Chapter One

Introduction and Channel Models

ADAPTIVE antenna arrays have long been an attractive solution to a plethora of problems related to signal detection and estimation. An array of antenna elements can overcome the directivity and beamwidth limitations of a single antenna element, and when it is combined with methods from statistical detection and estimation and control theory, a self-adjusting or adaptive system emerges. This key capability was recognized in 1967 by Widrow and his colleagues in their publication in the *IEEE Proceedings,* with which this book opens. This paper offers a valuable introduction to the adaptive antenna concepts.

A smart antenna system relies heavily on the spatial characteristics of the operational environment to improve the output signal. In order to study the performance of adaptive algorithms in radio operational environments (Chapters 2 and 3), it is essential to employ suitable channel models that provide both spatial and temporal information. For that reason, three papers are included in this chapter. There is still a lot of work to be done in terms of characterizing the radio channel and producing propagation models capable of providing all the information needed to efficiently study wideband systems that also exploit the spatial dimension. This need was recently underlined by the international standardization organisations, and several research activities are already under way (e.g., subgroup on spatial propagation models of the COST—European Union Forum for Cooperative Scientific Research—Action 259).

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

Reprinted from *IEEE Proceedings*, Vol 55, No. 12, pp. 2143–2159, December 1967.

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 $\theta=24^\circ$, 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

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

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 $\psi=\sin^{-1}(2\delta f_0)=15^\circ$.

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 $\theta=0^\circ$.

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

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

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

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

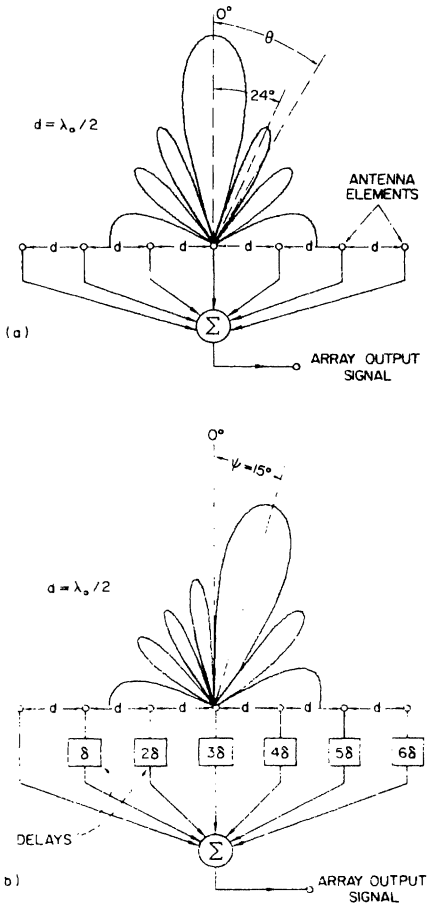


Fig. 1. Directivity pattern for a linear array. (a) Simple array. (b) Delays added.

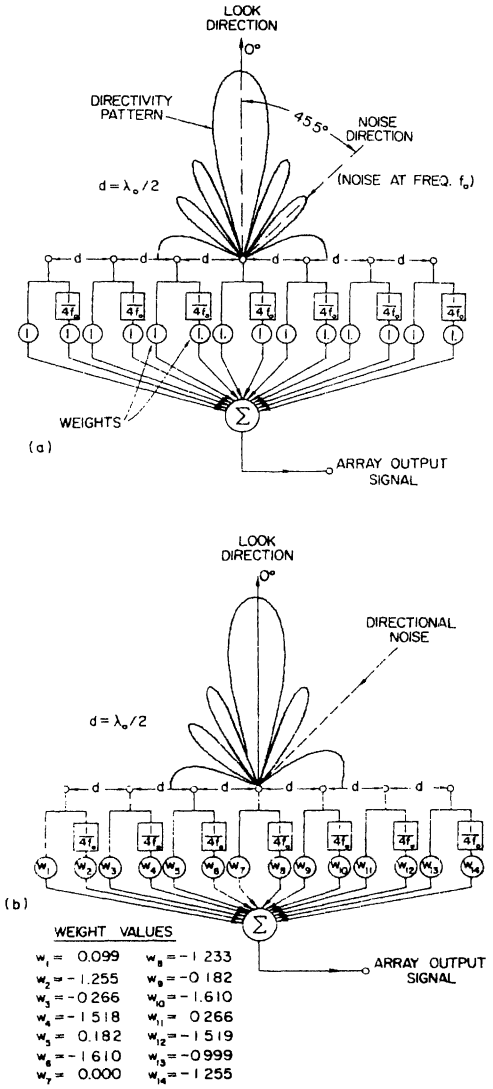


Fig. 2. Directivity pattern of linear array. (a) With equal weighting. (b) With weighting for noise elimination.

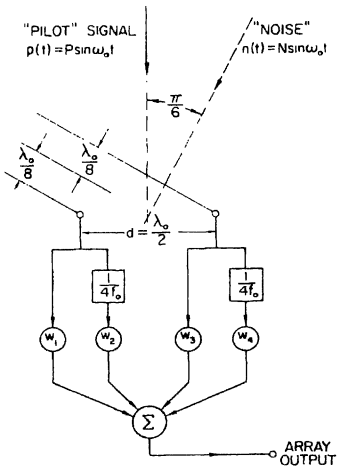


Fig. 3. Array configuration for noise elimination example.

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.

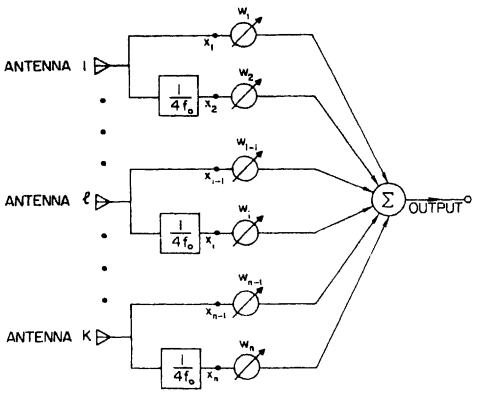


Fig. 4. Adaptive array configuration for receiving narrowband signals.

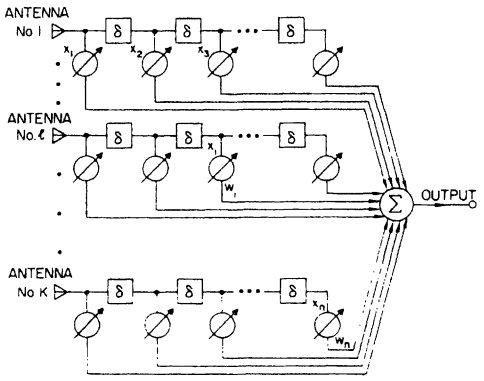


Fig. 5. Adaptive array configuration for receiving broadband signals.

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 $\varepsilon(j)$:

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

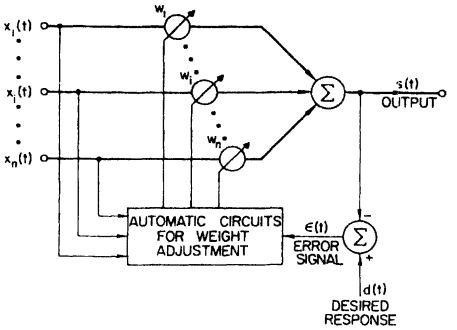


Fig. 6. Basic adaptive element.

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

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

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 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

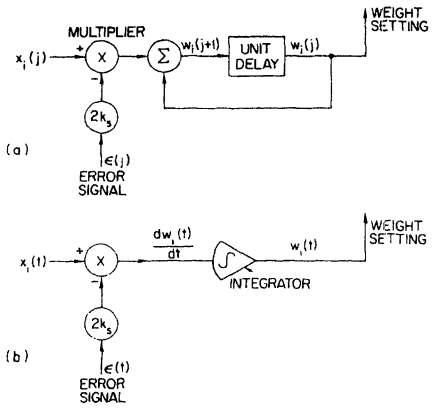


Fig. 7. Block diagram representation of LMS algorithm. (a) Digital realization. (b) Analog realization.

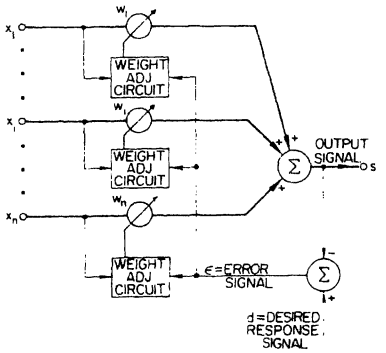


Fig. 8. Analog digital implementation of LMS weight-adjustment algorithm.

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

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

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

Consider the diagonal matrix $[I + 2k_sE]$. As long as its diagonal terms are all of magnitude less than unity

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

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 $\sum_{i=0}^{j}[I+2k_sE]^i$ becomes

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

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

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

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 $[I+2k_sE]$ must all have magnitude less than unity, and since all eigenvalues in 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 $\Phi(x, x)$. This convergence condition on k_s can be related to the total input power as follows.

Since

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

where

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

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

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

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

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

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

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[23] 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

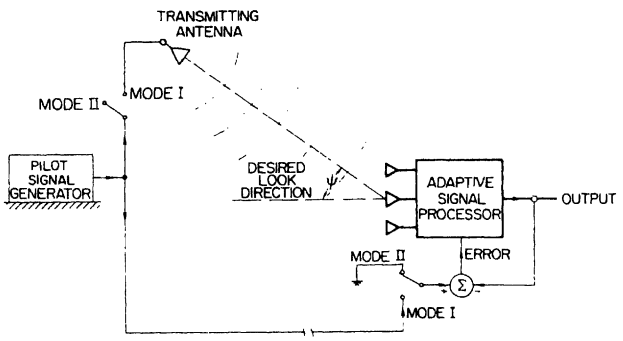


Fig. 9. Adaptation with external pilot-signal generator. Mode I: adaptation with pilot signal present; Mode II: adaptation with pilot signal absent.

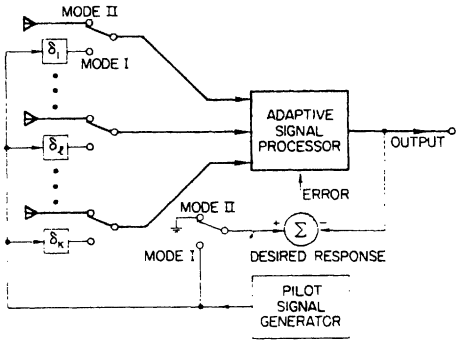


Fig. 10. Two-mode adaptation with internal pilot-signal generator. Mode I: adaptation with pilot signal present; Mode II: adaptation with pilot signal absent.

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)

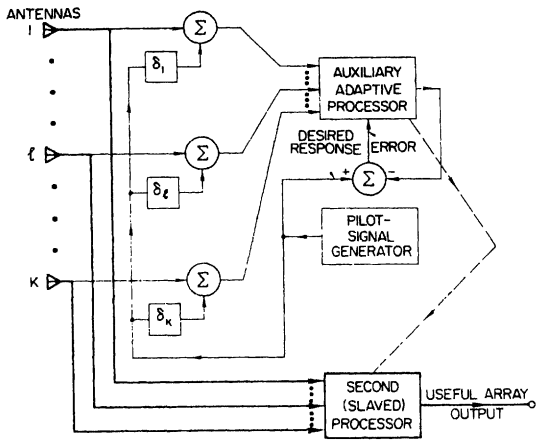


Fig. 11. Single-mode adaptation with pilot signal.

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-

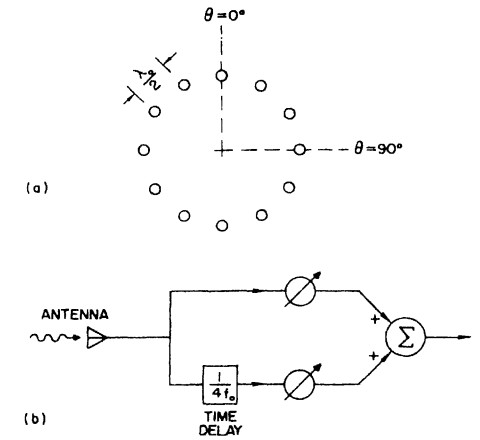


Fig. 12. Array configuration and processing for narrowband experiments.

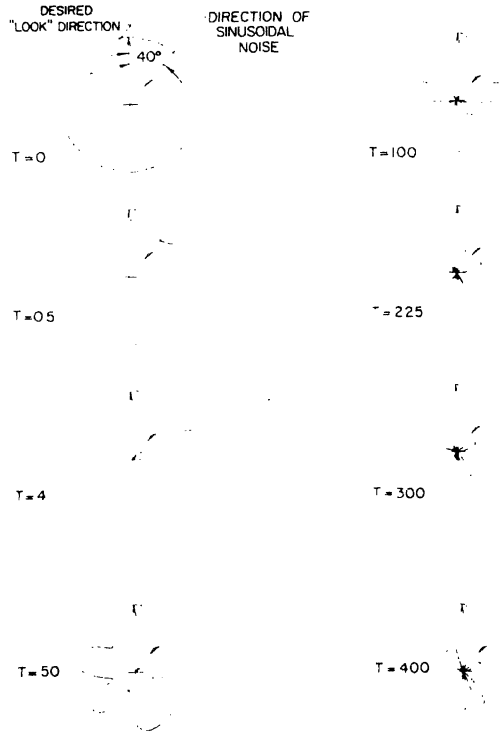


Fig. 13. Evolution of the directivity pattern while learning to eliminate a directional noise and uncorrelated noises. (Array configuration of Fig. 12.) T = number of elapsed cycles of frequency f_0 (total number of adaptations = $20T$).

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

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

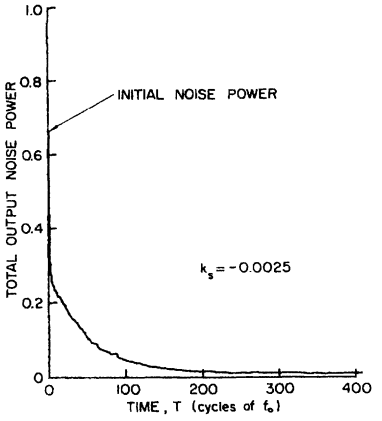


Fig. 14. Learning curve for narrowband system of Fig. 12, with noise from one direction only.

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

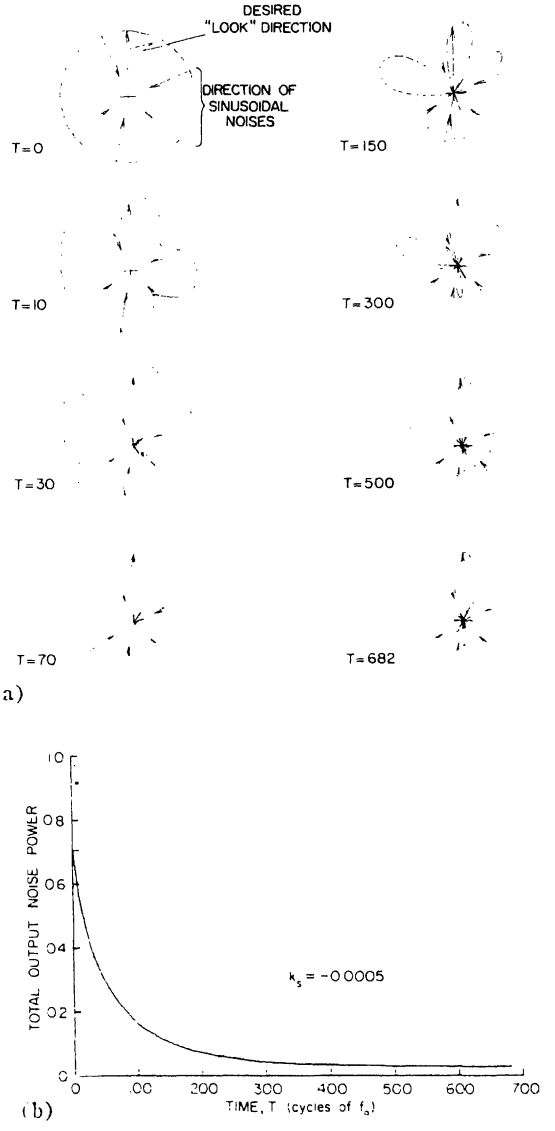


Fig. 15. Evolution of the directivity pattern while learning to eliminate five directional noises and uncorrelated noises. (Array configuration of Fig. 12.) (a) Sequence of directivity patterns during adaptation. (b) Learning curve (total number of adaptations = 20T).

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-

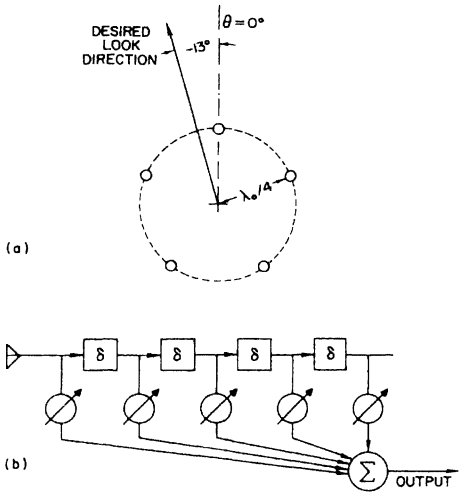


Fig. 16. Array configuration and processing for broadband experiments. (a) Array geometry. (b) Individual element signal processor.

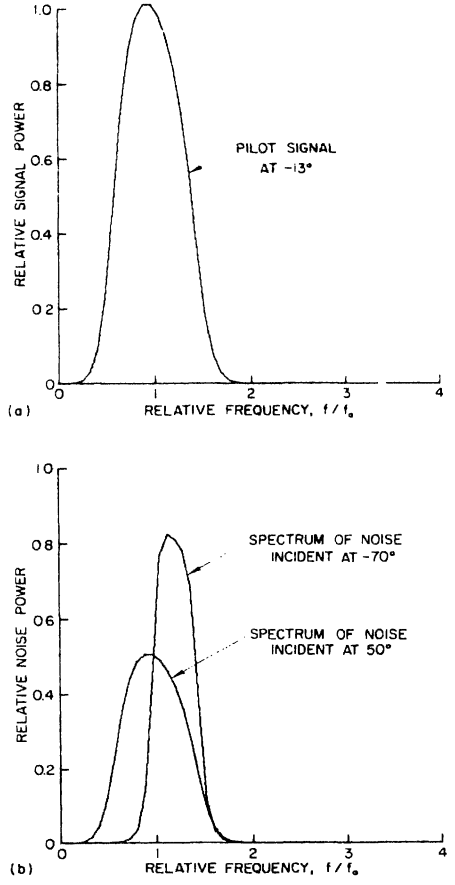


Fig. 17. Frequency spectra for broadband experiments. (a) Pilot signal at $\theta = -13°$. (b) Incident noises at $\theta = 50°$ and $\theta = -70°$.

