ of the weight vector after a large number of iterations can then be shown to converge to the Wiener solution given by (19). Taking the expected value of both sides of (23), we obtain a difference equation in the expected value of the weight vector

$$E[\boldsymbol{W}(j + 1)] = E[\boldsymbol{W}(j)] - 2k_sE[\{d(j) - \boldsymbol{W}^T(j)\boldsymbol{X}(j)\}\boldsymbol{X}(j)] = [\boldsymbol{I} + 2k_s\boldsymbol{\Phi}(x, x)]E[\boldsymbol{W}(j)] - 2k_s\boldsymbol{\Phi}(x, d) \quad (24)$$

where $\boldsymbol{I}$ is the identity matrix. With an initial weight vector $\boldsymbol{W}(0)$, $j+1$ iterations of (24) yield

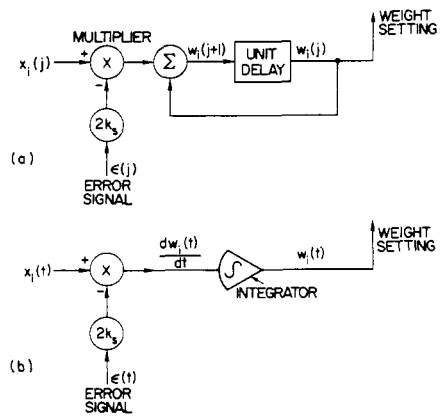

Fig. 7. Block diagram representation of LMS algorithm. (a) Digital realization. (b) Analog realization.

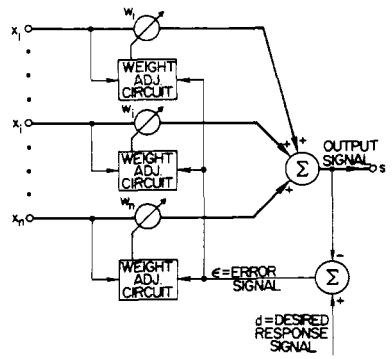

Fig. 8. Analog/digital implementation of LMS weight-adjustment algorithm.

$$E[\boldsymbol{W}(j + 1)] = [\boldsymbol{I} + 2k_s\boldsymbol{\Phi}(x, x)]^{j+1}\boldsymbol{W}(0) - 2k_s\sum_{i=0}^{j}[\boldsymbol{I} + 2k_s\boldsymbol{\Phi}(x, x)]^i\boldsymbol{\Phi}(x, d). \quad (25)$$

Equation (25) may be put in diagonal form by using the appropriate similarity transformation $\boldsymbol{Q}$ for the matrix $\boldsymbol{\Phi}(x, x)$, that is,

$$\boldsymbol{\Phi}(x, x) = \boldsymbol{Q}^{-1}\boldsymbol{E}\boldsymbol{Q}$$

where

$$\boldsymbol{E} \triangleq \begin{bmatrix} e_1 & 0 & \cdots & 0 \\ 0 & e_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e_n \end{bmatrix}$$

is the diagonal matrix of eigenvalues. The eigenvalues are all positive, since $\boldsymbol{\Phi}(x, x)$ is positive definite [see (16)]. Equation (25) may now be expressed as

$$E[\boldsymbol{W}(j + 1)] = [\boldsymbol{I} + 2k_s\boldsymbol{Q}^{-1}\boldsymbol{E}\boldsymbol{Q}]^{j+1}\boldsymbol{W}(0) - 2k_s\sum_{i=0}^{j}[\boldsymbol{I} + 2k_s\boldsymbol{Q}^{-1}\boldsymbol{E}\boldsymbol{Q}]^i\boldsymbol{\Phi}(x, d) = \boldsymbol{Q}^{-1}[\boldsymbol{I} + 2k_s\boldsymbol{E}]^{j+1}\boldsymbol{Q}\boldsymbol{W}(0) - 2k_s\boldsymbol{Q}^{-1}\sum_{i=0}^{j}[\boldsymbol{I} + 2k_s\boldsymbol{E}]^i\boldsymbol{Q}\boldsymbol{\Phi}(x, d). \quad (26)$$

Consider the diagonal matrix $[\boldsymbol{I} + 2k_s\boldsymbol{E}]$. As long as its diagonal terms are all of magnitude less than unity

$$\lim_{j\to\infty} [\boldsymbol{I} + 2k_s\boldsymbol{E}]^{j+1} \to 0$$

and the first term of (26) vanishes as the number of iterations increases. The second term in (26) generally converges to a nonzero limit. The summation factor $\sum_{i=0}^{j}[\boldsymbol{I}+2k_s\boldsymbol{E}]^i$ becomes

$$\lim_{j\to\infty} \sum_{i=0}^{j} [\boldsymbol{I} + 2k_s\boldsymbol{E}]^i = -\frac{1}{2k_s}\boldsymbol{E}^{-1}$$

where the formula for the sum of a geometric series has been used, that is,

$$\sum_{i=0}^{\infty} (1 + 2k_se_p)^i = \frac{1}{1-(1+2k_se_p)} = \frac{-1}{2k_se_p}.$$

Thus, in the limit, (26) becomes

$$\lim_{j\to\infty} E[\boldsymbol{W}(j+1)] = \boldsymbol{Q}^{-1}\boldsymbol{E}^{-1}\boldsymbol{Q}\boldsymbol{\Phi}(x, d)$$
$$= \boldsymbol{\Phi}^{-1}(x, x)\boldsymbol{\Phi}(x, d).$$

Comparison of this result with (19) shows that as the number of iterations increases without limit, the expected value of the weight vector converges to the Wiener solution.

Convergence of the mean of the weight vector to the Wiener solution is insured if and only if the proportionality constant k_s is set within certain bounds. Since the diagonal terms of $[\boldsymbol{I}+2k_s\boldsymbol{E}]$ must all have magnitude less than unity, and since all eigenvalues in $\boldsymbol{E}$ are positive, the bounds on k_s are given by

$$|1 + 2k_se_{\max}| < 1$$

or

$$\frac{-1}{e_{\max}} < k_s < 0 \tag{27}$$

where $e_{\max}$ is the maximum eigenvalue of $\boldsymbol{\Phi}(x, x)$. This convergence condition on k_s can be related to the total input power as follows.

Since

$$e_{\max} \leqq \text{trace}\,[\boldsymbol{\Phi}(x, x)] \tag{28}$$

where

$$\text{trace}\,[\boldsymbol{\Phi}(x, x)] \triangleq E[\boldsymbol{X}^T(j)\boldsymbol{X}(j)]$$
$$= \sum_{i=1}^{n} E[x_i^2] \triangleq \text{total input power},$$

it follows that satisfactory convergence can be obtained with

$$\frac{-1}{\sum_{i=1}^{n} E[x_i^2]} < k_s < 0.$$

In practice, when slow, precise adaptation is desired, k_s is usually chosen such that

$$\frac{-1}{\sum_{i=1}^{n} E[x_i^2]} \ll k_s < 0. \tag{29}$$

It is the opinion of the authors that the assumption of independent successive input samples used in the foregoing convergence proof is overly restrictive. That is, convergence of the mean of the weight vector to the LMS solution can be achieved under conditions of highly correlated input samples. In fact, the computer-simulation experiments described in this paper *do not* satisfy the condition of independence.