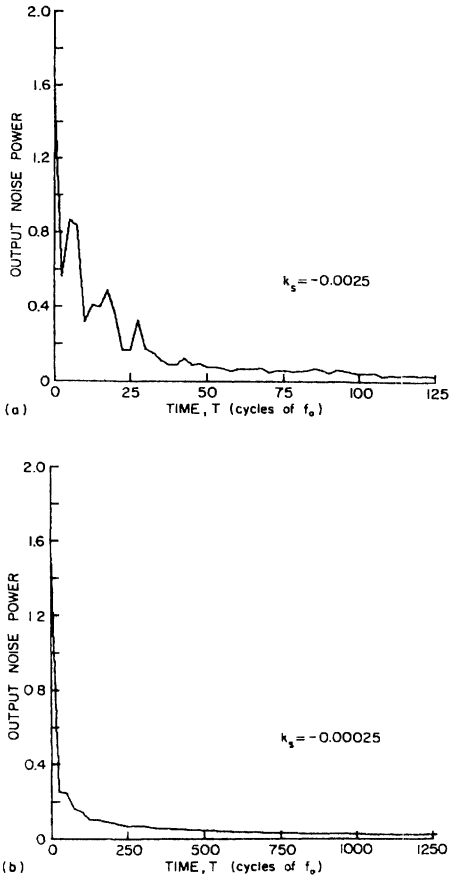


Fig. 18. Learning curves for broadband experiments. (a) Rapid learning ($M = 13$ percent). (b) Slow learning ($M = 1.3$ percent).

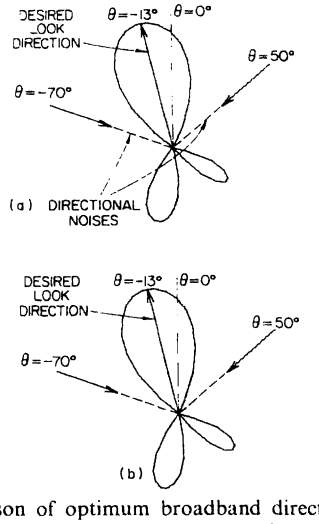


Fig. 19. Comparison of optimum broadband directivity pattern with experimental pattern after former has been adapted during 625 cycles of f_0. (Plotted at frequency f_0.) (a) Optimum pattern. (b) Adapted with $k_s = -0.00025$.

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.[11]

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.

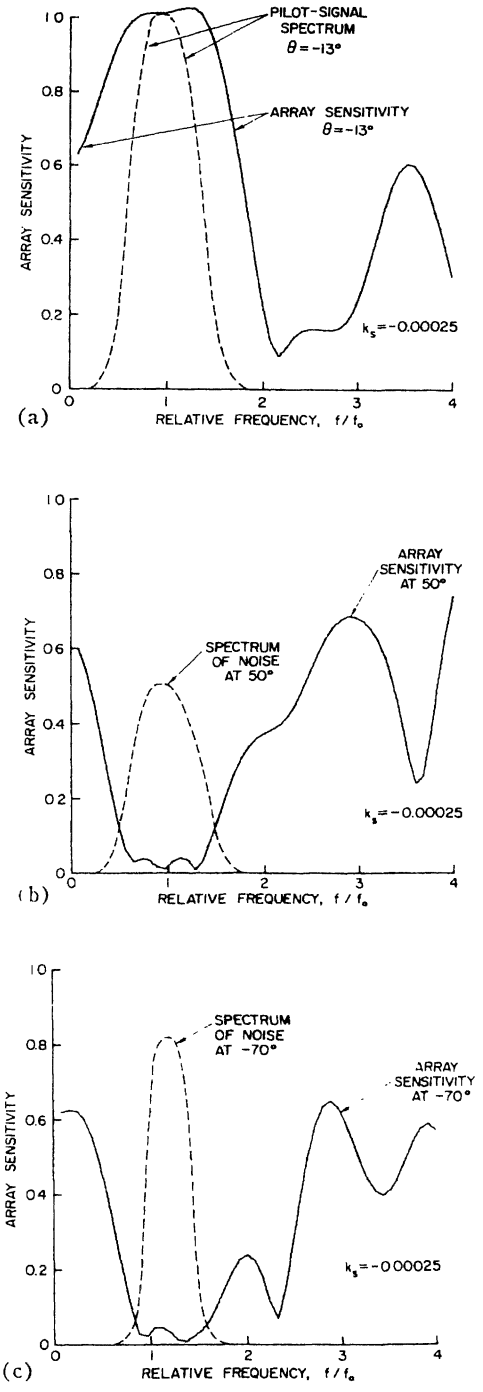


Fig. 20. Array sensitivity versus frequency, for broadband experiment of Fig. 19. (a) Desired look direction, $\theta = -13$. (b) Sensitivity in one noise direction, $\theta = 50$. (c) Sensitivity in the other noise direction, $\theta = -70$.

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. Pattent 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.

Abstract

Throughout the history of wireless communications, spatial antenna diversity has been important in improving the radio link between wireless users. Historically, microscopic antenna diversity has been used to reduce the fading seen by a radio receiver, whereas macroscopic diversity provides multiple listening posts to ensure that mobile communication links remain intact over a wide geographic area. In recent years, the concepts of spatial diversity have been expanded to build foundations for emerging technologies, such as smart (adaptive) antennas and position location systems. Smart antennas hold great promise for increasing the capacity of wireless communications because they radiate and receive energy only in the intended directions, thereby greatly reducing interference. To properly design, analyze, and implement smart antennas and to exploit spatial processing in emerging wireless systems, accurate radio channel models that incorporate spatial characteristics are necessary. In this tutorial, we review the key concepts in spatial channel modeling and present emerging approaches. We also review the research issues in developing and using spatial channel models for adaptive antennas.

Overview of Spatial Channel Models for Antenna Array Communication Systems

RICHARD B. ERTEL AND PAULO CARDIERI, VIRGINIA POLYTECHNIC INSTITUTE
KEVIN W. SOWERBY, UNIVERSITY OF AUCKLAND, NEW ZEALAND
THEODORE S. RAPPAPORT AND JEFFREY H. REED,
VIRGINIA POLYTECHNIC INSTITUTE

With the advent of antenna array systems for both interference cancellation and position location applications comes the need to better understand the spatial properties of the wireless communications channel. These spatial properties of the channel will have an enormous impact on the performance of antenna array systems; hence, an understanding of these properties is paramount to effective system design and evaluation.

The challenge facing communications engineers is to develop realistic channel models that can efficiently and accurately predict the performance of a wireless system. It is important to stress here that the level of detail about the environment a channel model must provide is *highly dependent* on the type of system under consideration. To predict the performance of single-sensor narrowband receivers, it may be acceptable to consider only the received signal power and/or time-varying amplitude (fading) distribution of the channel. However, for emerging wideband multisensor arrays, in addition to signal power level, information regarding the signal multipath delay and angle of arrival (AOA) is needed.

Classical models provide information about signal power level distributions and Doppler shifts of the received signals. These models have their origins in the early days of cellular radio when wideband digital modulation techniques were not readily available. As shown subsequently, many of the emerging spatial models in the literature utilize the fundamental principles of the classical channel models. However, modern spatial channel models build on the classical understanding of fading and Doppler spread, and incorporate additional concepts such as time delay spread, AOA, and adaptive array antenna geometries.

In this article, we review the fundamental channel models that have led to the present-day theories of spatial diversity from both mobile user and base station perspectives. The evolution of these models has paralleled that of cellular systems. Early models only accounted for amplitude and time-varying properties of the channel. These models were then enhanced by adding time delay spread information, which is important when dealing with digital transmission performance. Now, with the introduction of techniques and features that depend on the spatial distribution of the mobiles, spatial information is required in the channel models. As shown in the next sections, more accurate models for the distribution of the scatterers surrounding the mobile and base station are needed. The differentiation between the mobile and base station is important. Classical work has demonstrated that models must account for the physical geometry of scattering objects in the vicinity of the antenna of interest. The number and locations of these scattering objects are dependent on the heights of the antennas, particularly regarding the local environment.

This article, then, explores some of the emerging models for spatial diversity and adaptive antennas, and includes the physical mechanisms and motivations behind the models. A literature survey of existing RF channel measurements with AOA information is also included. The article concludes with a summary and suggestions for future research.

Wireless Multipath Channel Models

This section describes the physical properties of the wireless communication channel that must be modeled. In a wireless system, a signal transmitted into the channel interacts with the environment in a very complex way. There are reflections from large objects, diffraction of the electromagnetic waves

This work was partially supported by the DARPA GloMo program, Virginia Tech's Federal Highways Research Center of Excellence, Virginia Tech's Bradley Foundation, the Brazilian National Science Council — CNPq, and NSF Presidential Faculty Fellowship.

Reprinted from *IEEE Personal Communications Magazine*, Vol 5, No. 1, pp. 10–22, February 1998.

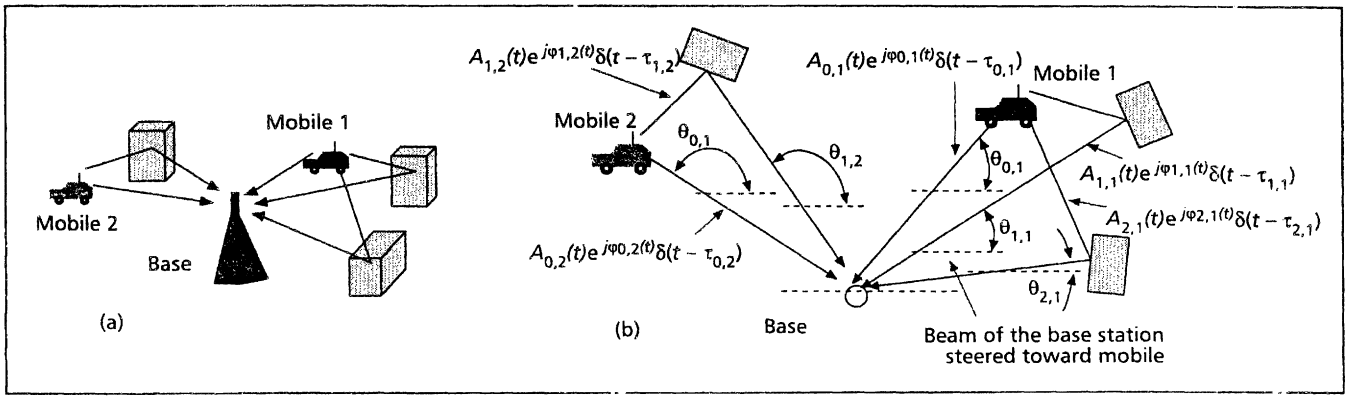


Figure 1. *Multipath propagation channel: a) side view; b) top view.*

around objects, and signal scattering. The result of these complex interactions is the presence of many signal components, or *multipath* signals, at the receiver. Another property of wireless channels is the presence of Doppler shift, which is caused by the motion of the receiver, the transmitter, and/or any other objects in the channel. A simplified pictorial of the multipath environment with two mobile stations is shown in Fig. 1. Each signal component experiences a different path environment, which will determine the amplitude $A_{l,k}$, carrier phase shift $\varphi_{l,k}$, time delay $\tau_{l,k}$, AOA $\theta_{l,k}$, and Doppler shift f_d of the lth signal component of the kth mobile. In general, each of these signal parameters will be time-varying.

The early classical models, which were developed for narrowband transmission systems, only provide information about signal amplitude level distributions and Doppler shifts of the received signals. These models have their origins in the early days of cellular radio [1–4] when wideband digital modulation techniques were not readily available.

As cellular systems became more complex and more accurate models were required, additional concepts, such as time delay spread, were incorporated into the model. Representing the RF channel as a time-variant channel and using a baseband complex envelope representation, the channel impulse response for mobile 1 has classically been represented as [5]

$$h_1(t,\tau)=\sum_{l=0}^{L(t)-1} A_{l,1}(t)e^{j\varphi_{l,1}(t)}\delta\big(t-\tau_{l,1}(t)\big) \qquad (1)$$

where $L(t)$ is the number of multipath components and the other variables have already been defined. The amplitude $A_{l,k}$ of the multipath components is usually modeled as a Rayleigh distributed random variable, while the phase shift $\varphi_{l,k}$ is uniformly distributed.

The time-varying nature of a wireless channel is caused by the motion of objects in the channel. A measure of the time rate of change of the channel is the *Doppler power spectrum*, introduced by M. J. Gans in 1972 [2]. The Doppler power spectrum provides us with statistical information on the variation of the frequency of a tone received by a mobile traveling at speed v. Based on the flat fading channel model developed by R. H. Clarke in 1968, Gans assumed that the received signal at the mobile station came from all directions and was uniformly distributed. Under these assumptions and for a $\lambda/4$ vertical antenna, the Doppler power spectrum is given by [5]

$$S(f)=\begin{cases}\dfrac{1.5}{\pi f_m\sqrt{1-\left(\dfrac{f-f_c}{f_m}\right)^2}} & |f-f_c|<f_m\\ 0 & \textit{elsewhere}\end{cases}$$

where f_m is the maximum Doppler shift given by v/λ where λ is the wavelength of the transmitted signal at frequency f_c.

Figure 2 shows the received signals at the base station, assuming that mobiles 1 and 2 have transmitted narrow pulses at the same time. Also shown is the output of an antenna array system adapted to mobile 1.

The channel model in Eq. 1 does not consider the AOA of each multipath component shown in Figs. 1 and 2. For narrowband signals, the AOA may be included into the vector channel impulse response using

$$\vec{h}_1(t,\tau)=\sum_{l=0}^{L(t)-1} A_{l,1}(t)e^{j\varphi_{l,1}(t)}\vec{a}\big(\theta_l(t)\big)\delta\big(t-\tau_{l,1}(t)\big) \qquad (2)$$

where $\vec{a}(\theta_l(t))$ is the *array response vector*. The array response vector is a function of the array geometry and AOA. Figure 3 shows the case for an arbitrary array geometry when the array and signal are restricted to two-dimensional space. The resulting array response vector is given by

$$a\big(\theta_l(t)\big)=\begin{bmatrix}\exp(-j\psi_{l,1})\\ \exp(-j\psi_{l,2})\\ \exp(-j\psi_{l,3})\\ \cdots\\ \exp(-j\psi_{l,m})\end{bmatrix}$$

where $\psi_{l,i}(t)=[x_i\cos(\theta_l(t))+y_i\sin(\theta_l(t))]\beta$ and $\beta=2\pi/\lambda$ is the wavenumber.

The spatial channel impulse response given in Eq. 2 is a summation of several multipath components, each of which has its own amplitude, phase, and AOA. The distribution of these parameters is dependent on the type of environment. In particular, the angle spread of the channel is known to be a function of both the environment and the base station antenna heights. In the next section, we describe macrocell and microcell environments and discuss how the environment affects the signal parameters.

Macrocell vs. Microcell

Macrocell Environment – Figure 4 shows the channel on the forward link for a macrocell environment. It is usually assumed that the scatterers surrounding the mobile station are about the same height as or are higher than the mobile. This implies that the received signal at the mobile antenna arrives from all directions after bouncing from the surrounding scatterers as illustrated in Fig. 4.

Under these conditions, Gans' assumption that the AOA is uniformly distributed over $[0, 2\pi]$ is valid. The classical

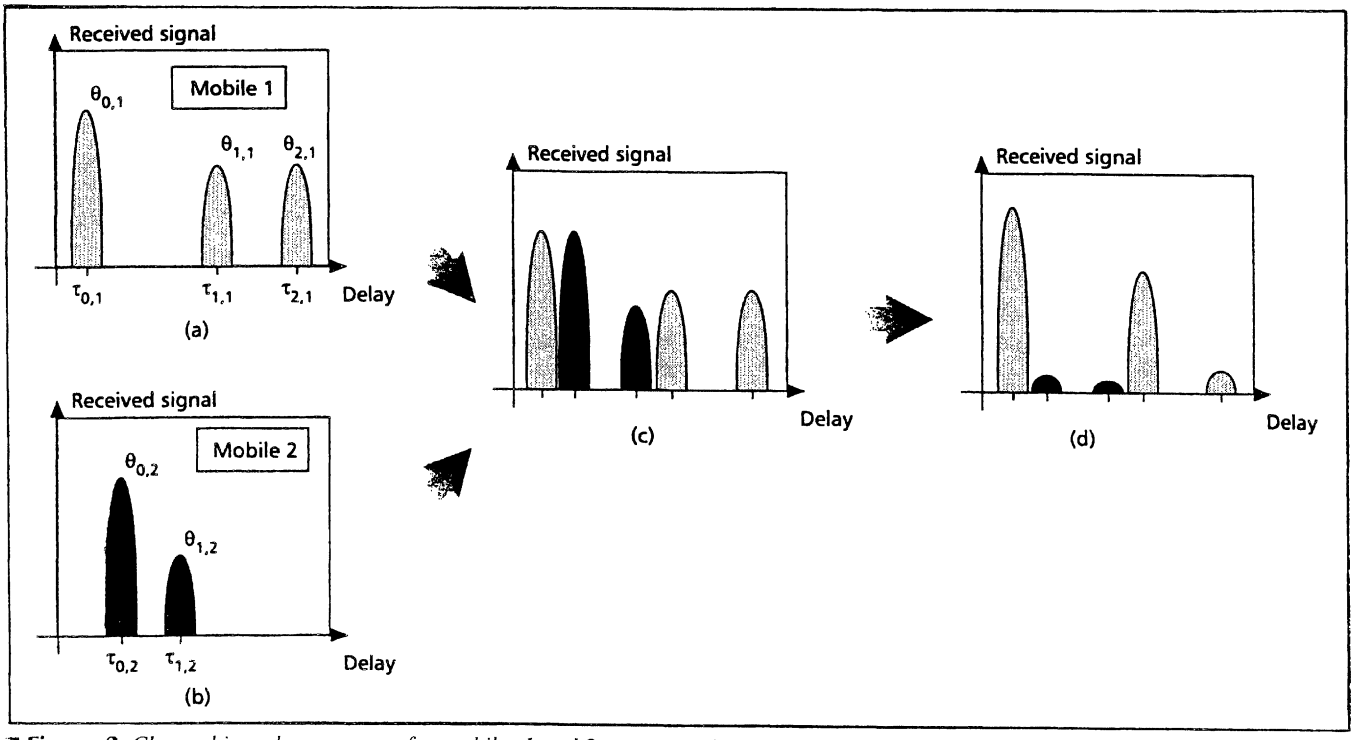


Figure 2. *Channel impulse responses for mobiles 1 and 2: a) received signal from mobile 1 to the base station; b) received signal from mobile 2 to the base station; c) combined received signal from mobiles 1 and 2 at the base station; d) received signal at the base station when a beam steered toward mobile 1 is employed.*

Rayleigh fading envelope with deep fades approximately $\lambda/2$ apart emanates from this model [5].