Time Constants and Learning Curve with LMS Adaptation

State-variable methods, which are widely used in modern control theory, have been applied by Widrow[1] and Koford and Groner[2] to the analysis of stability and time constants (related to rate of convergence) of the LMS algorithm. Considerable simplifications in the analysis have been realized by expressing transient phenomena of the system adjustments (which take place during the adaptation process) in terms of the normal coordinates of the system. As shown by Widrow,[1] the weight values undergo transients during adaptation. The transients consist of sums of exponentials with time constants given

$$\tau_p = \frac{1}{2(-k_s)e_p},\ p = 1, 2, \cdots, n \tag{30}$$

where e_p is the pth eigenvalue of the input-signal correlation matrix $\boldsymbol{\Phi}(x, x)$.

In the special case when all eigenvalues are equal, all time constants are equal. Accordingly,

$$\tau = \frac{1}{2(-k_s)e}.$$

One very useful way to monitor the progress of an adaptive process is to plot or display its "learning curve." When mean-square error is the performance criterion being used, one can plot the expected mean-square error at each stage of the learning process as a function of the number of adaptation cycles. Since the underlying relaxation phenomenon which takes place in the weight values is of exponential nature, and since from (15) the mean-square error is a quadratic form in the weight values, the transients in the mean-square-error function must also be exponential in nature.

When all the time constants are equal, the mean-square-error learning curve is a pure exponential with a time constant

$$\tau_{\text{mse}} = \frac{\tau}{2} = \frac{1}{4(-k_s)e}.$$

The basic reason for this is that the square of an exponential function is an exponential with half the time constant.

Estimation of the rate of adaptation is more complex when the eigenvalues are unequal.

When actual experimental learning curves are plotted, they are generally of the form of noisy exponentials because of the inherent noise in the adaptation process. The slower the adaptation, the smaller will be the amplitude of the noise apparent in the learning curve.

Misadjustment with LMS Adaptation

All adaptive or learning systems capable of adapting at real-time rates experience losses in performance because their system adjustments are based on statistical averages taken with limited sample sizes. The faster a system adapts, in general, the poorer will be its expected performance.

When the LMS algorithm is used with the basic adaptive element of Fig. 8, the expected level of mean-square error will be greater than that of the Wiener optimum system whose weights are set in accordance with (19). The longer the time constants of adaptation, however, the closer the expected performance comes to the Wiener optimum performance. To get the Wiener performance, i.e., to achieve the minimum mean-square error, one would have to know the input statistics *a priori*, or, if (as is usual) these statistics are unknown, they would have to be measured with an arbitrarily large statistical sample.

When the LMS adaptation algorithm is used, an excess mean-square error therefore develops. A measure of the extent to which the adaptive system is misadjusted as compared to the Wiener optimum system is determined in a performance sense by the ratio of the excess mean-square error to the minimum mean-square error. This dimensionless measure of the loss in performance is defined as the "misadjustment" M. For LMS adaptation of the basic adaptive element, it is shown by Widrow[1] that

$$\text{Misadjustment } M = \frac{1}{2}\sum_{p=1}^{n}\frac{1}{\tau_p}. \tag{31}$$

The value of the misadjustment depends on the time constants (settling times) of the filter adjustment weights. Again, in the special case when all the time constants are equal, *M is proportional to the number of weights and inversely proportional to the time constant.* That is,

$$M = \frac{n}{2\tau} = \frac{n}{4\tau_{\text{mse}}}. \tag{32}$$

Although the foregoing results specifically apply to statistically stationary processes, the LMS algorithm can also be used with nonstationary processes. It is shown by Widrow[2,3] that, under certain assumed conditions, the rate of adaptation is optimized when the loss of performance resulting from adapting too rapidly equals twice the loss in performance resulting from adapting too slowly.

Adaptive Spatial Filtering

If the radiated signals received by the elements of an adaptive antenna array were to consist of signal components plus undesired noise, the signal would be reproduced (and noise eliminated) as best possible in the least-squares sense if the desired response of the adaptive processor were made to be the signal itself. This signal is not generally available for adaptation purposes, however. If it were available, there would be no need for a receiver and a receiving array.

In the adaptive antenna systems to be described here, the desired response signal is provided through the use of an artificially injected signal, the "pilot signal", which is completely known at the receiver and usually generated there. The pilot signal is constructed to have spectral and directional characteristics similar to those of the incoming signal of interest. These characteristics may, in some cases, be known *a priori* but, in general, represent estimates of the parameters of the signal of interest.

Adaptation with the pilot signal causes the array to form a beam in the pilot-signal direction having essentially flat spectral response and linear phase shift within the passband of the pilot signal. Moreover, directional noises impinging on the antenna array will cause reduced array response (nulling) in their directions within their passbands. These notions are demonstrated by experiments which will be described in the following.

Injection of the pilot signal could block the receiver and render useless its output. To circumvent this difficulty, two adaptation algorithms have been devised, the "one-mode" and the "two-mode." The two-mode process alternately adapts on the pilot signal to form the beam and then adapts on the natural inputs with the pilot signal off to eliminate noise. The array output is usable during the second mode, while the pilot signal is off. The one-mode algorithm permits listening at all times, but requires more equipment for its implementation.

The Two-Mode Adaptation Algorithm

Fig. 9 illustrates a method for providing the pilot signal wherein the latter is actually transmitted by an antenna located some distance from the array in the desired look direction. Fig. 10 shows a more practical method for providing the pilot signal. The inputs to the processor are connected either to the actual antenna element outputs (during "mode II"), or to a set of delayed signals derived from the pilot-signal generator (during "mode I"). The filters $\delta_1, \cdots, \delta_K$ (ideal time-delays if the array elements are identical) are chosen to result in a set of input signals identical with those that would appear if the array were actually receiving a radiated plane-wave pilot signal from the desired "look" direction, the direction intended for the main lobe of the antenna directivity pattern.

During adaptation in mode I, the input signals to the adaptive processor derive from the pilot signal, and the desired response of the adaptive processor is the pilot signal

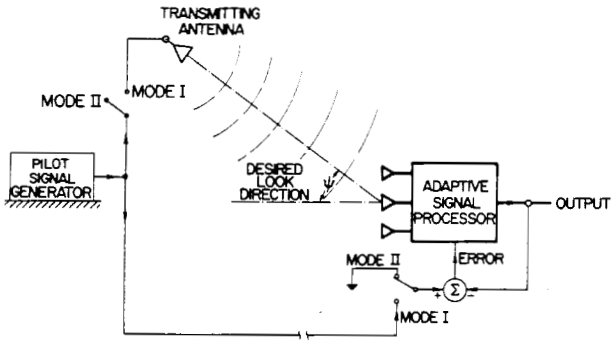

Fig. 9. Adaptation with external pilot-signal generator. Mode I: adaptation with pilot signal present; Mode II: adaptation with pilot signal absent.