However, the AOA of the received signal at the base station is quite different. In a macrocell environment, typically, the base station is deployed higher than the surrounding scatterers. Hence, the received signals at the base station result from the scattering process in the vicinity of the mobile station, as shown in Fig. 5. The multipath components at the base station are restricted to a smaller angular region, θ_{BW}, and the distribution of the AOA is no longer uniform over $[0.2\pi]$. Other AOA distributions are considered later in this article.

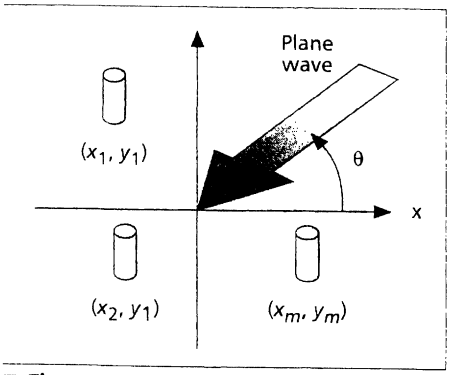


Figure 3. *Arbitrary antenna array configuration.*

The base station model of Fig. 5 was used to develop the theory and practice of base station diversity in today's cellular system and has led to rules of thumb for the spacing of diversity antennas on cellular towers [3].

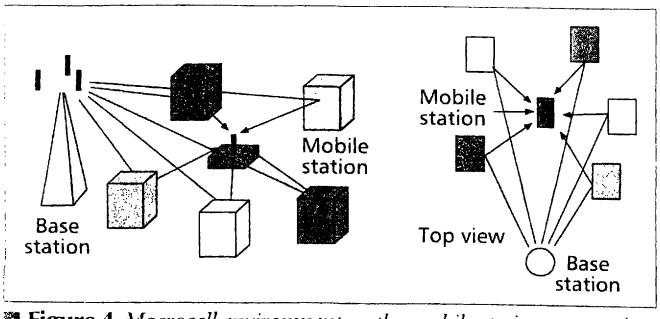


Figure 4. *Macrocell environment — the mobile station perspective.*

Microcell Environment – In the microcell environment, the base station antenna is usually mounted at the same height as the surrounding objects. This implies that the scattering spread of the AOA of the received signal at the base station is larger than in the macrocell case since the scattering process also happens in the vicinity of the base station. Thus, as the base station antenna is lowered, the tendency is for the multipath AOA spread to increase. This change in the behavior of the received signal is very important as far as antenna array applications are concerned. Studies have shown that statistical characteristics of the received signal are functions of the angle spread. Lee [3] and Adachi [6] found that the correlation between the signals received at two base station antennas increases as the angle spread decreases.

This section has presented some of the physical properties of a wireless communication channel. A mathematical expression that describes the time-varying spatial channel impulse response was given in Eq. 2. In the next section, several models that provide varying levels of information about the spatial channel are presented.

Space: The Final Frontier

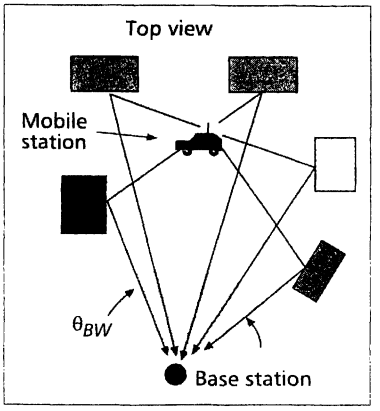
Details of the Spatial Channel Models

In the past when the distribution of angle of arrival of multipath signals was unknown, researchers assumed uniform distribution over $[0, 2\pi]$ [7]. In this section, a number of more realistic spatial channel models are introduced. The defining equations (or geometry) and the key results for the models

are described. Also provided is an extensive list of references.

Table 1 lists some representative active research groups in the field and their Web site addresses where more information on the subject can be found. (Note that this is by no means an exhaustive list.)

The Gaussian Wide Sense Stationary Uncorrelated Scattering (GWSSUS), Gaussian Angle of Arrival (GAA), Typical Urban (TU), and Bad Urban (BU) models described below were developed in a series of papers at the Royal Institute of Technology and may be downloaded from the Web site. Further details of the Geometrically Based Single Bounce (GBSB) models are given in theses at Virginia Tech, which are available at http://etd.vt.edu/etd/index.html.



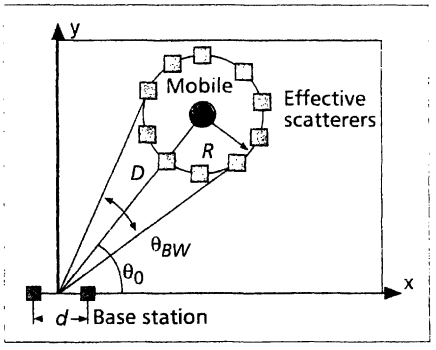
■ **Figure 5.** *Macrocell — base station perspective.*

These various models were developed and used for different applications. Some of the models were intended to provide information about only a single channel characteristic, such as angle spread, while others attempt to capture all the properties of the wireless channel. In the discussion of the models, an effort is made to identify the original motivation of the model and to convey the information the model is intended to provide.

Lee's Model

In Lee's model, scatterers are evenly spaced on a circular ring about the mobile as shown in Fig. 6. Each of the scatterers is intended to represent the effect of many scatterers within the region, and hence are referred to as *effective scatterers*. The model was originally used to predict the correlation between the signals received by two sensors as a function of element spacing. However, since the correlation matrix of the received signal vector of an antenna array can be determined by considering the correlation between each pair of elements, the model has application to any arbitrary array size.

The level of correlation will determine the performance of spatial diversity methods [3, 9]. In general, larger angle spreads and element spacings result in lower correlations, which provide an increased diversity gain. Measurements of the correlation observed at both the base station and the mobile are consistent with a narrow angle spread at the base station and a large angle spread at the mobile. Correlation measurements made at the base station indicate that the typical radius of scatterers is from 100 to 200 wavelengths [3].



■ **Figure 6.** *Lee's model.*

Assuming that N scatterers are uniformly placed on the circle with radius R and oriented such that a scatterer is located on the line of sight, the discrete AOAs are [9]

$$\theta_i \approx \frac{R}{D}\sin\left(\frac{2\pi}{N}i\right) \text{ for } i = 0,1,\ldots,N-1.$$

From the discrete AOAs, the correlation of the signals between any two elements of the array can be found using [9]

$$\rho(d,\theta_0,R,D) = \frac{1}{N}\sum_{i=0}^{N-1}\exp\left[-j2\pi d\cos(\theta_0+\theta_i)\right],$$

where d is the element spacing and θ_0 is measured with respect to the line between the two elements as shown in Fig. 6.

The original model provided information regarding only signal correlations. Motivated by the need to consider small-scale fading in diversity systems, Stapleton *et al.* proposed an extension to Lee's model that accounts for Doppler shift by imposing an angular velocity on the ring of scatterers [10, 11]. For the model to give the appropriate maximum Doppler shift, the angular velocity of the scatterers must equal v/R where v is the vehicle velocity and R is the radius of the scatterer ring [11]. Using this model to simulate a Rayleigh fading spatial channel model, the BER for a $\pi/4$ differential quadrature phase shift keyed (DQPSK) signal was simulated. The results were compared with measurements taken in a typical suburban environment. The resulting BER estimates were within a factor of two of the actual measured BER, indicating a reasonable degree of accuracy for the model [10].

When the model is used to provide

Research group	Web site
Center for Communications Research — University of Bristol	http://www.fen.bris.ac.uk/elec/research/ccr/ccr.html
Center for Personkommunikation — Aalborg University	http://www.kom.auc.dk/CPK/
Center for Wireless Telecommunications — Virginia Tech	http://www.cwt.vt.edu/
Mobile and Portable Radio Research Group — Virginia Tech	http://www.mprg.ee.vt.edu
Research Group for RF Communications — University of Kaiserlautern	http://www.e-technik.uni-kl.de/
Royal Institute of Technology	http://www.s3.kth.se
Smart Antenna Research Group — Stanford University	http://www-isl.stanford.edu/groups/SARG
Telecommunications and Information Systems Engineering — University of Texas at Austin	http://www.ece.utexas.edu/projects/tise/
Wireless Technology Group — McMaster University	http://www.crl.mcmaster.ca

■ **Table 1.** *Some active research groups in the field of adaptive antenna arrays.*

joint AOA and time of arrival (TOA) channel information, one finds that the resulting power delay profile is "U-shaped" [12]. By considering the intersections of the effective scatterers by ellipses of constant delay, one finds that there is a high concentration of scatterers in ellipses with minimum delay, a high concentration of scatterers in ellipses with maximum delay, and a lower concentration of scatterers between. Higher concentrations of scatterers with a given delay correspond with larger powers, and hence larger values on the power delay profile. The U-shaped" power delay profile is not consistent with measurements. Therefore, an extension to Lee's model is proposed in [11] in which additional scatterer rings are added to provide different power delay profiles.

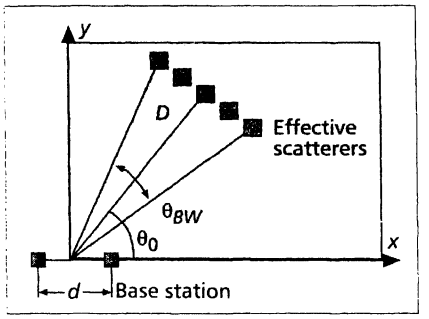


Figure 7. *Discrete uniform geometry.*

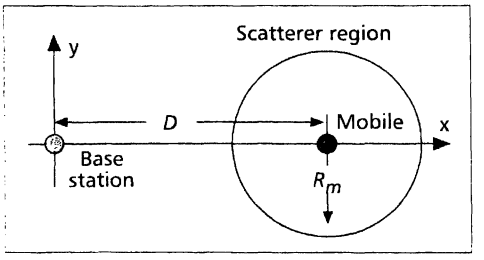


Figure 8. *Circular scatterer density geometry.*

While the model is quite useful in predicting the correlation between any two elements of the array, and hence the array correlation matrix, it is not well suited for simulations requiring a complete model of the wireless channel.

Discrete Uniform Distribution

A model similar to Lee's model in terms of both motivation and analysis was proposed in [9]. The model (referred to here as the discrete uniform distribution) evenly spaces N scatterers within a narrow beamwidth centered about the line of sight to the mobile as shown in Fig. 7. The discrete possible AOAs, assuming N is odd, are given by [9]

$$\theta_i = \frac{1}{N-1}\theta_{BW} i, \quad i = -\frac{N-1}{2}, \ldots, \frac{N-1}{2}.$$

From this, the correlation of the signals present at two antenna elements with a separation of d is found to be

$$\rho(d, \theta_0, \theta_{BW}) = \frac{1}{N} \sum_{i=-\frac{N-1}{2}}^{\frac{N-1}{2}} \exp\left[-j2\pi d \cos(\theta_0 + \theta_i)\right].$$

Measurements reported in [9] suggest that the AOA statistics in rural and suburban environments are Gaussian distributed (see the discussion of the GAA model later). However, in practice the AOA will be discrete (i.e., a finite number of samples from a Gaussian distribution), and therefore it is not valid to use a continuous AOA distribution to estimate the correlation present between different antenna elements in the array. The correlation that results from a continuous AOA distribution decreases monotomically with element spacing, whereas the correlation that results from a discrete AOA has damped oscillations present. Therefore, a continuous AOA distribution will underestimate the correlation that exists between the elements in the array [9].

In [9], a comparison is made between the correlation obtained using the discrete uniform distribution model, Lee's model, and a continuous Gaussian AOA as a function of element spacing. The comparisons indicate that, for small element separations (two wavelengths), the three models have nearly identical correlations. For larger element separations (greater than two wavelengths), the correlation values using the continuous Gaussian AOA are close to zero, while the two discrete models have oscillation peaks with correlations as high as 0.2 even beyond four wavelengths. Additionally, it was found that the correlation of the discrete uniform distribution falls off more quickly than the correlation in Lee's model.

Again, while the model is useful for predicting the correlation between any pair of elements in the array (which can be used to calculate the array correlation matrix), it fails to include all the phenomena, such as delay spread and Doppler spread, required for certain types of simulations.

Geometrically Based Single-Bounce Statistical Channel Models

Geometrically Based Single-Bounce (GBSB) Statistical Channel Models are defined by a spatial scatterer density function. These models are useful for both simulation and analysis purposes. Use of the models for simulation involves randomly placing scatterers in the scatterer region according to the form

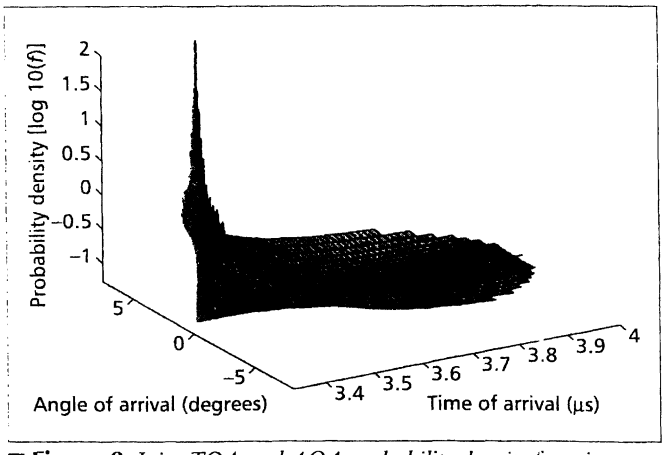


Figure 9. *Joint TOA and AOA probability density function at the base station, circular model (log-scale).*

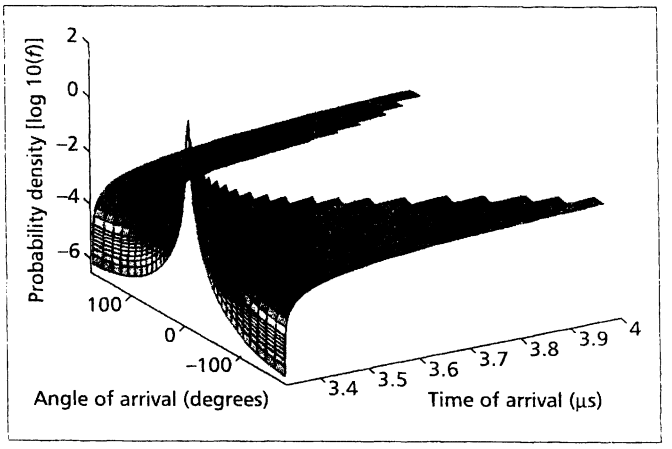


Figure 10. *Joint TOA and AOA probability density function at the mobile, circular model (log-scale).*

of the spatial scatterer density function. From the location of each of the scatterers, the AOA, TOA, and signal amplitude are determined.

From the spatial scatterer density function, it is possible to derive the joint and marginal TOA and AOA probability density functions. Knowledge of these statistics can be used to predict the performance of an adaptive array. Furthermore, knowledge of the underlying structure of the resulting array response vector may be exploited by beamforming and position location algorithms.

The shape and size of the spatial scatterer density function required to provide an accurate model of the channel is subject to debate. Validation of these models through extensive measurements remains an active area of research.

Geometrically Based Circular Model (Macrocell Model) – The geometry of the Geometrically Based Single Bounce Circular Model (GBSBCM) is shown in Fig. 8. It assumes that the scatterers lie within radius R_m about the mobile. Often the requirement that $R_m < D$ is imposed. The model is based on the assumption that in macrocell environments where antenna heights are relatively large, there will be no signal scatterering from locations near the base station. The idea of a circular region of scatterers centered about the mobile was originally proposed by Jakes [13] to derive theoretical results for the correlation observed between two antenna elements. Later, it was used to determine the effects of beamforming on the Doppler spectrum [14, 15] for narrowband signals. It was shown that the rate and the depth of the envelope fades are significantly reduced when a narrow-beam beamformer is used.

The joint TOA and AOA density function obtained from the model provides some insights into the properties of the model. Using a Jacobian transformation, it is easy to derive the joint TOA and AOA density function at both the base station and the mobile. The resulting joint probability density functions (PDFs) at the base station and the mobile are shown in the box on this page [16].

The joint TOA and AOA PDFs for the GBSBCM are shown in Figs. 9 and 10 for the case of D = 1 km and R_m = 100 m from the base station and mobile perspectives, respectively. The circular model predicts a relatively high probability of multipath components with small excess delays along the line of sight. From the base-station perspective, all of the multipath components are restricted to lie within a small range of angles.

The appropriate values for the radius of scatterers can be determined by equating the angle spread predicted by the model (which is a function of R_m) with measured values. Measurements reported in [9] suggest that typical angle spreads for macrocell environments with a T-R separation of 1 km are approximately two to six degrees. Also, it is stated that the angle spread is inversely proportional to the T-R separation, which leads to a radius of scatterers that ranges from 30 to 200 m [16]. In [3], it is stated that the active scattering region around the mobile is about 100–200 wavelengths for 900 MHz, which provides a range of 30–60 m, roughly the width of wide urban streets.