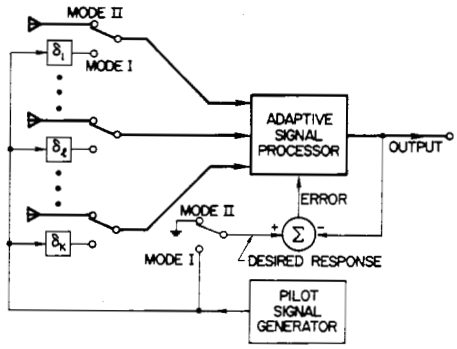

Fig. 10. Two-mode adaptation with internal pilot-signal generator. Mode I: adaptation with pilot signal present; Mode II: adaptation with pilot signal absent.

itself. If a sinusoidal pilot signal at frequency f_0 is used, for example, adapting the weights to minimize mean-square error will force the gain of the antenna array in the look direction to have a specific amplitude and a specific phase shift at frequency f_0.

During adaptation in mode II, all signals applied to the adaptive processor are received by the antenna elements from the actual noise field. In this mode, the adaptation process proceeds to eliminate all received signals, since the desired response is set to zero. Continuous operation in mode II would cause all the weight values to tend to zero, and the system would shut itself off. However, by alternating frequently between mode I and mode II and causing only small changes in the weight vector during each mode of adaptation, it is possible to maintain a beam in the desired look direction and, in addition, to minimize the reception of incident-noise power.

The pilot signal can be chosen as the sum of several sinusoids of differing frequencies. Then adaptation in mode I will constrain the antenna gain and phase in the look direction to have specific values at each of the pilot-signal frequencies. Furthermore, if several pilot signals of different simulated directions are added together, it will be possible to constrain the array gain simultaneously at various frequencies and angles when adapting in mode I. This feature affords some control of the bandwidth and beamwidth in the look direction. The two-mode adaptive process essentially minimizes the mean-square value (the total power)

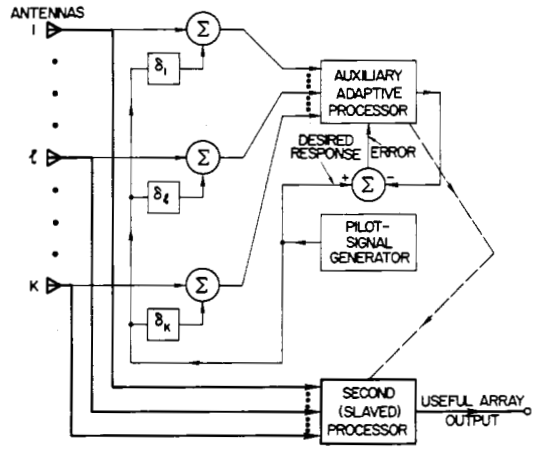

Fig. 11. Single-mode adaptation with pilot signal.

of all signals received by the antenna elements which are uncorrelated with the pilot signals, subject to the constraint that the gain and phase in the beam approximate predetermined values at the frequencies and angles dictated by the pilot-signal components.

The One-Mode Adaptation Algorithm

In the two-mode adaptation algorithm the beam is formed during mode I, and the noises are eliminated in the least-squares sense (subject to the pilot-signal constraints) in mode II. Signal reception during mode I is impossible because the processor is connected to the pilot-signal generator. Reception can therefore take place only during mode II. This difficulty is eliminated in the system of Fig. 11, in which the actions of both mode I and mode II can be accomplished simultaneously. The pilot signals and the received signals enter into an auxiliary, adaptive processor, just as described previously. For this processor, the desired response is the pilot signal $p(t)$. A second weighted processor (linear element) generates the actual array output signal, but it performs no adaptation. Its input signals do not contain the pilot signal. It is slaved to the adaptive processor in such a way that its weights track the corresponding weights of the adapting system, so that it never needs to receive the pilot signal.

In the single-mode system of Fig. 11, the pilot signal is on continuously. Adaptation to minimize mean-square error will force the adaptive processor to reproduce the pilot signal as closely as possible, and, at the same time, to reject as well as possible (in the mean-square sense) all signals received by the antenna elements which are uncorrelated with the pilot signal. Thus the adaptive process forces a directivity pattern having the proper main lobe in the look direction in the passband of the pilot signal (satisfying the pilot signal constraints), and it forces nulls in the directions of the noises and in their frequency bands. Usually, the stronger the noises, the deeper are the corresponding nulls.

Computer Simulation of Adaptive Antenna Systems

To demonstrate the performance characteristics of adaptive antenna systems, many simulation experiments, involving a wide variety of array geometries and signal-

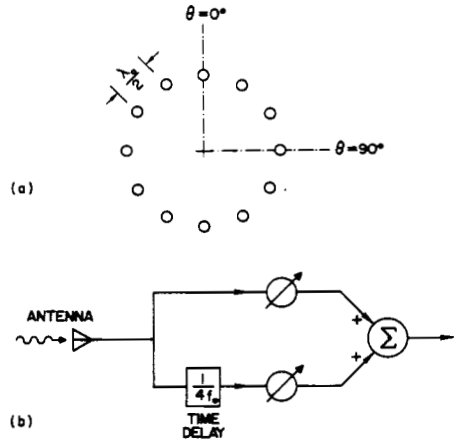

Fig. 12. Array configuration and processing for narrowband experiments.

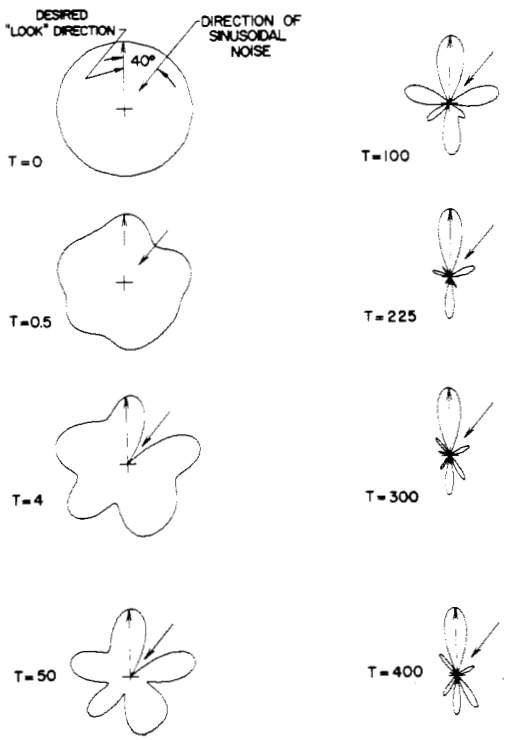

Fig. 13. Evolution of the directivity pattern while learning to eliminate a directional noise and uncorrelated noises. (Array configuration of Fig. 12.) T = number of elapsed cycles of frequency f_0 (total number of adaptations = $20T$).

and noise-field configurations, have been carried out using an IBM 1620-II computer equipped with a digital output plotter.

For simplicity of presentation, the examples outlined in the following are restricted to planar arrays composed of ideal isotropic radiators. In every case, the LMS adaptation algorithm was used. All experiments were begun with the initial condition that all weight values were equal.

Narrowband Processor Experiments

Fig. 12 shows a twelve-element circular array and signal processor which was used to demonstrate the performance of the narrowband system shown in Fig. 4. In the first computer simulation, the two-mode adaptation algorithm was used. The pilot signal was a unit-amplitude sine wave (power = 0.5, frequency f_0) which was used to train the array to look in the $\theta = 0°$ direction. The noise field consisted of a sinusoidal noise signal (of the same frequency and power as the pilot signal) incident at angle $\theta = 40°$, and a small amount of random, uncorrelated, zero-mean, "white" Gaussian noise of variance (power) = 0.1 at each antenna element. In this simulation, the weights were adapted using the LMS two-mode algorithm.

Fig. 13 shows the sequence of directivity patterns which evolved during the "learning" process. These computer-plotted patterns represent the decibel sensitivity of the array at frequency f_0. Each directivity pattern is computed from the set of weights resulting at various stages of adaptation. The solid arrow indicates the direction of arrival of the interfering sine-wave noise source. Notice that the initial directivity pattern is essentially circular. This is due to the symmetry of the antenna array elements and of the initial weight values. A timing indicator T, the number of elapsed cycles of frequency f_0, is presented with each directivity pattern. The total number of adaptations equals $20T$ in these experiments. Note that if $f_0 = 1$ kHz, $T = 1$ corresponds to 1 ms real time; if $f_0 = 1$ MHz, $T = 1$ corresponds to 1 μs, etc.

Several observations can be made from the series of directivity patterns of Fig. 13. Notice that the sensitivity of the array in the look direction is essentially constant during the adaptation process. Also notice that the array sensitivity drops very rapidly in the direction of the sinusoidal noise source; a deep notch in the directivity pattern forms in the noise direction as the adaptation process progresses. After the adaptive transients died out, the array sensitivity in the noise direction was 27 dB below that of the array in the desired look direction.

The total noise power in the array output is the sum of the sinusoidal noise power due to the directional noise source plus the power due to the "white" Gaussian, mutually uncorrelated noise-input signals. The total noise power generally drops as the adaptation process commences, until it reaches an irreducible level.

A plot of the total received noise power as a function of T is shown in Fig. 14. This curve may be called a "learning curve." Starting with the initial weights, the total output noise power was 0.65, as shown in the figure. After adaptation, the total output noise power was 0.01. In this noise field, the signal-to-noise ratio of the array[1] after adaptation was better than that of a single isotropic receiving element by a factor of about 60.

A second experiment using the same array configuration and the two-mode adaptive process was performed to investigate adaptive array performance in the presence of several interfering directional noise sources. In this example, the noise field was composed of five directional sinus-

[1] Signal-to-noise ratio is defined as

$$\text{SNR} = \frac{\text{array output power due to signal}}{\text{array output power due to noise}}.$$

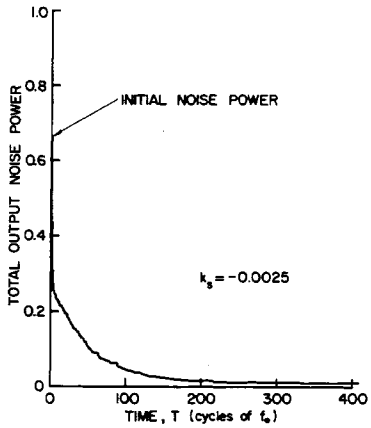

Fig. 14. Learning curve for narrowband system of Fig. 12, with noise from one direction only.

TABLE I

SENSITIVITIES OF ARRAY IN DIRECTIONS OF THE FIVE NOISE SOURCES OF FIG. 15, AFTER ADAPTATION

Noise Direction (degrees)	Noise Frequency (times f_0)	Array Sensitivity in Noise Direction, Relative to Sensitivity in Desired Look Direction (dB)
67	1.10	−26
134	0.95	−30
191	1.00	−28
236	0.90	−30
338	1.05	−38

oidal noises, each of amplitude 0.5 and power 0.125, acting simultaneously, and, in addition, superposed uncorrelated "white" Gaussian noises of power 0.5 at each of the antenna elements. The frequencies of the five directional noises are shown in Table I.

Fig. 15(a) shows the evolution of the directivity pattern, plotted at frequency f_0, from the initial conditions to the finally converged (adapted) state. The latter was achieved after 682 cycles of the frequency f_0. The learning curve for this experiment is shown in Fig. 15(b). The final array sensitivities in the five noise directions relative to the array sensitivity in the desired look direction are shown in Table I. The signal-to-noise ratio was improved by a factor of about 15 over that of a single isotropic radiator. In Fig. 15(b), one can roughly discern a time constant approximately equal to 70 cycles of the frequency f_0. Since there were 20 adaptations per cycle of f_0, the learning curve time constant was approximately $\tau_{mse} = 1400$ adaptations. Within about 400 cycles of f_0, the adaptive process virtually converges to steady state. If f_0 were 1 MHz, 400 μs would be the real-time settling time. The misadjustment for this process can be roughly estimated by using (32), although actually all eigenvalues were not equal as required by this equation:

$$M = \frac{n}{4\tau_{mse}} = \frac{24}{4\tau_{mse}} = \frac{6}{1400} = 0.43 \text{ percent.}$$

This is a very low value of misadjustment, indicating a very slow, precise adaptive process. This is evidenced by the

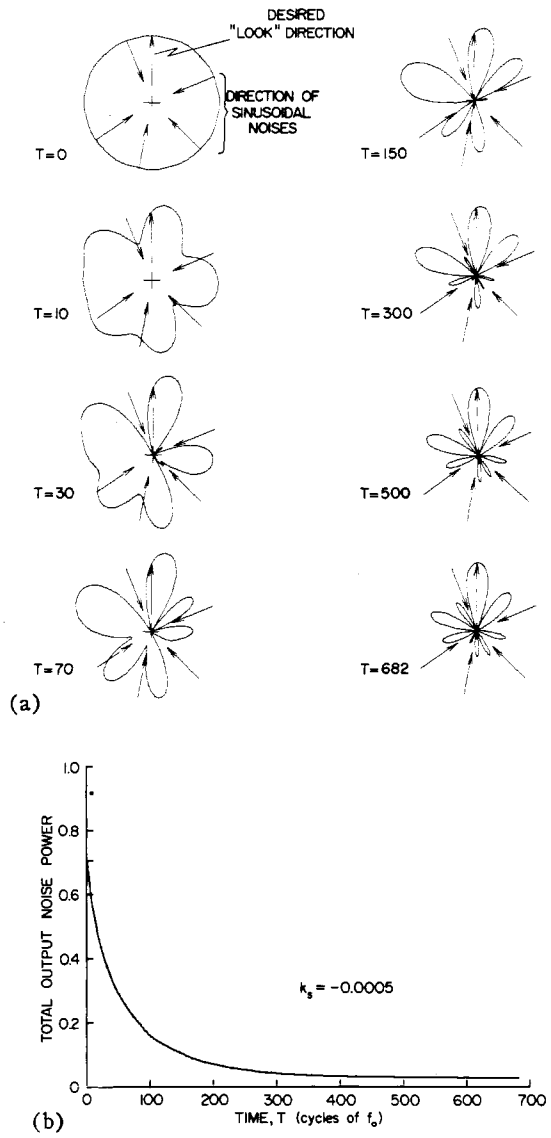

Fig. 15. Evolution of the directivity pattern while learning to eliminate five directional noises and uncorrelated noises. (Array configuration of Fig. 12.) (a) Sequence of directivity patterns during adaptation. (b) Learning curve (total number of adaptations = 20T).

learning curve Fig. 15(b) for this experiment, which is very smooth and noise-free.