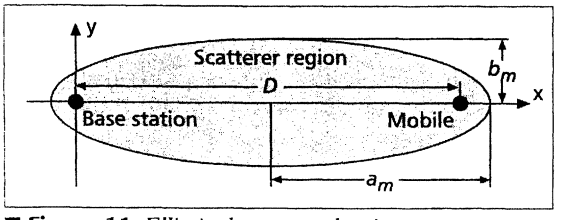


Figure 11. *Elliptical scatterer density geometry.*

The GBSBCM can be used to generate random channels for simulation purposes. Generation of samples from the GBSBCM is accomplished by uniformly placing scatterers in the circular scatterer region about the mobile and then calculating the corresponding AOA, TOA, and power levels.

Geometrically Based Elliptical Model (Microcell Wideband Model) – The Geometrically Based Single Bounce Elliptical Model (GBSBEM) assumes that scatterers are uniformly distributed within an ellipse, as shown in Fig. 11, where the base station and mobile are the foci of the ellipse. The model was proposed for microcell environments where antenna heights are relatively low, and therefore multipath scattering near the base station is just as likely as multipath scatterering near the mobile [17, 18].

A nice attribute of the elliptical model is the physical interpretation that only multipath signals which arrive with an absolute delay $\leq \tau_m$ are accounted for by the model. Ignoring components with larger delays is possible since signals with longer delays will experience greater path loss, and hence have relatively low power compared to those with shorter delays. Therefore, provided that τ_m is chosen sufficiently large, the model will account for nearly all the power and AOA of the multipath signals.

The parameters a_m and b_m are the semimajor axis and semiminor axis values, which are given by

$$a_m = \frac{c\tau_m}{2},$$

$$b_m = \frac{1}{2}\sqrt{c^2\tau_m^2 - D^2}.$$

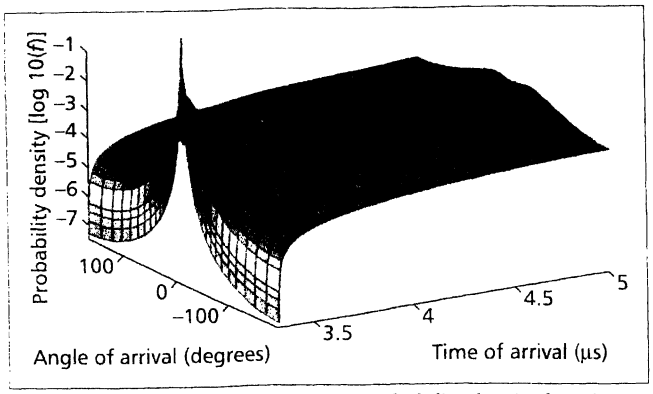
At the base station:

$$f_{\tau,\theta_b}(\tau,\theta_b) = \begin{cases} \dfrac{\left(D^2 - \tau^2c^2\right)\left(D^2c + \tau^2c^3 - 2\tau c^2 D\cos(\theta_b)\right)}{4\pi R_m^2\left(D\cos(\theta_b) - \tau c\right)^3} & : \quad \dfrac{D^2 - 2\tau cD\cos(\theta_b) + \tau^2c^2}{\tau c - D\cos(\theta_b)} \leq 2R_m \\ 0 & : \quad \textit{else.} \end{cases}$$

At the mobile:

$$f_{\tau,\theta_m}(\tau,\theta_m) = \begin{cases} \dfrac{\left(D^2 - \tau^2c^2\right)\left(D^2c + \tau^2c^3 - 2\tau c^2 D\cos(\theta_m)\right)}{4\pi R_m^2\left(D\cos(\theta_m) - \tau c\right)^3} & : \quad \dfrac{D^2 - \tau^2c^2}{D\cos(\theta_m) - \tau c} \leq 2R_m \\ 0 & : \quad \textit{else,} \end{cases}$$

where θ_b and θ_m are the angle of arrival measured relative to the line of sight from the base station and the mobile, respectively.



■ **Figure 12.** *Joint TOA and AOA probability density function, elliptical model (log-scale).*

where c is the speed of light and τ_m is the maximum TOA to be considered. To gain some insight into the properties of this model, consider the resulting joint TOA and AOA density function. Using a transformation of variables of the original uniform scatterer spatial density function, it can be shown that the joint TOA and AOA density function observed at the base station is given by [16]

$$f_{\tau,\theta_b}(\tau,\theta_b) = \begin{cases} \dfrac{\left(D^2-\tau^2c^2\right)\left(D^2c+\tau^2c^3-2\tau c^2 D\cos(\theta_b)\right)}{4\pi a_m b_m\left(D\cos(\theta_b)-\tau c\right)^3} & \dfrac{D}{c} \le \tau \le \tau_m \\ 0 & \text{elsewhere.} \end{cases}$$

where θ_b is AOA observed at the base station. A plot of the joint TOA and AOA PDF is shown in Fig. 12 for the case of D = 1 km and τ_m = 5 μs. From the plot of the joint TOA and AOA PDF, it is apparent that the GBSBEM results in a high probability of scatterers with minimum excess delay along the line of sight.

The choice of τ_m will determine both the delay spread and angle spread of the channel. Methods for selecting an appropriate value of τ_m are given in [18]. Table 2 summarizes the techniques for selecting τ_m where L_r is the reflection loss in dB, n is the path loss exponent, and τ_0 is the minimum path delay.

To generate multipath profiles using the GBSBEM, the most efficient method is to uniformly place scatterers in the ellipse and then calculate the corresponding AOA, TOA, and power levels from the coordinates of the scatterer. Uniformly placing scatterers in an ellipse may be accomplished by first uniformly placing the scatterers in a unit circle and then scaling each x and y coordinate by a_m and b_m, respectively [16].

Gaussian Wide Sense Stationary Uncorrelated Scattering

The GWSSUS is a statistical channel model that makes assumptions about the form of the received signal vector [19–22]. The primary motivation of the model is to provide a general equation for the received signal correlation matrix. In the GWSSUS model, scatterers are grouped into clusters in space. The clusters are such that the delay differences within each cluster are not resolvable within the transmission signal bandwidth. By including multiple clusters, frequency-selective fading channels can be modeled using the GWSSUS. Figure 13 shows the geometry assumed for the GWSSUS model corresponding to d = 3 clusters. The mean AOA for the kth cluster is denoted θ_{0k}. It is assumed that the location and delay associated with each cluster remains constant over several data bursts, b. The form of the received signal vector is

$$\mathbf{x}_b(t) = \sum_{k=1}^{d} \mathbf{v}_{k,b}\, s(t-\tau_k),$$

where $v_{k,b}$ is the superposition of the steering vectors during the bth data burst within the kth cluster, which may be expressed as

$$\mathbf{v}_{k,b} = \sum_{i=1}^{N_k} \alpha_{k,i} e^{j\Phi_{k,i}}\, \mathbf{a}\left(\theta_{0k}-\theta_{k,i}\right),$$

where N_k denotes the number of scatterers in the kth cluster, $\alpha_{k,i}$ is the amplitude, $\phi_{k,i}$ is the phase, $\theta_{k,i}$ is the angle of arrival of the ith reflected scatterer of the kth cluster, and $\mathbf{a}(\theta)$ is the array response vector in the direction of θ [9]. It is assumed that the steering vectors are independent for different k.

If N_k is sufficiently large (approximately 10 or more [19]) for each cluster of scatterers, the central limit theorem may be applied to the elements of $\mathbf{v}_{k,b}$. Under this condition, the elements of $\mathbf{v}_{k,b}$ are Gaussian distributed. Additionally, it is assumed that $\mathbf{v}_{k,b}$ is wide sense stationary. The time delays τ_k are assumed to be constant over several bursts, b, whereas the phases $\phi_{k,i}$ change much more rapidly. The vectors $\mathbf{v}_{k,b}$ are assumed to be zero mean, complex Gaussian wide sense, stationary random processes where b plays the role of the time argument. The vector $\mathbf{v}_{k,b}$ is a multivariate Gaussian distribution, which is described by its mean and covariance matrix. When no line of sight component is present, the mean will be zero due to the random phase $\phi_{k,i}$, which is assumed to be uniformly distributed in the range 0 to 2π. When a direct path component is present, the mean becomes a scaled version of the corresponding array response vector $E\{\mathbf{v}_{k,b}\} \propto \mathbf{a}(\theta_{0k})$ [9]. The covariance matrix for the kth cluster is given by [21]

$$R_k = E\left\{\mathbf{v}_{k,b}\mathbf{v}_{k,b}^H\right\} = \sum_{i=1}^{N_k} \left|\alpha_{k,i}\right|^2 E\left\{\mathbf{a}\left(\theta_{0k}-\theta_{k,i}\right)\mathbf{a}^H\left(\theta_{0k}-\theta_{k,i}\right)\right\}.$$

The model provides a fairly general result for the form of the covariance matrix. However, it does not indicate the number or location of the scattering clusters, and hence requires some additional information for application to typical environments.

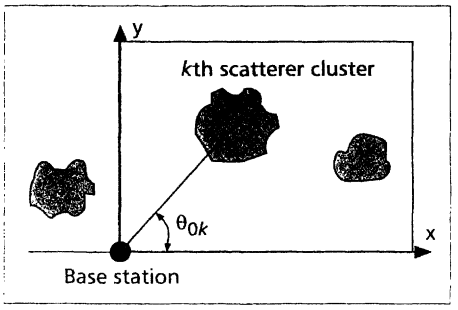
Criteria	Expression
Fixed maximum delay, τ_m	τ_m = constant
Fixed threshold T in dB	$\tau_m = \tau_0 10^{(T-L_r)/10n}$
Fixed delay spread, σ_τ	$\tau_m = 3.244\sigma_\tau + \tau_0$
Fixed max. excess delay, τ_e	$\tau_m = \tau_0 + \tau_e$

■ **Table 2.** *Methods for selection* τ_m.

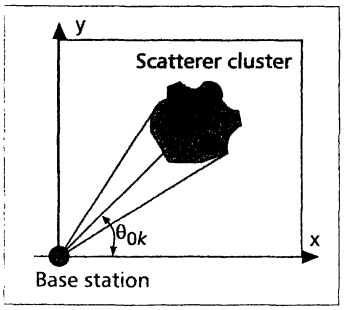
Gaussian Angle of Arrival

The Gaussian Angle of Arrival (GAA) channel model is a special case of the GWSSUS model described above where only a single cluster is considered (d = 1), and the AOA statistics are assumed to be Gaussian distributed about some nominal angle, θ_0, as shown in Fig. 14. Since only a single cluster is considered, the model is a narrowband channel model that is valid when the time spread of the channel is small compared to the inverse of the signal bandwidth; hence, time shifts may be modeled as simple phase shifts [23].

The statistics of the steering vector are distributed as a multivariate Gaussian random variable. Similar to the GWSSUS model, if no line of



■ **Figure 13.** *GWSSUS geometry.*



■ **Figure 14.** *GAA geometry.*

sight is present, then $E\{\mathbf{v}_{k,h}\} = 0$; otherwise, the mean is proportional to the array response vector $\mathbf{a}(\theta_{0k})$. For the special case of uniform linear arrays, the covariance matrix may be described by

$$\mathbf{R}(\theta_0,\sigma_\theta) \approx p\mathbf{a}(\theta_0)\mathbf{a}^H(\theta_0) \otimes \mathbf{B}(\theta_0,\sigma_\theta),$$

where the (k,l) element of $\mathbf{B}(\theta_0, \sigma_\theta)$ is given by

$$B(\theta_0,\sigma_\theta)_{kl} = \exp\left[-2(\pi\Delta(k-l))^2\sigma_\theta^2\cos^2\theta\right],$$

p is the receiver signal power, Δ is the element spacing, and $\otimes$ denotes element-wise multiplication [23].

Time-Varying Vector Channel Model (Raleigh's Model)

Raleigh's time-varying vector channel model was developed to provide both small scale Rayleigh fading and theoretical spatial correlation properties [24]. The propagation environment considered is densely populated with large dominant reflectors (Fig. 15). It is assumed that at a particular time the channel is characterized by L dominant reflectors. The received signal vector is then modeled as

$$\mathbf{x}(t) = \sum_{l=0}^{L(t)-1} \mathbf{a}(\theta_l)\alpha_l(t)s(t-\tau) + \mathbf{n}(t).$$

where $\mathbf{a}$ is the array response vector, $\alpha_l(t)$ is the complex path amplitude, $s(t)$ is the modulated signal, and $\mathbf{n}(t)$ is additive noise. This is equivalent to the impulse response given in Eq. 2.

The unique feature of the model is in the calculation of the complex amplitude term, $a_l(t)$, which is expressed as

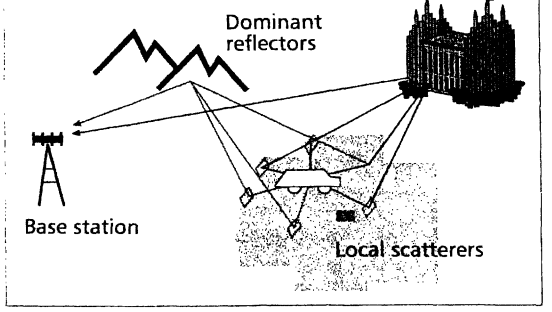
$$\alpha_l(t) = \beta_l(t)\cdot\sqrt{\Gamma_l}\cdot\psi(\tau_l),$$

where Γ_l accounts for log-normal fading, $\psi(\tau_l)$ describes the power delay profile, and $\beta_l(t)$ is the complex intensity of the radiation pattern as a function of time. The complex intensity is described by

$$\beta_l(t) = K\sum_{n=1}^{N_l} C_n(\theta_l)\exp\left(j\omega_d\cos(\Omega_{n,l})t\right),$$

where N_l is the number of signal components contributing to the lth dominant reflecting surface, K accounts for the antenna gains and transmit signal power, $C_n(\theta_l)$ is the complex radiation of the nth component of the lth dominant reflecting surface in the direction of θ_l, ω_d is the maximum Doppler shift, and $\Omega_{n,l}$ is the angle toward the nth component of the lth dominant reflector with respect to the motion of the mobile. The resulting complex intensity, $\beta_l(t)$, exhibits a complex Gaussian distribution in all directions away from the mobile [24].

Both the time and spatial correlation properties of the model are compared to theoretical results in [24]. The comparison shows that there is good agreement between the two.



■ **Figure 15.** *Raleigh's model signal environment.*

Two Simulation Models (TU and BU)

Next we describe two spatial channel models that have been developed for simulation purposes.

The Typical Urban (TU) model is designed to have time properties similar to the GSM-TU defined in GSM 05.05, while the Bad Urban (BU) model was developed to model environments with large reflectors that are not in the vicinity of the mobile. Although the models are designed for GSM, DCS1800, and PCS1900 formats, extensions to other formats are possible [251].

Both of these models obtain the received signal vector using

$$\mathbf{x}(t) = \sum_{n=1}^{N} \alpha_n(t)\exp\left(-j2\pi f_c\frac{l_n(t)}{c} + \beta\right)s\left(t - \frac{l_n(t)}{c} + \Delta_l\right)\mathbf{a}(\theta_n(t))$$

where N is the number of scatterers, f_c the carrier frequency, c is the speed of light, $l_n(t)$ the path propagation distance, β a random phase, and Δ_l random delay. In general, the path propagation distance $l_n(t)$ will vary continuously with time; hence, Doppler fading occurs naturally in the model.

Typical Urban (TU) – In the TU model, 120 scatterers are randomly placed within a 1 km radius about the mobile [25]. The position of the scatterers is held fixed over the duration in which the mobile travels a distance of 5 m. At the end of the 5 m, the scatterers are returned to their original position with respect to the mobile. At each 5-m interval, random phases are assigned to the scatterers as well as randomized shadowing effects, which are modeled as log-normal with distance with a standard deviation of 5–10 dB [25]. The received signal is determined by brute force from the location of each of the scatterers. An exponential path loss law is also applied to account for large-scale fading [21]. Simulations have shown that the TU model and the GSM-TU model have nearly identical power delay profiles, Doppler spectrums, and delay spreads [25]. Furthermore, the AOA statistics are approximately Gaussian and similar to those of the GAA model described above.

Bad Urban (BU) – The BU is identical to the TU model with the addition of a second scatterer cluster with another 120 scatterers offset 45° from the first, as shown in Fig. 16. The scatterers in the second cluster are assigned 5 dB less average power than the original cluster [25]. The presence of the second cluster results in an increased angle spread, which in turn reduces the off-diagonal elements of the array covariance matrix. The presence of the second cluster also causes an increase in the delay spread.

Uniform Sectored Distribution

The defining geometry of Uniform Sectored Distribution (USD) is shown in Fig. 17 [26]. The model assumes that scatterers are uniformly distributed within an angle distribution of θ_{BW} and a radial

range of ΔR centered about the mobile. The magnitude and phase associated with each scatterer is selected at random from a uniform distribution of [0,1] and [0, 2π], respectively. As the number of scatterers approaches infinity, the signal fading envelope becomes Rayleigh with uniform phase [26]. In [26], the model is used to study the effect of angle spread on spatial diversity techniques. A key result is that beam-steering techniques are most suitable for scatterer distributions with widths slightly larger than the beamwidths.