Broadband Processor Experiments

Fig. 16 shows the antenna array configuration and signal processor used in a series of computer-simulated broadband experiments. In these experiments, the one-mode or simultaneous adaptation process was used to adjust the weights. Each antenna or element in a five-element circular array was connected to a tapped delay line having five variable weights, as shown in the figure. A broadband pilot signal was used, and the desired look direction was chosen (arbitrarily, for purposes of example) to be $\theta = -13°$. The frequency spectrum of the pilot signal is shown in Fig. 17(a). This spectrum is approximately one octave wide and is centered at frequency f_0. A time-delay increment of $1/(4f_0)$ was used in the tapped delay line, thus providing a delay between adjacent weights of a quarter cycle at fre-

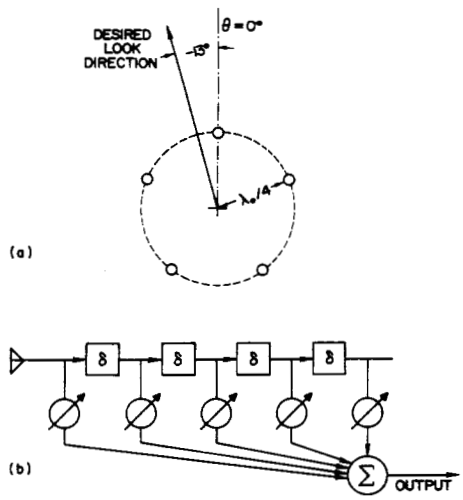

Fig. 16. Array configuration and processing for broadband experiments. (a) Array geometry. (b) Individual element signal processor.

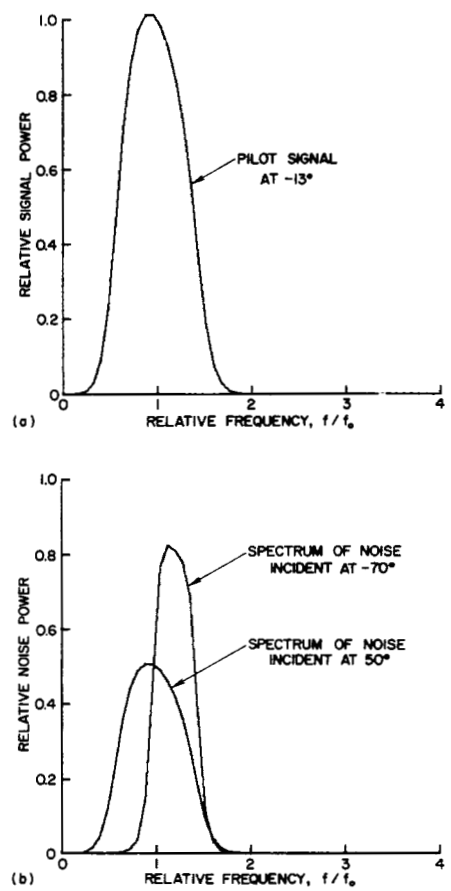

Fig. 17. Frequency spectra for broadband experiments. (a) Pilot signal at $\theta = -13°$. (b) Incident noises at $\theta = 50°$ and $\theta = -70°$.

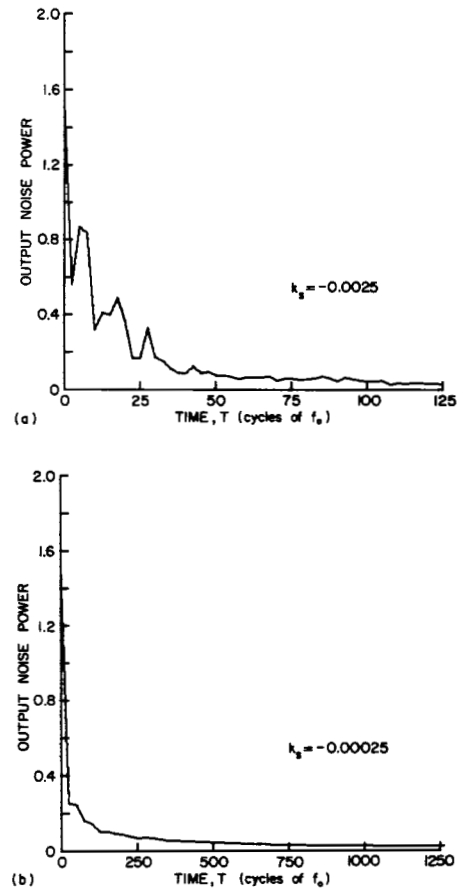

Fig. 18. Learning curves for broadband experiments. (a) Rapid learning ($M = 13$ percent). (b) Slow learning ($M = 1.3$ percent).

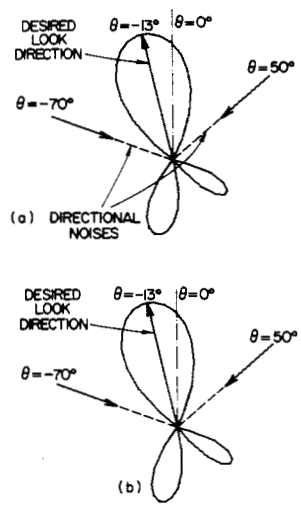

Fig. 19. Comparison of optimum broadband directivity pattern with experimental pattern after former has been adapted during 625 cycles of f_0. (Plotted at frequency f_0.) (a) Optimum pattern. (b) Adapted with $k_s = -0.00025$.

quency f_0, and a total delay-line length of one wavelength at this frequency.

The computer-simulated noise field consisted of two wideband directional noise sources[2] incident on the array at angles $\theta = 50°$ and $\theta = -70°$. Each source of noise had power 0.5. The noise at $\theta = 50°$ had the same frequency spectrum as the pilot signal (though with reduced power); while the noise at $\theta = -70°$ was narrower and centered at a slightly higher frequency. The noise sources were uncorrelated with the pilot signal. Fig. 17(b) shows these frequency spectra. Additive "white" Gaussian noises (mutually uncorrelated) of power 0.0625 were also present in each of the antenna-element signals.

To demonstrate the effects of adaptation rate, the experiments were performed twice, using two different values (-0.0025 and -0.00025) for k_s, the scalar constant in (23). Fig. 18(a) and (b) shows the learning curves obtained under these conditions. The abscissa of each curve is expressed in cycles of f_0, the array center frequency; and, as before, the array was adapted at a rate of twenty times per cycle of f_0. Note that the faster learning curve is a much more noisy one.

Since the statistics of the pilot signal and directional noises in this example are known (having been generated in the computer simulation), it is possible to check measured values of misadjustment against theoretical values. Thus the $\Phi(x, x)$ matrix is known, and its eigenvalues have been computed.[3]

Using (30) and (31) and the known eigenvalues, the misadjustment for the two values of k_s is calculated to give the following values:

k_s	Theoretical Value of M	Experimental Value of M
-0.0025	0.1288	0.134
-0.00025	0.0129	0.0170

The theoretical values of misadjustment check quite well with corresponding measured values.

From the known statistics the optimum (in the least-squares sense) weight vector W_{LMS} can be computed, using (19). The antenna directivity pattern for this optimum weight vector W_{LMS} is shown in Fig. 19(a). This is a broadband directivity pattern, in which the relative sensitivity of the array versus angle of incidence θ is plotted for a broadband received signal having the same frequency spectrum as the pilot signal. This form of directivity pattern has few side lobes, and nulls which are generally not very deep. In Fig. 19(b), the broadband directivity pattern which resulted from adaptation (after 625 cycles of f_0, with $k_s = -0.0025$) is plotted for comparison with the optimum broadband pattern. Note that the patterns are almost indistinguishable from each other.

The learning curves of Fig. 18(a) and (b) are composed of decaying exponentials of various time constants. When k_s is set to -0.00025, in Fig. 18(b), the misadjustment is about 1.3 percent, which is a quite small, but practical value. With this rate of adaptation, it can be seen from Fig. 18(b) that adapting transients are essentially finished after about 500 cycles of f_0. If f_0 is 1 MHz, for example, adaptation could be completed (if the adaptation circuitry is fast enough) in about 500 μs. If f_0 is 1 kHz, adaptation could be completed in about one-half second. Faster adaptation is possible, but there will be more misadjustment. These figures are typical for an adaptive antenna with broadband noise inputs with 25 adaptive weights. For the same level of misadjustment, convergence times increase approximately linearly with the number of weights.[1]

The ability of this adaptive antenna array to obtain "frequency tuning" is shown in Fig. 20. This figure gives the sensitivities of the adapted array (after 1250 cycles of f_0 at $k_s = -0.00025$) as a function of frequency for the desired look direction, Fig. 20(a), and for the two noise directions, Fig. 20(b) and (c). The spectra of the pilot signal and noises are also shown in the figures.

In Fig. 20(a), the adaptive process tends to make the sensitivity of this simple array configuration as close as possible to unity over the band of frequencies where the pilot signal has finite power density. Improved performance might be attained by adding antenna elements and by adding more taps to each delay line; or, more simply, by band-limiting the output to the passband of the pilot signal. Fig. 20(b) and (c) shows the sensitivities of the array in the directions of the noises. Illustrated in this figure is the very striking reduction of the array sensitivity in the directions of the noises, within their specific passbands. The same idea is illustrated by the nulls in the broadband directivity patterns which occur in the noise directions, as shown in Fig. 19. After the adaptive transients subsided in this experiment, the signal-to-noise ratio was improved by the array over that of a single isotropic sensor by a factor of 56.

Implementation

The discrete adaptive processor shown in Figs. 7(a) and 8 could be realized by either a special-purpose digital apparatus or a suitably programmed general-purpose machine. The antenna signals would need analog-to-digital conversion, and then they would be applied to shift registers or computer memory to realize the effects of the tapped delay lines as illustrated in Fig. 5. If the narrowband scheme shown in Fig. 4 is to be realized, the time delays can be implemented either digitally or by analog means (phase shifters) before the analog-to-digital conversion process.

The analog adaptive processor shown in Figs. 7(b) and 8 could be realized by using conventional analog-computer

[2] Broadband directional noises were computer-simulated by first generating a series of uncorrelated ("white") pseudorandom numbers, applying them to an appropriate sampled-data (discrete, digital) filter to achieve the proper spectral characteristics, and then applying the resulting correlated noise waveform to each of the simulated antenna elements with the appropriate delays to simulate the effect of a propagating wavefront.

[3] They are: 10.65, 9.83, 5.65, 5.43, 3.59, 3.44, 2.68, 2.13, 1.45, 1.35, 1.20, 0.99, 0.66, 0.60, 0.46, 0.29, 0.24, 0.20, 0.16, 0.12, 0.01, 0.087, 0.083, 0.075, 0.069.

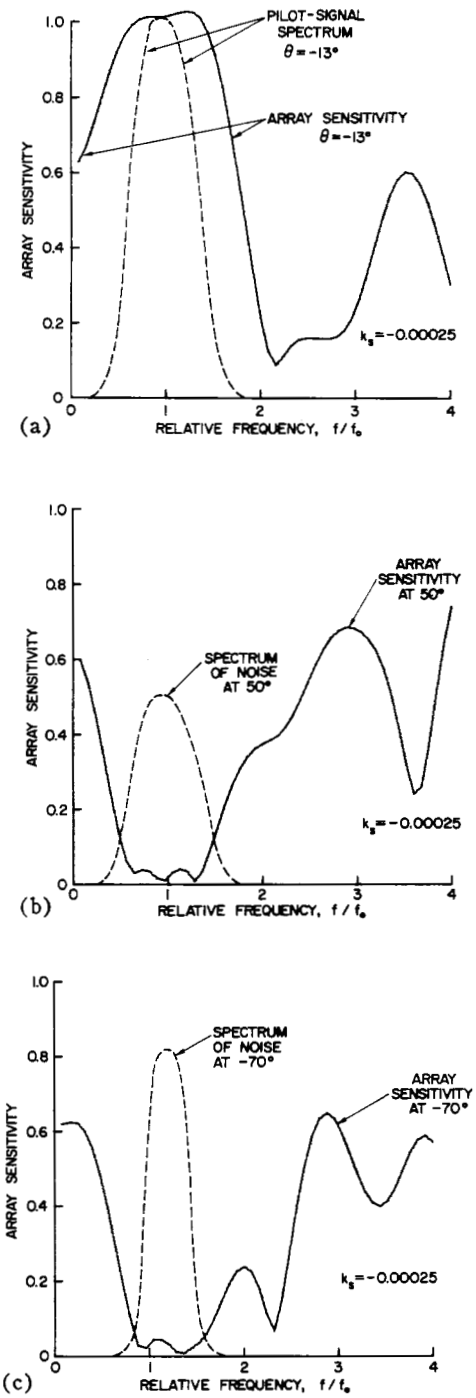

Fig. 20. Array sensitivity versus frequency, for broadband experiment of Fig. 19. (a) Desired look direction, $\theta = -13°$. (b) Sensitivity in one noise direction, $\theta = 50°$. (c) Sensitivity in the other noise direction, $\theta = -70°$.

apparatus, such as multipliers, integrators, summers, etc. More economical realizations that would, in addition, be more suitable for high-frequency operation might use field-effect transistors as the variable-gain multipliers, whose control (gate) signals could come from capacitors used as integrators to form and store the weight values. On the other hand, instead of using a variable resistance structure to form the vector dot products, the same function could be achieved using variable-voltage capacitors, with ordinary capacitors again storing the weight values. The resulting structure would be a capacitive voltage divider rather than a resistive one. Other possible realizations of analog weights include the use of a Hall-effect multiplier combiner with magnetic storage[24] and also the electrochemical memistor of Widrow and Hoff.[25]

Further efforts will be required to improve existing weighting elements and to develop new ones which are simple, cheap, and adaptable according to the requirements of the various adaptation algorithms. The realization of the processor ultimately found to be useful in certain applications may be composed of a combination of analog and digital techniques.