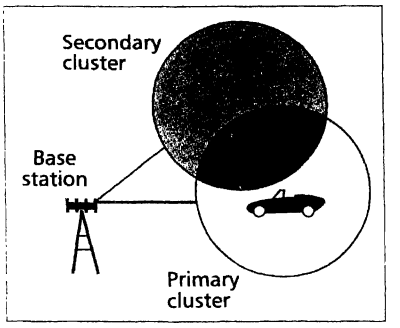


Figure 16. *Bad Urban vector channel model geometry.*

Modified Saleh-Valenzuela's Model

Saleh and Valenzuela developed a multipath channel model for indoor environment based on the clustering phenomenon observed in experimental data [27]. The clustering phenomenon refers to the observation that multipath components arrive at the antenna in groups. It was found that both the clusters and the rays within a cluster decayed in amplitude with time. The impulse response of this model is given by

$$h(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\alpha_{ij}\delta\left(t - T_i - \tau_{ij}\right) \qquad (3)$$

where the sum over i corresponds to the clusters and the sum over j represents the rays within a cluster. The variables α_{ij} are Rayleigh distributed with the mean square value described by a double-exponential decay given by

$$\overline{\alpha_{ij}^2} = \overline{\alpha_{00}^2}\exp\left(-T_i/\Gamma\right)\exp\left(\tau_{ij}/\gamma\right)$$

where Γ and γ are the cluster and ray time decay constant, respectively. Motivated by the need to include AOA in the channel mode, Spencer *et al.* proposed an extension to the Saleh-Valenzuela's model [28], assuming that time and the angle are statistically independent, or

$$h(t,\theta) = h(t)h(\theta).$$

Similar to the time impulse response in Eq. 3, the proposed *angular impulse response* is given by

$$h(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\alpha_{ij}\delta\left(\theta - \Theta_i - \omega_{ij}\right)$$

where α_{ij} is the amplitude of the jth ray in the ith cluster. The variable Θ_i is the mean angle of the ith cluster and is assumed to be uniformly distributed over [0, 2π]. The variable ω_{ij} corresponds to the ray angle within a cluster and is modeled as a Laplacian distributed random variable with zero mean and standard deviation σ:

$$f(\omega) = \frac{1}{\sqrt{2}\sigma}\exp\left(-\left|\frac{\sqrt{2}\omega}{\sigma}\right|\right).$$

This model was proposed based on indoor measurements which will be discussed in the fourth section.

Extended Tap-Delay-Line Model

A wideband channel model that is an extension of the traditional statistical tap-delay-line model and includes AOA information was developed by Klein and Mohr [29]. The channel impulse response is represented by

$$h(\tau,t,\theta) = \sum_{w=1}^{W} a_w(t)\delta(\tau - \tau_w)\delta(\theta - \theta_w).$$

This model is composed by W taps, each with an associated time delay τ_w, complex amplitude α_w, and AOA θ_w. The joint density functions of the model parameters should be determined from measurements. As shown in [29], measurements can provide histograms of the joint distribution of $|a|$, τ, and θ, and the density functions, which are proportional to these histograms, can be chosen.

Elliptical Subregions Model (Lu, Lo, and Litva's Model)

Lu *et al.* [30] proposed a model of multipath propagation based on the distribution of the scatterers in elliptical subregions, as shown in Fig. 18. Each subregion (shown in a different shade) corresponds to one range of the excess delay time. This approach is similar to the GBSBEM proposed by Liberti and Rappaport [18] in that an ellipse of scatterers is considered. The primary difference between the two models is in the selection of the number of scatterers and the distribution of those scatterers. In the GBSBEM, the scatterers were uniformly distributed within the entire ellipse. In Lu, Lo, and Litva's model, the ellipse is first subdivided into a number of elliptical subregions. The number of scatterers within each subregion is then selected from a Poisson random variable, the mean of which is chosen to match the measured time delay profile data.

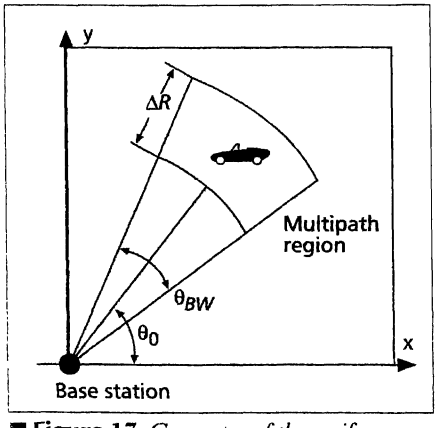


Figure 17. *Geometry of the uniform sectored distribution.*

It was also assumed that the multipath components arrive in clusters due to the multiple reflecting points of the scatterers. Thus, assuming that there are L scatterers with K_l reflecting points each, the model proposed is represented by

$$h(t,t_0) = \sum_{i=1}^{L} E_t\left(\theta_i^{(t)}\right) \times \sum_{k=0}^{K_i}\alpha_{ik}\exp\left(-\left(2\pi f_{ik}t_0 + \gamma_{ik}\right)\right)\delta(t - \tau_{ik})E_r(\theta_{ik})$$

where α_{ik}, τ_{ik}, and γ_{ik} correspond to the amplitude, time delay, and phase of the signal component from the ikth reflecting point, respectively. f_{ik} is the Doppler frequency shift of each individual path, θ_{ik} is the angle between the ikth path and the receiver-to-transmitter direction, and $\theta_i^{(t)}$ is the angle of the ith scatterer as seen from the transmitter. $E_t(\theta)$ and $E_r(\theta)$ are the radiation patterns of the transmit and receive antennas, respectively. The variable θ_{ik} was assumed to be Gaussian distributed.

Simulation results using this model were presented in [30], showing that a 60° beamwidth antenna reduces the mean RMS delay spread by about 30–43 percent. These results are consistent with similar measurements made in Toronto using a sectorized antenna [31].

Measurement-Based Channel Model

A channel model in which the parameters are based on measurement was proposed by Blanz et al. [32]. The idea behind this approach is to characterize the propagation environment, in terms of scattering points, based on measurement data. The time-variant impulse response takes the form

$$h(\tau,t) = \int_0^{2\pi} \vartheta(\tau,t,\theta) * g(\tau,\theta) * f(\tau) d\theta$$

where $f(\tau)$ is the impulse response representing the joint transfer characteristic of the transmission system components (modulator, demodulator, filters, etc.), and $g(\tau, \theta)$ is the characteristic of the base station antenna. The term $\vartheta(\tau, t, \theta)$ is the time-variant directional distribution of channel impulse response seen from the base station. This distribution is time-variant due to mobile motion and depends on the location, orientation, and velocity of the mobile station antenna and the topographical and morphographical properties of the propagation area as well. Measurement is used to determine the distribution $\vartheta(\tau, t, \theta)$.

Ray Tracing Models

The models presented so far are based on statistical analysis and measurements, and provide us with the average path loss and delay spread, adjusting some parameters according to the environment (indoor, outdoor, obstructed, etc.). In the past few years, a *deterministic* model, called *ray tracing*, has been proposed based on the geometric theory and reflection, diffraction, and scattering models. By using site-specific information, such as building databases or architecture drawings, this technique deterministically models the propagation channel [33–36], including the path loss exponent and the delay spread. However, the high computational burden and lack of detailed terrain and building databases make ray tracing models difficult to use. Although some progress has been made in overcoming the computational burden, the development of an effective and efficient procedure for generating terrain and building data for ray tracing is still necessary.

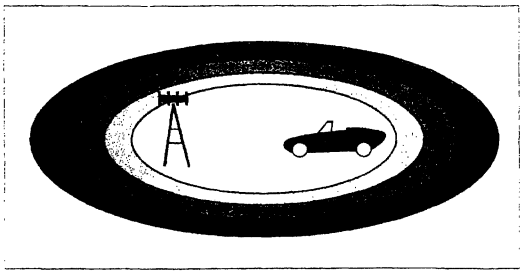


Figure 18. *Elliptical subregions spatial scatterer density.*

Channel Model Summary

Table 3 summarizes each of the spatial models presented above.

Spatial Signal Measurements

There have been only a few publications relating to spatial channel measurements. In this section, references are given to these papers, and the key results are described.

In [38], TOA and AOA measurements are presented for outdoor macrocellular environments. The measurements were made using a rotating 9° azimuth beam directional receiver antenna with a 10 MHz bandwidth centered at 1840 MHz. Three environments near Munich were considered, including rural, suburban, and urban areas with base station antenna heights of 12.3 m, 25.8 m, and 37.5 m, respectively. The key observations made include [38]:

- Most of the signal energy is concentrated in a small interval of delay and within a small AOA in rural, suburban, and even many urban environments.
- By using directional antennas, it is possible to reduce the time dispersion.

Another set of TOA and AOA measurements is reported in [39] for urban areas. The measurements were made using a two-element receiver that was mounted on the test vehicle with an elevation of 2.6 m. The transmitting antenna was placed 30 m high on the side of a building. A bandwidth of 10 MHz with a carrier frequency of 2.33 GHz was used. The delay-Doppler spectra observed at the mobile was used to obtain the delay-AOA spectra. The second antenna element is used to remove the ambiguity in AOA that would occur if only the Doppler spectra were known. The results indicate that it is possible to account for most of the major features of the delay-AOA spectra by considering the large buildings in the environment.

Motivated by diversity combining methods, earlier measurements were concerned primarily with determining the correlation between the signals at two antenna elements as a function of the element separation distance. These studies found that, at the mobile, relatively small separation distances were required to obtain a small degree of correlation between the elements, whereas at the base station very large spacing was needed. These findings indicate that there is a relatively small angle spread observed at the base station [6].

Previously, an extension to Saleh-Valenzuela's indoor model, including AOA information, was presented. This extension was proposed based on indoor measurements of delay spread and AOA at 7 GHz made at Brigham Young University [40]. The AOAs were measured using a 60 cm parabolic dish antenna that had a 3 dB beamwidth of 6°. The results showed a clustering pattern in both time and angle domain, which led to the proposed channel model described in [28]. Also, it was observed that the cluster mean angle of arrival was uniformly distributed [0, 2π]. The distribution of the angle of arrival of the rays within a cluster presented a sharp peak at the mean, leading to the Laplacian distribution modeling. The standard deviation found for this distribution was around 25°. Based on these measurements, a channel model including delay spread and AOA information was proposed, supposing that time and angle were independent variables.

In [41], two-dimensional AOA and delay spread measurement and estimation were presented. The measurements were made in downtown Paris using a channel sounder at 900 MHz and a horizontal rectangular planar array at the receiver. The estimation of AOA, including azimuth and elevation angle, was performed using 2D unitary ESPRIT [42] with a time resolution of 0.1μs and angle resolution of 5°. The results presented confirmed assumptions made in urban propagation, such as the wave-guiding mechanism of streets and the exponential decay of the power delay profile. Also, it was observed that 90 percent of the received power was contained in the paths with elevation between 0° and 40° with the low elevated paths contributing a larger amount.

Finally, in [43] measurements are used to show the variation in the spatial signature with both time and frequency. Two measures of change are given, the relative angle change given by

Model	Description	References
Lee's Model	Effective scatterers are evenly spaced on a circular ring about the mobile Predicts correlation coefficient using a discrete AOA model Extension accounts for Doppler shift	[3, 9, 11]
Discrete Uniform Distribution	N scatterers are evenly spaced over an AOA range Predicts correlation coefficient using a discrete AOA model Correlation predicted by this model falls off more quickly than the correlation in Lee's model	[9]
Geometrically Based Circular Model (Macrocell Model)	Assumes that the scatterers lie within circular ring about the mobile AOA, TOA, joint TOA and AOA, Doppler shift, and signal amplitude information is provided Intended for macrocell environments where antenna heights are relatively large	[12–14, 16, 37]
Geometrically Based Elliptical Model (Microcell Wideband Model)	Scatterers are uniformly distributed in an ellipse where the base station and the mobile are the foci of the ellipse AOA, TOA, joint TOA and AOA, Doppler shift, and signal amplitude information is provided Intended for microcell environments where antenna heights are relatively low	[17, 18]
Gaussian Wide Sense Stationary Uncorrelated Scattering (GWSSUS)	N scatterers are grouped into clusters in space such that the delay differences within each cluster are not resolvable within the transmission signal BW Provides an analytical model for the array covariance matrix	[19–22]
Gaussian Angle of Arrival (GAA)	Special case of the GWSSUS model with a single cluster and angle of arrival statistics assumed to be Gaussian distributed about some nominal angle Narrowband channel model Provides an analytical model for the array covariance matrix	[23]
Time-Varying Vector Channel Model (Raleigh's Model)	Assumes that the signal energy leaving the region of the mobile is Rayleigh faded Angle spread is accounted for by dominant reflectors Provides both Rayleigh fading and theoretical spatial correlation properties	[20]
Typical Urban	Simulation model for GSM, DCS1800, and PCS1900 Time domain properties are similar to the GSM-TU defined in GSM 05.05 120 scatterers are randomly placed within a 1 km radius about the mobile Received signal is determined by brute force from the location of each of the scatterers and the time-varying location of the mobile AOA statistics are approximately Gaussian	[21, 22, 25]
Bad Urban	Simulation model for GSM, DCS1800, and PCS1900 Accounts for large reflectors not in the vicinity of the mobile Identical to the TU model with the addition of a second scatterer cluster offset 45° from the first	[21, 22, 25]
Uniform Sectored Distribution	Assumes that scatterers are uniformly distributed within an angle distribution of θ_{BW} and a radial range of ΔR centered about the mobile Magnitude and phase associated with each scatterer are selected at random from a uniform distribution of [0,1] and [0, 2π], respectively	[26]
Modified Saleh-Valenzuela's Model	An extension to the Saleh-Valenzuela model, including AOA information in the channel model Assumes that time and the angle are statistically independent Based on indoor measurements	[28]
Extended Tap-Delay-Line Model	Wideband channel model Extension of the traditional statistical tap-delay-line model which includes AOA information The joint density functions of the model parameters should be determined from measurements	[29]
Spatio-Temporal Model (Lu, Lo, and Litva's Model)	Model of multipath propagation based on the distribution of the scatterers in elliptical subregions, corresponding to a range of excess delay time Similar to the GBSBEM	[30]
Measurement-Based Channel Model	Parameters are based on measurement Characterizes the propagation environment in terms of scattering points	[32]
Ray Tracing Models	Deterministic model based on the geometric theory and reflection, diffraction, and scattering models Uses site-specific information, such as building databases or architecture drawings	[33–36]

Table 3. *Summary of spatial channel models.*

$$\text{Relative Angle Change (\%)} = 100 \times \sqrt{1 - \left| \frac{\mathbf{a}_i^*}{\|\mathbf{a}_i\|} \cdot \frac{\mathbf{a}_j^*}{\|\mathbf{a}_j\|} \right|^2}$$

and the relative amplitude change, found using

$$\text{Relative Amplitude Change(dB)} = 20\log_{10} \frac{\|\mathbf{a}_j\|}{\|\mathbf{a}_i\|}$$

where $\mathbf{a}_i$ and $\mathbf{a}_j$ are the two spatial signatures (array response vectors) being compared. The measurements indicate that when the mobile and surroundings are stationary, there are relatively small changes in the spatial signature. Likewise, there are moderate changes when objects and the environment are in motion and large changes when the mobile itself is moving. Also, it was found that the spatial signature changes significantly with a change in carrier frequency. In particular, the measurements found that the relative amplitude change in the spatial signal could exceed 10 dB with a frequency change of only 10 MHz. This result indicates that the uplink spatial signature cannot be directly applied for downlink beamforming in most of today's cellular and PCS systems that have 45 MHz and 80 MHz separation between the uplink and downlink frequencies, respectively.

Application of Spatial Channel Models

The effect that classical channel properties such as delay spread and Doppler spread have on system performance has been an active area of research for several years and hence is fairly well understood. The spatial channel models include the AOA properties of the channel, which are often characterized by the angle spread. The angle spread has a major impact on the correlation observed between the pairs of elements in the array. These correlation values specify the received signal vector covariance matrix, which is known to determine the performance of linear combining arrays [44]. In general, the higher the angle spread the lower the correlation observed between any pair of elements in the array. The various spatial channel models provide different angle spreads and hence will predict different levels of system performance.

The channel models presented here have various applications in the analysis and design of systems that utilize adaptive antenna arrays. Some of the models were developed to provide analytical models of the spatial correlation function, while others are intended primarily for simulation purposes.

Simulation

More and more companies are relying on detailed simulations to help design and develop today's wireless networks. The application of adaptive antennas is no exception. However, to obtain reliable results, accurate spatial channel models are needed. With accurate simulations of adaptive antenna array systems, researchers will be able to predict the capacity improvement, range extension, and other performance measures of the system, which in turn will determine the cost effectiveness of adaptive array technologies.

Algorithm Development

The availability of channel models also opens up the possibility of developing new maximum likelihood smart antennas and AOA estimation algorithms based on these channel models. Good analytical models that will provide insights into the structure of the spatial channel are needed.

Conclusions

As antenna technology advances, radio system engineers are increasingly able to utilize the spatial domain to enhance system performance by rejecting interfering signals and boosting desired signal levels. However, to make effective use of the spatial domain, design engineers need to understand and appropriately model spatial domain characteristics, particularly the distribution of scatterers, angles of arrival, and the Doppler spectrum. These characteristics tend to be dependent on the height of the transmitting and receiving antennas relative to the local environment. For example, the distributions expected in a microcellular environment with relatively low base station antenna heights are usually quite different from those found in traditional macrocellular systems with elevated base station antennas.

This article has provided a review of a number of spatial propagation models. These models can be divided into three groups:

- General statistically based models
- More site-specific models based on measurement data
- Entirely site-specific models.

The first group of models (Lee's Model, Discrete Uniform Distribution Model, Geometrically Based Single Bounce Statistical Model, Gaussian Wide Sense Stationary Uncorrelated Scattering Model, Gaussian Angle of Arrival Model, Uniform Sectored Distributed Model, Modified Saleh-Valenzuela's Model, Spatio-Temporal Model) are useful for general system performance analysis. The models in the second group (Extended Tap Delay Line Model and Measurement-Based Channel Model) can be expected to yield greater accuracy but require measurement data as an input. An example from the third group of models is Ray Tracing, which has the potential to be extremely accurate but requires a comprehensive description of the physical propagation environment as well as measurements to validate the models.

Further research is required to validate and enhance the models described in this article. Bearing in mind that an objective of modeling is to substantially reduce the amount of physical measurement required in the system planning process, it is important for design engineers to have reliable models of AOA, TDOA, delay spread, and the power of the multipath components. Further measurement programs that focus on spatial domain signal characteristics are required. These programs would greatly benefit from the development of improved measurement equipment.

Armed with improved spatial channel modeling tools and a greater understanding of signal propagation, engineers can begin to meet the challenges inherent in designing future high-capacity/high-quality wireless communication systems, including the effective use of smart antennas.