Relaxation Algorithms and Their Implementation

Algorithms other than the LMS procedure described in the foregoing exist that may permit considerable decrease in complexity with specific adaptive circuit implementations. One method of adaptation which may be easy to implement electronically is based on a relaxation algorithm described by Southwell.[26] This algorithm uses the same error signal as used in the LMS technique. An estimated mean-square error formed by squaring and averaging this error signal over a finite time interval is used in determining the proper weight adjustment. The relaxation algorithm adjusts one weight at a time in a cyclic sequence. Each weight in its turn is adjusted to minimize the measured mean-square error. This method is in contrast to the simultaneous adjustment procedure of the LMS steepest-descent algorithm. The relaxation procedure can be shown to produce a misadjustment that increases with the *square* of the number of weights, as opposed to the LMS algorithm whose misadjustment increases only linearly with the number of weights. For a given level of misadjustment, the adaptation settling time of the relaxation process increases with the square of the number of weights.

For implementation of the Southwell relaxation algorithm, the configurations of the array and adaptive processor remain the same, as does the use of the pilot signal. The relaxation algorithm will work with either the two-mode or the one-mode adaptation process. Savings in circuitry may result, in that changes in the adjustments of the weight values depend only upon error measurements and not upon configurations of error measurements and simultaneous input-signal measurements. Circuitry for implementing the LMS systems as shown in Fig. 7(a) and (b) may be more complicated.

The relaxation method may be applicable in cases where the adjustments are not obvious "weight" settings. For example, in a microwave system, the adjustments might be a system of motor-driven apertures or tuning stubs in a waveguide or a network of waveguides feeding an antenna. Or the adjustments may be in the antenna geometry itself. In such cases, the mean-square error can still be measured, but it is likely that it would not be a simple quadratic function of the adjustment parameters. In any event, some very interesting possibilities in automatic optimization are presented by relaxation adaptation methods.

Other Applications and Further Work on Adaptive Antennas

Work is continuing on the proper choice of pilot signals to achieve the best trade-off between response in the desired look direction and rejection of noises. The subject of "null-steering," where the adaptive algorithm causes the nulls of the directivity pattern to track moving noise sources, is also being studied.

The LMS criterion used as the performance measure in this paper minimizes the mean-square error between the array output and the pilot signal waveform. It is a useful performance measure for signal *extraction* purposes. For signal *detection*, however, maximization of array output signal-to-noise ratio is desirable. Algorithms which achieve the maximum SNR solution are also being studied. Goode[27] has described a method for synthesizing the optimal Bayes detector for continuous waveforms using Wiener (LMS) filters. A third criterion under investigation has been discussed by Kelley and Levin[28] and, more recently, applied by Capon *et al.*[29] to the processing of large aperture seismic array (LASA) data. This filter, the maximum-likelihood array processor, is constrained to provide a *distortionless* signal estimate and simultaneously minimize output noise power. Griffiths[30] has discussed the relationship between the maximum likelihood array processor and the Wiener filter for discrete systems.

The examples given have illustrated the ability of the adaptive antenna system to counteract directional interfering noises, whether they are monochromatic, narrowband, or broadband. Although adaptation processes have been applied here exclusively to receiving arrays, they may also be applied to transmitting systems. Consider, for example, an application to aid a low-power transmitter. If a fixed amplitude and frequency pilot signal is transmitted from the receiving site on a slightly different frequency than that of the carrier of the low-power information transmitter, the transmitter array could be adapted (in a receiving mode) to place a beam in the direction of this pilot signal, and, therefore, by reciprocity the transmitting beam would be directed toward the receiving site. The performance of such a system would be very similar to that of the retrodirective antenna systems,[5],[6] although the methods of achieving such performance would be quite different. These systems may be useful in satellite communications.

An additional application of interest is that of "signal seeking." The problem is to find a coherent signal of unknown direction in space, and to find this signal by adapting the weights so that the array directivity pattern receives this signal while rejecting all other noise sources. The desired response or pilot signal for this application is the received signal itself processed through a narrowband filter. The use of the output signal of the adaptive processor to provide its own desired response is a form of unsupervised learning that has been referred to as "bootstrap learning."[31] Use of this adaptation algorithm yields a set of weights which accepts all correlated signals (in the desired passband) and rejects all other received signals. This system has been computer simulated and shown to operate as expected. However, much work of a theoretical and experimental nature needs to be done on capture and rejection phenomena in such systems before they can be reported in detail.

Conclusion

It has been shown that the techniques of adaptive filtering can be applied to processing the output of the individual elements in a receiving antenna array. This processing results in reduced sensitivity of the array to interfering noise sources whose characteristics may be unknown *a priori*. The combination of array and processor has been shown to act as an automatically tunable filter in both space and frequency.

Acknowledgment

The authors are indebted to Dr. M. E. Hoff, Jr., for a number of useful discussions in the early development of these ideas, and to Mrs. Mabel Rockwell who edited the manuscript.

References

[1] B. Widrow, "Adaptive filters I: Fundamentals," Stanford Electronics Labs., Stanford, Calif., Rept. SEL-66-126 (Tech. Rept. 6764-6), December 1966.

[2] J. S. Koford and G. F. Groner, "The use of an adaptive threshold element to design a linear optimal pattern classifier," *IEEE Trans. Information Theory*, vol. IT-12, pp. 42–50, January 1966.

[3] K. Steinbuch and B. Widrow, "A critical comparison of two kinds of adaptive classification networks," *IEEE Trans. Electronic Computers (Short Notes)*, vol. EC-14, pp. 737–740, October 1965.

[4] C. H. Mays, "The relationship of algorithms used with adjustable threshold elements to differential equations," *IEEE Trans. Electronic Computers (Short Notes)*, vol. EC-14, pp. 62–63, February 1965.

[5] L. C. Van Atta, "Electromagnetic reflection," U.S. Patent 2 908 002, October 6, 1959.

[6] "Special Issue on Active and Adaptive Antennas," *IEEE Trans. Antennas and Propagation*, vol. AP-12, March 1964.

[7] C. V. Jakowatz, R. L. Shuey, and G. M. White, "Adaptive waveform recognition," *4th London Symp. on Information Theory*. London: Butterworths, September 1960, pp. 317–326.

[8] L. D. Davisson, "A theory of adaptive filtering," *IEEE Trans. Information Theory*, vol. IT-12, pp. 97–102, April 1966.