References

[1] R. H. Clarke, "A Statistical Theory of Mobile-Radio Reception," *Bell Sys. Tech. J.*, vol. 47, 1968, pp. 957–1000.

[2] M. J. Gans, "A Power Spectral Theory of Propagation in the Mobile Radio Environment," *IEEE Trans. Vehic. Tech.*, vol.VT-21, Feb. 1972, pp. 27–38.

[3] W. C. Y. Lee, *Mobile Communications Engineering*, New York: McGraw Hill, 1982.

[4] Turin, "A Statistical Model for Urban Multipath Propagation," *IEEE Trans. Vehic. Tech.*, vol. VT-21, no. 1, Feb. 1972, pp. 1–11.

[5] T. S. Rappaport, *Wireless Communications: Principles & Practice*, Upper SaddleRiver, NJ: Prentice Hall PTR, 1996.

[6] F. Adachi *et al.*, "Crosscorrelation Between the Envelopes of 900 MHz Signals Received at a Mobile Radio Base Station Site," *IEE Proc.*, vol. 133, Pt. F, no. 6, Oct. 1986, pp. 506–12.

[7] J. D. Parsons, *The Mobile Radio Propagation Channel*, New York: John Wiley & Sons, 1992.

[8] W. C. Y. Lee, *Mobile Cellular Telecommunications Systems*, New York: McGraw Hill, 1989.

[9] D. Aszetly, "On Antenna Arrays in Mobile Communication Systems: Fast Fading and GSM Base Station Receiver Algorithms," Ph.D. dissertation, Royal Inst. Technology, Mar. 1996.
[10] S. P. Stapleton, X. Carbo, and T. McKeen, "Spatial Channel Simulator for Phased Arrays," *Proc. IEEE VTC.*, 1994, pp 1789–92.
[11] S. P. Stapleton, X. Carbo, and T. Mckeen, "Tracking and Diversity for a Mobile Communications Base Station Array Antenna," *Proc. IEEE VTC*, 1996, pp. 1695–99.
[12] R. B. Ertel, "Vector Channel Model Evaluation," Tech.rep., SW Bell Tech. Resources, Aug. 1997.
[13] J. William C. Jakes, ed., *Microwave Mobile Communications*, New York: John Wiley & Sons, 1974.
[14] P. Petrus, "Novel Adaptive Array Algorithms and Their Impact on Cellular System Capacity," Ph.D. dissertation, Virginia Polytechnic Inst. and State Univ., Mar. 1997.
[15] P. Petrus, J. H. Reed, and T. S. Rappaport, "Effects of Directional Antennas at the Base Station on the Doppler Spectrum," *IEEE Commun. Lett.*, vol. 1, no. 2, Mar. 1997.
[16] R. B. Ertel, "Statistical Analysis of the Geometrically Based Single Bounce Channel Models," unpublished notes, May 1997.
[17] J. C. Liberti, "Analysis of CDMA Cellular Radio Systems Employing Adaptive Antennas," Ph.D. dissertation, Virginia Polytechnic Inst. and State Univ , Sept. 1995.
[18] J. C. Liberti and T. S. Rappaport, "A Geometrically Based Model for Line of Sight Multipath Radio Channels," *IEEE VTC*, Apr. 1996, pp. 844–48.
[19] P. Zetterberg and B. Ottersten, "The Spectrum Efficiency of a Basestation Antenna Array System for Spatially Selective Transmission," *IEEE VTC*, 1994.
[20] P. Zetterberg, "Mobile Communication with Base Station Antenna Arrays: Propagation Modeling and System Capacity," Tech. rep., Royal Inst. Technology, Jan 1995.
[21] P. Zetterberg and P L. Espensen, "A Downlink Beam Steering Technique for GSM/DCS1800/PCS1900," IEEE PIMRC, Taipei, Taiwan, Oct., 1996.
[22] P. Zetterberg, P. L. Espensen, and P. Mogensen, "Propagation, Beamsteering and Uplink Combining Algorithms for Cellular Systems," ACTS Mobile Commun. Summit, Granada, Spain, Nov., 1996.
[23] B. Ottersten "Spatial Division Multiple Access (SDMA) in Wireless Communications, ' *Proc. Nordic Radio Symp.*, 1995.
[24] G. G. Raleigh and A. Paulraj, "Time Varying Vector Channel Estimation for Adaptive Spatial Equalization," *Proc. IEEE Globecom*, 1995, pp. 218–24.
[25] P. Mogensen *et al.*, "Algorithms and Antenna Array Recommendations," Tech.rep. A020/AUC/A12/DR/P/1/xx-D2.1.2, Tsunami (II), Sept. 1996.
[26] O. Norklit and J. B. Anderson, "Mobile Radio Environments and Adaptive Arrays," *Proc. IEEE PIMRC*, 1994, pp. 725–28.
[27] A. M. Saleh and R. A. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE JSAC*, vol. SAC-5, Feb. 1987.
[28] Q. Spencer *et al.*, "A Statistical Model for Angle of Arrival in Indoor Multipath Propagation," *IEEE VTC*, 1997, pp 1415–19.
[29] A. Klein and W. Mohr, "A Statistical Wideband Mobile Radio Channel Model Including the Direction of Arrival," *Proc. IEEE 4th Int'l. Symp. Spread Spectrum Techniques & Applications*, 1996, pp. 102–06.
[30] M. Lu, T. Lo, and J. Litva, "A Physical Spatio-Temporal Model of Multipath Propagation Channels," *Proc. IEEE VTC*, 1997, pp. 180–814.
[31] E. S. Sousa *et al.*, "Delay Spread Measurements for the Digital Cellular Channel in Toronto," *IEEE Trans. Vehic. Tech.*, vol. 43, no. 4, Nov. 1994, pp. 837–47.
[32] J. J. Blanz, A. Klein, and W. Mohr, "Measurement-Based Parameter Adaptation of Wideband Spatial Mobile Radio Channel Models," *Proc. IEEE 4th Int'l. Symp. Spread Spectrum Techniques & Applications*, 1996, pp. 91–97.
[33] K. R. Schaubach, N. J. Davis IV, and T. S. Rappaport, "A Ray Tracing method for Predicting Path Loss and Delay Spread in Microcellular Environment," *IEEE VTC*, May 1992, pp. 932–35.
[34] J. Rossi and A. Levi, "A Ray Model for Decimetric Radiowave Propagation in an Urban Area," *Radio Science*, vol. 27, no. 6, 1993, pp. 971–79.
[35] R. A. Valenzuela, "A Ray Tracing Approach for Predicting Indoor Wireless Transmission," *IEEE VTC*, 1993, pp. 214–18.
[36] S. Y. Seidel and T. S. Rappaport, "Site-Specific Propagation Prediction for Wireless In-Building Personal Communication System Design," *IEEE Trans. Vehic. Tech.*, vol. 43, no. 4, Nov. 1994.
[37] P. Petrus, J. H. Reed, and T. S. Rappaport, "Geometrically Based Statistical Channel Model for Macrocellular Mobile Environments," *IEEE Proc. GLOBECOM, 1996*, pp. 1197–1201.
[38] A. Klein *et al.*, "Direction-of-Arrival of Partial Waves in Wideband Mobile Radio Channels for Intelligent Antenna Concepts," *IEEE VTC*, 1996, pp. 849–53.
[39] H. J. Thomas, T. Ohgane, and M. Mizuno, "A Novel Dual Antenna Measurement of the Angular Distribution of Received Waves in the Mobile Radio Environment as a Function of Position and Delay Time," *Proc. IEEE VTC*, vol. 1, 1992, pp 546–49.
[40] Q. Spencer *et al.*, "Indoor Wideband Time/Angle of Arrival Multipath Propagation Results," *IEEE VTC*, 1997, pp. 1410–14.
[41] J. Fuhl, J-P Rossi, and E. Bonek, "High Resolution 3-D Direction-of-Arrival Determination for Urban Mobile Radio," *IEEE Trans. Antennas and Propagation*, vol. 45, no. 4, Apr. 1997, pp. 672–81.
[42] M. D. Zoltowski, M. Haardt, and C. P. Mathews, "Closed-Form 2-D Angle Estimation with Rectangular Arrays in Element Space or Beamspace via Unitary ESPRIT," *IEEE Trans. Signal Processing*, vol., vol. 44, Feb. 1996, pp. 316–28.
[43] S. S. Jenget *et al.*, "Measurements of Spatial Signature of an Antenna Array," *Proc. IEEE 6th PIMRC*, vol. 2, Sept. 1995, pp. 669–72.
[44] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*, New York: John Wiley & Sons, 1980.

Additional Reading

[1] M. J. Devasirvatham, "A Comparison of Time Delay Spread and Signal Level Measurements with Two Dissimilar Office Buildings," *IEEE Trans. Antennas and Propagation*, vol. AP-35, No. 3, Mar. 1994, pp 319–24.
[2] M. J. Feuerstein *et al.*, "Path Loss, Delay Spread, and Outage Models as Functions of Antenna for Microcellular Systems Design," *IEEE Trans. Vehic. Tech.*, vol. 43, no. 3, Aug. 1994, pp. 487–98.
[3] J. C. Liberti and T. S. Rappaport, "Analytical Results for Capacity Improvements in CDMA," *IEEE Trans. Vehic. Tech.*, VT-43, no. 3, Aug. 1994.
[4] D. Molkdar, "Review on Radio Propagation into and within Buildings," *IEE Proc.*, vol. 138, no. 1, Feb. 1991.
[5] A. F. Naguib, A. Paulraj, and T. Kailath, "Capacity Improvement with Base-Station Antenna Array," *IEEE Trans. Vehic. Tech.*, VT-43, no. 3, Aug. 1994, pp. 691–98.
[6] A. F. Naguib and A. Paulraj, "Performance Enhancement and Trade-offs of a Smart Antenna in CDMA Cellular Network," *IEEE VTC*, 1995, pp. 225–29.
[7] S. Y. Seidel *et al.*, "Path Loss, Scattering and Multipath Propagation Statistics for European Cities for Digital Cellular and Microcellular Radiotelephone," *IEEE Trans. Vehic, Tech,*, vol. VT-40, no. 4, Nov. 1991, pp. 721–30.
[8] S. C. Swales, M. A. Beach, and J. P. MacGeehan, "The Performance Enhancement of Multi-beam Adaptive Base-station Antenna for Cellular Land Mobile Radio System," *IEEE Trans. Vehic. Tech.*, vol.VT-39, Feb 1990, pp. 56–67.

Antenna Systems for Base Station Diversity in Urban Small and Micro Cells

Patrick C. F. Eggers, Jørn Toftgård, and Alex M. Oprea

Abstract—**This paper describes cross-correlation properties for compact urban base station antenna configurations, nearly all resulting in very low envelope cross-correlation coefficients of about 0.1 to 0.3. A focus is set on polarization diversity systems for their potential in improving link quality when hand-held terminals are involved. An expression is given for the correlation function of compound space and polarization diversity systems. Dispersion and envelope dynamic statistics are presented for the measured environments. For microcell applications, it is found that systems such as GSM having a bandwidth of 200 kHz or less can use narrowband cross-correlation analysis directly.**

I. Introduction

As the bandwidths of newer digital land/mobile and personal communications systems increase, radio channel time dispersion can produce noticeable frequency-selective fading within the band. The capacity demand on such systems is also increasing, leading to network layouts with smaller cell sizes. A way to assist with cell layout, and to combat interference and fading degradation, is to use base station diversity. A normal figure of merit for narrowband antenna diversity systems is the envelope cross-correlation coefficient $\rho_{12_{env}}$. For phase-modulated systems, the phase decorrelation is also relevant especially for predetection combining, and here the complex cross-correlation coefficient ρ_{12} can be used.

The work presented in this paper falls in two areas. The paper describes investigations of the overall diversity potential for antenna systems used in urban small and microcells (predominantly range lengths up to 3000 m and 300 m, respectively), and the degree of overall radio channel dispersion by considering the temporal and spatial domain separately.

Two base station sites have been investigated. One microcell site has antenna heights of a few meters above rooftops and range lengths from 50 to 800 m. On the other site, the antennas were elevated approximately 30 m above rooftops and ranges from 200 to 3000 m were used. Antenna configurations with vertical and horizontal separations have been used, as well as compound space and polarization diversity configurations. Hand-held terminals in existing mobile systems and upcoming personal systems will most likely have low polarization and will be subject to varying handset orientation. Under such conditions, polarization diversity is an obvious way of improving link quality.

Manuscript received March 1992; revised December 1992. This work was supported by the Danish Technical Research Board. This work was presented in part at the Fifth Digital Mobile Radio Conference, Helsinki, Finland, December 1–3, 1992.

The authors are with the Mobile Communications Group, Aalborg University, DK-9220 Aalborg, Denmark.

IEEE Log Number 9210927.

Digital transmission using antenna diversity in frequency-selective fading channels has been investigated in [1]. The investigations were based on simulations and assuming fully decorrelated branch signals with uncorrelated scattering [(US), i.e., full decorrelation along the temporal axis], with the same average power delay profile (PDP) at each branch. This paper presents an extensive measurement campaign, performed in an urban environment. The radio channel has been sampled simultaneously in the spatial and temporal domains on two antenna branches. Achievable figures of the degree of antenna signal correlation, dispersion, and fading for practical systems are presented, giving a base of comparison to some of the assumptions used in [1].

II. Mobile Radio Channel Characterization

The mobile radio channel is often modeled by a superposition of the following three major effects:

a) Large variations, i.e., range dependency (median pathloss)

b) Local variations, i.e., shadowing, (generally considered log-normal distributed)

c) Short-term variations, i.e., multipath fading (shown to follow Rayleigh or Rician distributions)

In general, the mobile radio channel is both time variant and nonstationary. The measurements reported here have mostly been performed in low-density and low-velocity traffic situations. The roof of the measuring van (2.2 m) was higher than most of the other moving vehicles (scatterers). Thus, a time-invariant channel situation is assumed during the measurement, although channel condition details may have changed for a later measurement at the same location (note it would be very difficult to find exactly the same location in practice). This assumption allows a comparison of the statistics. The measurement runs used have been short, so that global variation statistics (pathloss) are assumed "frozen." Thus, we consider the multipath fading being superimposed with local variations (shadowing) affecting only the mean values of the envelope signal. A removal of the estimated shadowing component is possible. This will stabilize the mean to a quasiconstant value. Thus, a wide sense stationary (WSS) short-term fading channel is extracted for analysis (described in more detail later). Since the WSS mobile radio channel may be visualized as a linear filter at a given time, it can be described either by its impulse response or by its transfer function. Alternatively, the spatially variant behavior of the radio channel can be investigated. This has the advantage of providing a stable channel, invariant to change in mobile terminal velocity.

Reprinted from *IEEE Journal on Selected Areas in Communications,* Vol. 11, No. 7, pp. 1046–1057, September 1993.

The spatially variant impulse response $h(\tau, x)$ provides a direct illustration of the multipath phenomenon, where τ is the temporal lag (delay) and x is the spatial position along the travelled path. The frequency Doppler transfer function $H(f, f_d)$ gives an appropriate picture of the spatial-varying behavior via Doppler shifts. These two functions, which describe adequately a deterministic spatially variant channel, form a two-dimensional Fourier pair.

The following subsections describes the statistical parameters used for evaluation of the radio channel considered WSS in the spatial domain and fully stationary in the temporal domain.

A. Spatial Domain

Often, a CW measurement has been used to determine the correlation between fading signals from two diversity antennas. This corresponds to the DC response along the temporal axis, i.e., the complex envelope of the radio channel is found by integration as

$$g(x) = \int_{-\infty}^{\infty} h(\tau, x)\, d\tau \tag{1}$$

corresponding to the DC component of the spatial-variant frequency transfer function $H(0, x)$. For wide bandwidth systems, a Received Signal Strength Indicator (RSSI) signal may also be used for driving a diversity combiner. The RSSI signal represents the total power over the band and will, in frequency-selective channels, exhibit reduced fading dynamics. The RSSI signal is real and is found by integration of the spatially variant frequency transfer function weighted with the receiver filter, i.e.,

$$|g(x)|^2 = \int_{-\infty}^{\infty} |H(f, x)|^2 |H_{\text{filter}}(f)|^2\, df\,. \tag{2}$$

The normalized spatial complex correlation coefficient function is

$$\rho_{g1g2}(\Delta x) = \frac{C_{g1g2}(\Delta x)}{\sqrt{C_{g1g1}(0) \cdot C_{g2g2}(0)}} \tag{3}$$

where the cross covariance of lag Δx is

$$C_{g1g2}(\Delta x) = R_{g1g2}(\Delta x) - \mu_{g1}^{*} \cdot \mu_{g2}$$

and

$$R_{g1g2}(\Delta x) = E[g_1^{*}(x) \cdot g_2(x + \Delta x)]\,,$$
$$\mu_g = E[g(x)]\,. \tag{4}$$

R_{g1g2} and μ_g are the complex cross-correlation functions and complex mean of the CW or RSSI signals following (1) or (2). $E[\cdot]$ is the expectation value (mean value) operator. For a continuous random variable (RV) V with probability function (pdf) $f(v)$, $E[V] = \int_{-\infty}^{\infty} v f(v)\, dv$. Similarly, a discrete RV of sampled data $V = [v_1 \cdots v_N]$ is treated with ensemble averaging $E[V] = \langle V \rangle = (1/N) \sum_{n=1}^{N} v_n$. The parameter of most interest when considering antenna systems is the correlation function with zero lag, i.e., $\Delta x = 0$. The value of (3) at zero lag is normally just referred to as the correlation coefficient, i.e., $\rho_{g1g2}(0) = \rho_{12}$. The widely used envelope correlation coefficient $\rho_{12_{\text{env}}}$ is a specialization of (3) when the real envelope functions $|g(\tau, x)|$ are used instead of the complex functions $g(\tau, x)$. Thus, $\rho_{12_{\text{env}}}$ becomes a real function. For analysis purposes, the power correlation coefficient $\rho_{12_{\text{pow}}}$ has also been used (see Appendix A).