[9] E. M. Glaser, "Signal detection by adaptive filters," *IRE Trans. Information Theory*, vol. IT-7, pp. 87–98, April 1961.

[10] F. Bryn, "Optimum signal processing of three-dimensional arrays operating on gaussian signals and noise," *J. Acoust. Soc. Am.*, vol. 34, pp. 289–297, March 1962.

[11] H. Mermoz, "Adaptive filtering and optimal utilization of an antenna," U. S. Navy Bureau of Ships (translation 903 of Ph.D. thesis, Institut Polytechnique, Grenoble, France), October 4, 1965.

[12] S. W. W. Shor, "Adaptive technique to discriminate against coherent noise in a narrow-band system," *J. Acoust. Soc. Am.*, vol. 39, pp. 74–78, January 1966.

[13] B. Widrow and M. E. Hoff, Jr., "Adaptive switching circuits," *IRE WESCON Conv. Rec.*, pt. 4, pp. 96–104, 1960.

[14] N. G. Nilsson, *Learning Machines*. New York: McGraw-Hill, 1965.

[15] B. Widrow and F. W. Smith, "Pattern-recognizing control systems," *1963 Computer and Information Sciences (COINS) Symp. Proc.* Washington, D.C.: Spartan, 1964.

[16] L. R. Talbert *et al.*, "A real-time adaptive speech-recognition system," Stanford Electronics Labs., Stanford University, Stanford, Calif., Rept. SEL 63-064 (Tech. Rept. 6760-1), May 1963.

[17] F. W. Smith, "Design of quasi-optimal minimum time controllers," *IEEE Trans. Automatic Control*, vol. AC-11, pp. 71–77, January 1966.

[18] J. P. Burg, "Three-dimensional filtering with an array of seismome-

ters," *Geophysics*, vol. 29, pp. 693–713, October 1964.
[19] J. F. Claerbout, "Detection of *P* waves from weak sources at great distances," *Geophysics*, vol. 29, pp. 197–211, April 1964.
[20] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat.*, vol. 22, pp. 400–407, March 1951.
[21] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, vol. 23, pp. 462–466, March 1952.
[22] A. Dvoretzky, "On stochastic approximation," *Proc. 3rd Berkeley Symp. on Math. Stat. and Prob.*, J. Neyman, Ed. Berkeley, Calif.: University of California Press, 1956, pp. 39–55.
[23] B. Widrow, "Adaptive sampled-data systems," *Proc. 1st Internat'l Congress of the Internat'l Federation of Automatic Control* (Moscow, 1960). London: Butterworths, 1960.
[24] D. Gabor, W. P. L. Wilby, and R. Woodcock, "A universal nonlinear filter predictor and simulator which optimizes itself by a learning process," *Proc. IEE* (London), vol. 108 B, July 1960.
[25] B. Widrow and M. E. Hoff, Jr., "Generalization and information storage in networks of adaline 'neurons'," in *Self Organizing Systems 1962*, M. C. Yovits, G. T. Jacobi, and G. D. Goldstein, Eds. Washington, D. C.: Spartan, 1962, pp. 435–461.
[26] R. V. Southwell, *Relaxation Methods in Engineering Science*. London: Oxford University Press, 1940.
[27] B. B. Goode, "Synthesis of a nonlinear Bayes detector for Gaussian signal and noise fields using Wiener filters," *IEEE Trans. Information Theory (Correspondence)*, vol. IT-13, pp. 116–118, January 1967.
[28] E. J. Kelley and M. J. Levin, "Signal parameter estimation for seismometer arrays," M.I.T. Lincoln Lab., Lexington, Mass., Tech. Rept. 339, January 8, 1964.
[29] J. Capon, R. J. Greenfield, and R. J. Kolker, "Multidimensional maximum-likelihood processing of a large aperture seismic array," *Proc. IEEE*, vol. 55, pp. 192–211, February 1967.
[30] L. J. Griffiths, "A comparison of multidimensional Wiener and maximum-likelihood filters for antenna arrays," *Proc. IEEE (Letters)*, vol. 55, pp. 2045–2047, November 1967.
[31] B. Widrow, "Bootstrap learning in threshold logic systems," presented at the American Automatic Control Council (Theory Committee), IFAC Meeting, London, England, June 1966.

Corrections

The authors of the papers below have brought the following to the attention of the Editor.

Radio Frequency Power Measurements[1]

On page 842 in the example given in Section 3.4 the uncertainty in the comparison measurement should have been given as "± 8 percent."

In Fig. 12, page 842, Γ_2 should have been

$$\Gamma_2 = a_2/b_2.$$

Equation (11), page 842, should have read

$$b_2 = \Gamma_{Ge} a_2 + b_{Ge}.$$

Equation (12), page 842, should have read

$$\Gamma_{Ge} = S_{22} - (S_{21}S_{32})/S_{31}.$$

In the example given in Section 3.4.1, page 843, the uncertainty in the direct measurement should have been given as "one percent for a Γ_M equal to $0.2(\sigma_M = 1.5)$."

On lines 5 and 6, last paragraph, Section 3.4.1, page 843, the reflection coefficient looking into arm 2 should have been given as

$$\Gamma_{2i} = \Gamma_{Ge} = S_{22} - (S_{21}S_{32})/S_{31}.$$

Line 3, paragraph 2, Section 4.2, page 845, should have read "magnitude of the mount reflection coefficient $|\Gamma_M|$."

On page 850, the following reference should be added to the Bibliography under *Miscellaneous:*

[114] I. A. Harris, "The theory and design of coaxial resistor mounts for the frequency band 0–4000 Mc/s," *Proc. IEE (London)*, vol. 103, pt. C, pp. 1–10, March 1956.

Manuscript received September 18, 1967.
[1] A. Y. Rumfelt and L. B. Elwell, *Proc. IEEE*, vol. 55, pp. 837–850, June 1967.

Frequency Stabilization of Gas Lasers[1]

On page 1026, reference [44] should have read:

[44] P. W. Smith, "On the stabilization of a high-power single-frequency laser," to be published in *IEEE J. Quantum Electronics*.

Manuscript received September 14, 1967.
[1] G. Birnbaum, *Proc. IEEE*, vol. 55, pp. 1015–1026, June 1967.

Data Structure and Man-Machine Communication for Network Problems[1]

The manuscript received date should have read July 5, 1966, instead of August 31, 1966.

Manuscript received September 21, 1967.
[1] D. S. Evans and J. Katzenelson, *Proc. IEEE*, vol. 55, pp. 1135–1144, July 1967.

Reverse Recovery Processes in Silicon Power Rectifiers[1]

On page 1332, right column, the third line of the last paragraph was inadvertently omitted. The paragraph should have read as follows.

"Hitherto, reverse switching processes[1],[2] seem to have been treated mostly in consideration of the conditions of *communication engineering. Therefore, low-level injections* were assumed."

Manuscript received September 25, 1967.
[1] H. Benda and E. Spenke, *Proc. IEEE*, vol. 55, pp. 1331–1354, August 1967.