B. Temporal Domain

For the temporal domain, the power delay profile (PDP) is used as the basis for overall dispersion analysis

$$\text{PDP}(\tau) = \int_{-\infty}^{\infty} |h(\tau, x)|^2\, dx\,. \tag{5}$$

The PDP is an average power density impulse response. The frequency correlation function (FCF) follows by inverse Fourier transform via the Wiener-Khintchine Theorem as

$$\text{FCF}(\Delta f) = \mathcal{F}^{-1}[\text{PDP}(\tau)]\,. \tag{6}$$

The coherence bandwidth is found at a given correlation level from the normalized FCF and describes the overall dispersion strength represented in the frequency domain. The RMS delay spread (S) is widely used to describe the overall degree of dispersion in the time domain [6], and is given as the square root of the second central moment of the PDP

$$S^2 = E\left[(\tau - \overline{\tau})^2\right] = \int_{-\infty}^{\infty} (\tau - \overline{\tau})^2\, \frac{\text{PDP}(\tau)}{P_t}\, d\tau \tag{7}$$

where P_t is the total power in the PDP. For high temporal resolution systems, the absolute delay function must be taken into account in (5) through (7). The temporal resolution in our experiments is approximately 1 μs corresponding to a 300 m spatial resolution. As each measurement run was limited to 51 m, we consider the absolute delay from base to mobile being constant over the integration of PDP in (5). We can then use (7) without modifications.

III. The Measurement Setup

Usually, the choice of channel sounding technique depends upon the application foreseen for the collected data. To permit a study of the instantaneous behavior of the mobile radio channel later on, a direct pulse technique was chosen. The channel sounder, having a time resolution capability of approximately 1 μs and an instantaneous dynamic range of about 26 dB, was operated at 970 MHz. To achieve adequate resolution of Doppler shifts, a sampling interval of 6 cm was used. The following subsections describe the hardware implementation of the measurement system.

A. Transmitter

All frequencies used in the transmitter are derived from a stable 10 MHz frequency standard. The maximum transmitted power was a 5 W peak, the signal being radiated from a quarter-wavelength monopole mounted on the metal roof of the measuring vehicle. The cross-polarization discrimination (XPD) of the transmitter was difficult to measure, but field tests indicated a figure of up to 20 dB. This is quite high compared to XPD figures for normal vehicle mounted whip

antennas, but in our case the measuring van (Ford Transit-L) has a large flat roof, providing a better ground plane than most personal cars.

B. Receiver

A dual-branch receiver is used to investigate the complex cross-correlation between the two antenna signals. Each branch is a conventional quadrature detection receiver and, as in the transmitter, all frequencies used are synthesized from a stable 10 MHz frequency standard. The dominating output filters limit the bandwidth (−3 dB single sided) to 0.6 MHz. A linear input signal dynamic range of −45 to −105 dBm is available. The I and Q signals from the two branches of the receiver are simultaneously sampled at a 3.3 MHz rate and stored on a hard disk.

Two identical planar array directional antennas with an azimuth and elevation 3 dB beam width of 60° were used. Thus, the same illumination of the environment is achieved for both the copolarized and cross-polarized cases. The antennas have a gain of 8 dBi and a XPD better than 30 dB, ensuring sufficient suppression of the unwanted polarization to provide usable figures for the XPD of the radio environment.

C. Measurement Data

The length of each test run was 51 meters. The data was stored on a PC, consisting of three consecutive complex impulse responses for both receiver branches at every spatial sampling point. The data transfer rate of the PC limited the vehicle speed to 3 m/s.

IV. Measurement Scenarios

Two series of field trials were conducted in Aalborg, a medium-sized European city. The principal variable during these trials was the receiver antenna configuration at the base station. Fig. 1 shows a map of the test area with the base station and measurement site location.

In the first experiment, a typical urban environment was covered, corresponding to a microcell. The base station (BS1) was located on the roof of one of Aalborg University's buildings, some 18 m above ground level. The antenna position was just above the rooftop level of the surrounding buildings. The test route was chosen in a relatively flat and heavily built-up urban area consisting mostly of five-story buildings which are in close proximity to each other. A square of 0.75 by 0.75 km was covered by 14 measurement sites (shown by numbered circles on the map). Data were recorded on each of the 14 sites for 12 antenna configurations at the base station (Table I).

In the second experiment, a small cell was considered. The base station (BS2) antennas were located on Hotel Hvide Hus, a 15 story building in Aalborg, some 30 m higher than the rooftop level of the surrounding buildings. The first part of the route was similar to that of the first experiment. The second part was chosen in an undulating environment, starting some 1.5 km from the base station, across a 500 m wide river which splits the town in two. On the Nørre Sundby side of the river, the ground slopes upward reaching a height of 45 m at 2.6 km (site 14) from the base station and the area changes gradually into a suburban one. There were 18 measurement sites chosen (shown by numbered triangles on the map), half in Aalborg and half in Nørre Sundby, the furthest point being around 3 km away from the base station. Data were collected for each of the 18 measurement sites for six antenna configurations (Table II). The higher elevation and longer range lengths at BS2 will lower the effective beam width of the incoming scattered signal. Thus, larger antenna separations have been used at BS2 to obtain comparable decorrelation to the configurations used at BS1.

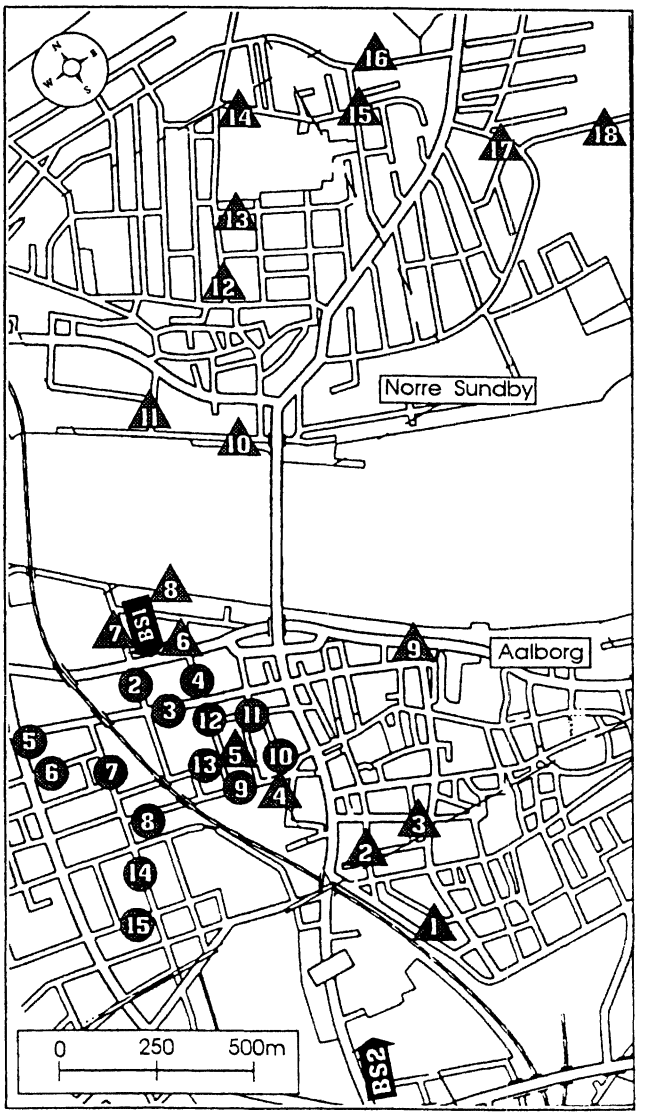


Fig. 1. Map showing test areas used with BS1 and BS2. *Circles:* BS1 measurement locations. *Triangles:* BS2 measurement locations. The numbers refer to each specific location.

V. Data Processing

The measured diversity data have been processed collectively for a block of measurements with the same base station antenna configuration. A measurement series consists of a measurement file for each location. The measured data are stored in the measurement files with a recorded impulse response for each diversity branch. Three consecutive in-

TABLE I
ANTENNA CONFIGURATIONS USED AT BS1. THE NUMBER IN THE BRACKETS REFER TO THE ELEVATION OF THE LOWEST ANTENNA WITH RESPECT TO THE BASE STATION ROOF.

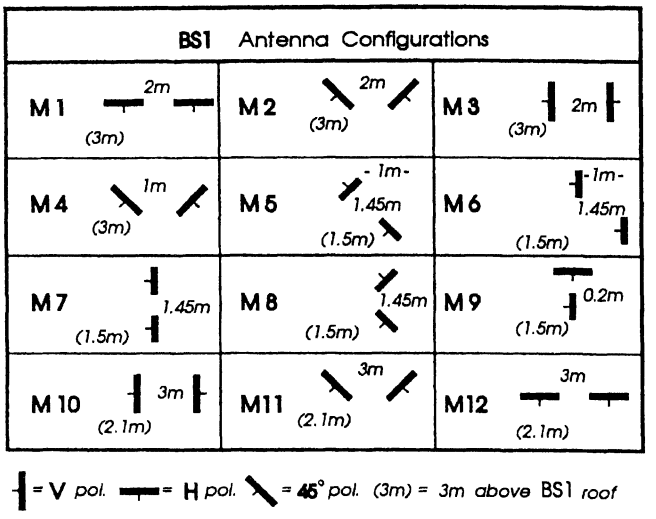
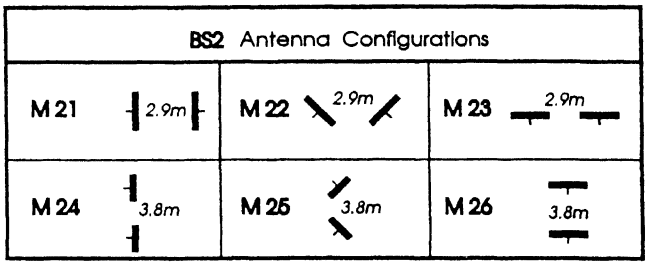


TABLE II
ANTENNA CONFIGURATIONS USED AT BS2. SYMBOLS USED, FOLLOW SAME NOMENCLATURE AS USED IN TABLE I.



stantaneous impulse responses were recorded for each spatial position. Through coherent (complex) averaging of these three impulse responses, a noise suppression of almost 5 dB was obtained.

To retrieve the fast fading components of the recorded signal, the log-normal shadowing component was compensated for. The log-normal component of the data was extracted by performing a moving average filtering in the spatial domain. Two consecutive rectangular windows with lengths of 10 and 6.7 m were used.

The PDP's for each branch were calculated by averaging the power of instantaneous impulse responses over the full 51 m run length. The statistical parameters extracted from the PDP's are delay spread (S), mean delay ($\overline{\tau}$), and total power (P_t). The FCF's were found by inverse Fourier transform of the PDP's. The coherence bandwidths have been found for 0.9, 0.7, and 0.5 correlation levels of the normalized FCF.

Three fading envelope signals have been generated from each data set—one CW signal and two RSSI signals. The RSSI signals have been extracted for the full bandwidth of the measuring equipment (1.2 MHz–3 dB double sided) and for a GSM-like bandwidth (200 kHz–3 dB double sided). A square cosine filter $|H(f)_{\text{filter}}|^2 = \text{COS}^2(0.5\pi(f - f_o)/200)$ kHz has been used for the GSM-like signal (truncated to 0 outside $(f - f_o) = \pm 200$ kHz). This filter provides a power spectrum closely resembling the main lobe of a GSM/GMSK power spectrum.

For the CW signals, several correlations between the antenna signals have been calculated: ρ_{12}, $\rho_{12_{\text{env}}}$, $\rho_{12_{\text{pow}}}$, and ρ_{II}, ρ_{IQ} (see Appendix A). For the RSSI signals, only envelope correlations have been calculated. For all three envelope signal types and each antenna, the 1% level of the cumulative envelope distribution (cdf) is found. This information is valuable in assessing necessary fading margins in radio network design. Also, assumptions of Rayleigh fading signals (see Appendix A) can be coarsely validated in a compact manner. For a true Rayleigh fading signal, the 1% cdf level corresponds to a signal level 20 dB below the mean of the signal.

VI. RESULTS OF THE ANALYSIS

In the following discussions, all data from all antenna configurations are considered in the figures (unless specifically noted otherwise).

A. Radio Environment

From [2] and [3], it is found that short-term fading statistics are independent of impulse response shape and only depend on RMS delay spread (S) [as long as the delay spread bandwidth product is much smaller than unity, i.e., $S \cdot BW << 1$]. This is referred to as a low-dispersive channel situation. For low-dispersive situations, the envelope correlation coefficient will be fully relevant, as all large phase and timing jitter variations are associated with fades in the instantaneous total power (RSSI signal). Thus, a classical envelope-controlled diversity combining scheme can also suppress the radio channel-introduced phase and timing jitter.

Previous investigations [4] have shown log-normal shadowing with $\sigma = 8$ dB from the area covered by BS1. Short-term effect investigations in [5] indicate Rayleigh-distributed fading behavior for most of the strong power components of the PDP, for the areas covered by BS2. Line-of-sight (LOS) peaks, however, show distinctly Rician distributed envelope fading. Fig. 2 shows the 1% level of the cumulative envelope distribution dependence versus normalized delay spread, i.e., $S \cdot BW$. For the plotting of the CW case, a 25 kHz bandwidth has been used. As expected, the CW signals are insensitive to radio dispersion whereas the RSSI signals reduce dynamics with higher dispersion and larger bandwidth. From Table IV, it follows that the mean 1% cumulative CW envelope levels are indeed close to −20 dB corresponding to Rayleigh fading. For the RSSI signals, it follows that there is a strong increase in the 1% envelope level for higher bandwidths, and that this increase is approximately 2 dB stronger for BS2 than for BS1. This is due to a higher degree of radio dispersion (S) at BS2 compared to BS1. The higher dispersion, the more frequency-selective the channel becomes and the probability of total power vanishing within the band is decreased.

The transmitter/receiver equipment filters limit the resolution of the measured channel impulse response. The RMS delay spreads found from the PDP's were compensated for

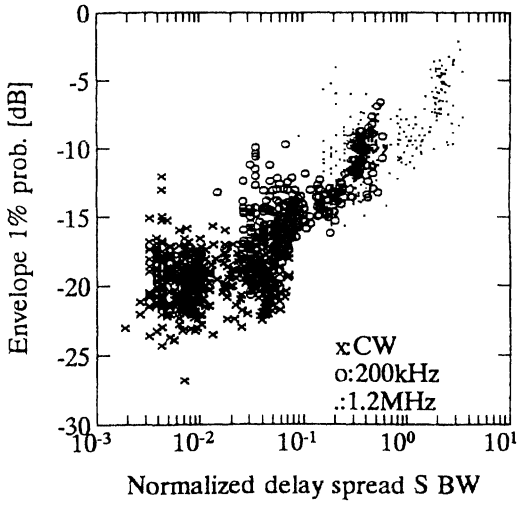


Fig. 2. Cumulative envelope distribution at 1% level versus normalized delay spread.

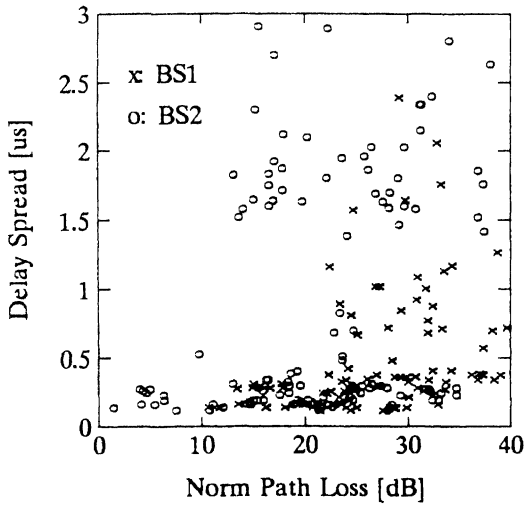
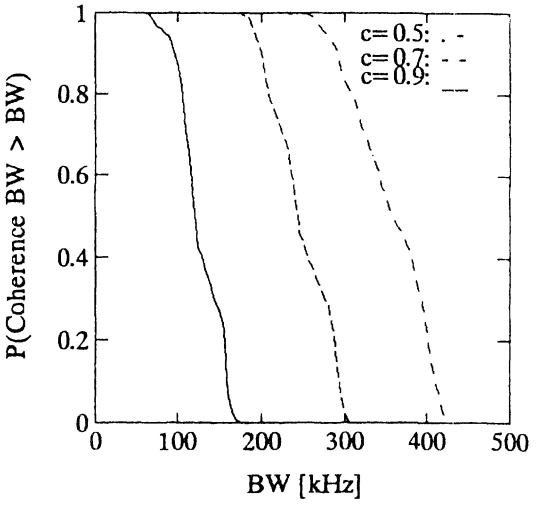
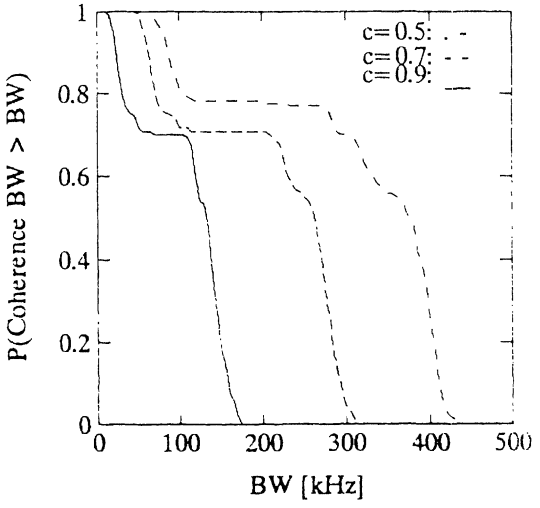


Fig. 3. RMS delay spread versus normalized (excess) path loss.



(a)



(b)

Fig. 4. Coherence bandwidth at 0.5, 0.7, and 0.9 correlation level for (a) BS1 and (b) BS2.

this time interference by simple variance subtraction [6], i.e.,

$$S^2_{\text{channel}} = S^2_{\text{measured}} - S^2_{\text{equipment}}. \quad (8)$$

It is, thus, possible to obtain RMS delay spread figures with finer resolution than the time resolution of the transmitter/receiver equipment initially would permit. This makes the RMS delay spread figure also relevant to systems having higher bandwidths than the 1.2 MHz used by the sounding system. Fig. 3 indicates a dependence between excess path loss and RMS delay spread, i.e., larger delay spreads for increasing excess path loss (where the excess path loss is the path loss relative to the free-space path loss). The indicated dependence can be explained by a lower probability of LOS paths with an increased excess path loss, as LOS paths will generally have lower excess attenuation than scattered and reflected paths. Thus, LOS paths tend to dominate (increase) the total power and reduce the effective width (dispersion) of the PDP.

Fig. 4 shows the cumulative coherence bandwidth for all measurement locations and antenna systems for BS1 and BS2. For BS1 [Fig. 4(a)], it is observed that the 0.7 coherence bandwidth is larger than 200 kHz for 90% of the locations. In most cases, this justifies using narrowband considerations for a GSM-like system in this environment. The coherence bandwidth function for BS2 [Fig. 4(b)] shows a large discontinuity around the 70–80% level. This is due to the rather harsh environment picked for this cell type. A river splits the town in two and allows free propagation for strong echoes across the river. Many PDP's exhibit a double-spike shape with about 4 μs separation (corresponding to twice the river width). This shows up as a strong periodicity in the FCF and, thus gives rise to the plateau in the cumulative coherence bandwidth function. For bandwidths over 500 kHz, a direct narrowband description alone cannot give a satisfactory description of the propagation conditions for the micro or the small cell environment investigated.

The GSM system has incorporated a slow frequency hopping option over 5 MHz. From Fig. 4, it follows that if the hopping frequency is just over 400 kHz neither the micro nor small cell environment will have locations with a FCF having

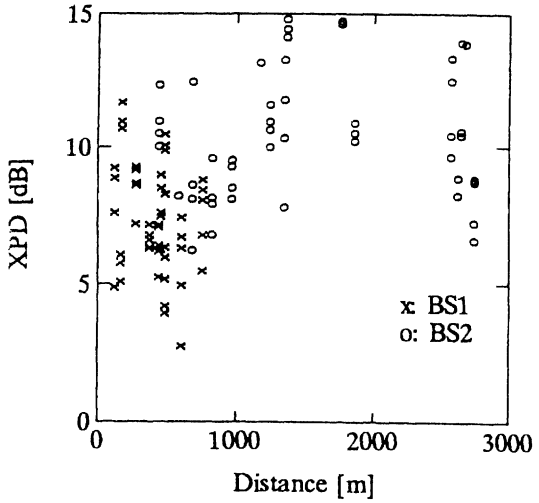


Fig. 5. Cross polarization discrimination versus range length.

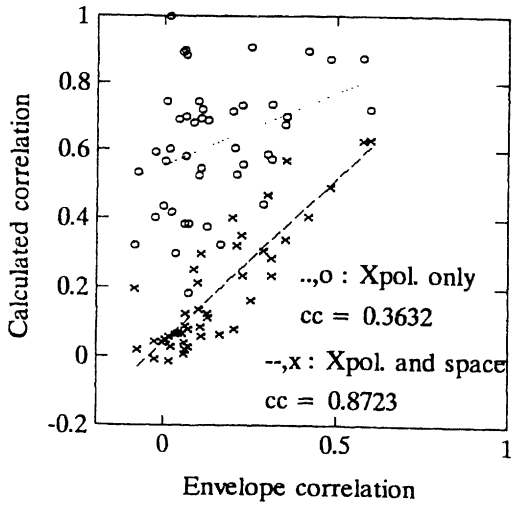


Fig. 6. Equation (9) with $\rho_{\text{env}X\text{pol}}$ only ($\rho_{\text{env}V} = 1$) and with measured $\rho_{\text{env}V}$ versus envelope correlation coefficient for physically slant antenna systems.

levels greater than 0.5. This indicates effective frequency diversity gain by use of the slow frequency hopping option.

B. Cross Polarization

The possibility of using polarization decorrelation ($\rho_{\text{pow}X\text{pol}}$) at a base station has been reported by [9], where the XPD was found in the range around 6 dB for urban areas at 920 MHz. Compact, compound, horizontal, and vertical space diversity antenna systems and polarization decorrelation have also been reported in [10], where macrocells (10–20 km cell radius) were investigated. For urban areas (partly the same area as used for the experiments in this paper), the XPD was around 4 dB in the 900 MHz band. For suburban areas, XPD was found in the range around 12 dB. Fig. 5 shows the XPD versus the range for all antenna configurations at BS1 and BS2. It follows from Fig. 5 that there is a large spread of XPD for both BS1 and BS2. From Table IV, it follows that the urban (BS1) mean XPD is higher (7.4 dB) for small and microcells compared to the macrocells previously investigated. This is expected, as the depolarization is related to the number of reflections, diffractions, etc., that each path is subject to. For small cells, fewer possible pertubating objects will be in the signal path. For longer range lengths, the mean XPD is 11.4 dB (BS2), which can be explained by the environment having a more suburban character with few obstructions between base and mobile. The higher XPD's found for small and microcells means a reduced diversity gain by using polarization decorrelation compared to macrocell case.

Appendix A gives a simple extension of the work of Vaughan [12]. The result of the analysis shows that polarization and space decorrelation effects are independent and multiplicative for compound antenna systems, as given in (9). The assumptions are noncorrelated Rayleigh fading polarizations, equal XPD (Γ) at the antennas, and the quadrature components of the vertical and horizontal polarizations exhibit approximately equal space decorrelation properties. In practice, this means that the radiation patterns of the antennas must be identical and rotationally symmetric. Experimental test cases are given in Appendix A to support these assumptions. The power correlations are related as:

$$\begin{aligned}\rho_{\text{pow}12} &= \rho_{\text{pow}V}\rho_{\text{pow}X\text{pol}}(\Gamma,\alpha) \\ &= \rho_{\text{pow}V}\frac{\Gamma+\frac{1}{\Gamma}-2}{\Gamma+\frac{1}{\Gamma}+\frac{1}{\tan^2(\alpha)}+\tan^2(\alpha)}\end{aligned} \tag{9}$$

where α is the antenna rotation angle (see Appendix A). This expression will also hold closely for envelope correlations, due to the strong relation as shown in Fig. 10. For co-located antennas, $\rho_{\text{env}V} = 1$, and only the cross-polarization term ($\rho_{\text{env}X\text{pol}}$) will provide decorrelation. In this situation, the expression is equal to the original one given in [12]. Thus, a compound space and polarization diversity system will achieve the envelope decorrelation of the product of space and polarization decorrelation effects. Fig. 6 shows the calculated correlation coefficient using (9) with $\alpha = 45°$, $\rho_{\text{env}V}$, and XPD figures from M1, M3, M10, M12, M21, M23, M24, and M26 against the compound antenna system correlation $\rho_{\text{env}12}$ found from M2, M11, M22, and M25 (Tables I and II). It is seen that (9) provides a very good fit to measurements (correlation $cc = 0.87$), whereas disregarding the space decorrelation ($\rho_{\text{env}V} = 1$) in (9) leads to a significantly decreased fit ($cc = 0.36$).

For predetection combiners, the polarization decorrelation potential is invariant with α but, for postdetection combiners, equal antenna branch powers (equal SNR in all branches) is an advantage. This can be obtained with tilted antennas with $\alpha = 45°$. Taking the polarization decorrelation term $\rho_{\text{env}X\text{pol}}(\Gamma, \alpha = 45°)$ from (9) with the mean XPD values from Table IV, we find that for BS1 $\rho_{\text{env}X\text{pol}} = 0.48$ and for BS2 $\rho_{\text{env}X\text{pol}} = 0.75$. These decorrelations seem worthwhile to utilize either alone or combined with space diversity decorrelation, as normally a figure of 0.7 for the envelope correlation coefficient is considered satisfactory for noticeable diversity gain. It follows from Table III that these mean polarization decorrelations are indeed found when comparing the pure space antenna systems. Hand-held terminals are likely

TABLE III
AVERAGE (<>) AND STANDARD DEVIATION (σ) VALUES OF CW ENVELOPE CORRELATION COEFFICIENT AND BRANCH POWER DIFFERENCE.

Antenna system	ρ_{12} CW <>	σ	P_1-P_2 [dB] <>	σ
M1	0.14	0.12	-1.0	0.4
M2	0.07	0.08	-1.4	1.8
M3	0.21	0.27	-1.7	1.2
M4	0.09	0.12	-1.6	1.9
M5	0.07	0.09	0.6	1.7
M6	0.23	0.22	-0.8	2.0
M7	0.70	0.22	-0.7	1.7
M8	0.20	0.15	0.8	2.4
M9	0.09	0.09	-7.6	2.1
M10	0.13	0.16	-1.1	0.8
M11	0.04	0.07	-0.7	1.2
M12	0.12	0.15	-1.0	0.9
M21	0.30	0.30	1.2	3.1
M22	0.20	0.17	-0.2	1.5
M23	0.37	0.26	-1.0	0.6
M24	0.64	0.25	0.2	1.1
M25	0.36	0.16	0.8	1.1
M26	0.54	0.24	0.2	1.7

TABLE IV
AVERAGE (<>) AND STANDARD DEVIATION (σ) VALUES OF S, XPD, AND 1% LEVEL OF CUMULATIVE RSSI ENVELOPE DISTRIBUTION (cdf) FOR CW, 200 kHz, AND 1.2 MHz BANDWIDTHS.

Base	S [μs] <>	σ	XPD [dB] <>	σ	cdf_{CW} [dB] <>	σ	cdf_{XO} [dB] <>	σ	cdf_{12} [dB] <>	σ
BS1	0.44	0.36	7.4	2.0	-19.8	1.6	-15.7	2.2	-11.0	2.4
BS2	0.82	0.81	11.4	3.0	-19.6	1.9	-13.8	3.1	-9.1	3.1

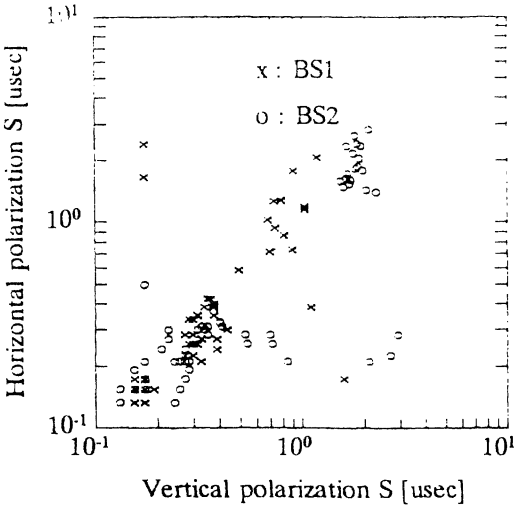


Fig. 7. RMS delay spread for decoupled (horizontal) polarization versus RMS delay spread for the co-polarized (vertical) signal.

to couple more power into the cross-polarized component, thereby increasing the polarization diversity gain potential.

The PDP's observed for both the co- and cross polarization showed very similar shapes except in the case of a dominant LOS peak in the copolarized case. Fig. 7 shows the RMS delay spread for the co- (vertical) and cross polarization. It is seen that most of the data lie around the 1:1 line, although some stray points are present. This indicates equal strength of the mean radio dispersion effects on the two polarizations.

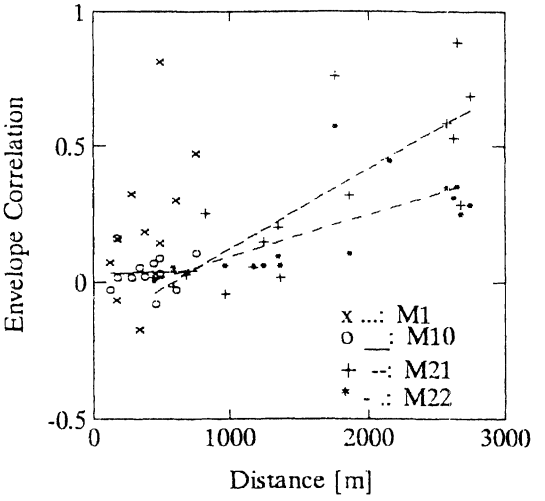


Fig. 8. Envelope correlation coefficient versus range length

C. Antenna Systems

The overall decorrelation properties of each antenna system are given in Table III. Recent theoretical work into base station cross correlation [11] has shown a model that can predict decorrelation of horizontal as well as vertical spaced antenna combinations. The model shows how the correlation increases with range length. The analysis in [11] is, though, mainly concentrated on longer range lengths (5 km) and larger antenna separations (10–20 λ) than used in our experiments. However, some comparisons can be made. From Table III, it follows that most antenna configurations have an average envelope correlation around 0.1 to 0.3 with a standard deviation around 0.2. Only the pure vertical space diversity systems show a significantly higher correlation of 0.6 to 0.7. For BS2, somewhat lower correlations are found compared with the examples given in [11] (cases with approximately 10 λ horizontal or vertical separation), with the horizontal separation being very sensitive to model parameter variation. For BS1, the base station antenna height is so low that the model assumptions in [11] are questionable and lead to much lower measured correlations than the model examples given. The modeling of compact spaced configurations, thus, seems to pose higher difficulty for cross correlation prediction, especially for very low base antenna heights as found in microcell implementations.

A trend of cross correlation depending on range (as expected from [11]) is clearly visible in Fig. 8, where cross correlation versus range is shown for some selected antenna systems. Local variations in environment may mask this effect (i.e., make it less pronounced), although the measurement area covering the locations around BS1 was chosen for its fairly homogenous nature with respect to building heights, street widths, and building density.

Fig. 9(a) shows the envelope correlation coefficient increase for the RSSI signals, compared to the CW signals, versus normalized delay spread $S \cdot BW$. There does not seem to be any clear dependence on radio dispersion for the overall data as expected from the high frequency selectivity of the channel, apart from a slight increase in correlation variability

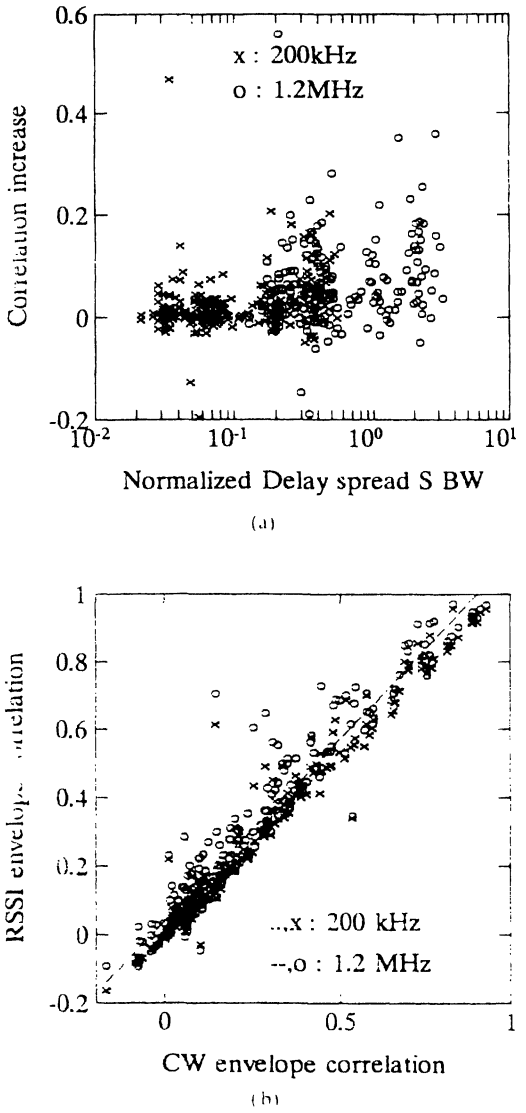


Fig. 9. (a) Envelope correlation increase with normalized delay spread. (b) RSSI versus CW envelope correlation coefficients.

with increasing dispersion. A very close agreement between CW and RSSI envelope correlations is seen in Fig. 9(b). Here it is shown that there is a slight mean and standard deviation increase of correlation for increasing bandwidth. Thus, the analysis for compound space and polarization systems, given in Appendix A, may also apply for wideband systems using RSSI driven combining even though the RSSI signals do not exhibit Rayleigh fading as assumed in the development of (9).

Generally, the antenna systems with pure vertical space separation diversity exhibit marginal decorrelation properties. All the other antenna systems exhibit very strong decorrelation properties, promising a high diversity gain.

VII. Conclusion

The main findings in this paper are in three areas.

1) From the general radio environment related investigations of the micro- and small cell areas, it follows that:

i) For the microcell area, the 0.7 coherence bandwidth exceeds 200 kHz at 90% of the locations. This justifies the use of narrowband considerations for GSM, in most cases.

ii) For the micro- and small cell areas, no location shows a 0.5 coherence bandwidth exceeding 400 kHz. Thus, the frequency hopping option of GSM will have a high frequency diversity gain potential.

iii) The cross polarization discrimination shows a large variance with location and mean values of 7.4 and 11.4 dB for the micro- and small cell areas. These means are larger than previously reported for macrocells in the same type environment.

2) The general evaluation of the antenna systems show:

i) The pure vertical space separation of an antenna system can yield acceptable decorrelation properties (mean ρ_{envV} around 0.6 to 0.7) for both micro- and small cell areas.

ii) All other antenna systems provide very high decorrelation efficiency (mean ρ_{env} around 0.1 to 0.3) and, thus, high diversity gain potential.

iii) Polarization diversity provides extra decorrelation in the compound antenna systems with about equal strength as the space decorrelation effect.

3) The analysis and experimental treatment of antenna system cross-correlation dependencies show that:

i) The polarization and space decorrelation effects in compound antenna systems are independent and multiplicative.

ii) RSSI signal decorrelation are largely insensitive to relative dispersion (bandwidth and RMS delay spread) for GSM-type systems.

The information contained in this paper verifies the effectiveness (diversity gain potential) of very compact antenna systems in micro- and small cell environments. The dimensions of the antenna systems lead to much easier mounting and site considerations than for standard mast-mounted macrocell implementations.

Appendix
Derivation and Experimental Verification of Compound Antenna System Cross-Correlation Function

We will now develop an expression for the CW envelope correlation coefficient for two tilted (orthogonal) antennas with respect to the CW envelope correlation coefficient for co-polarized antennas [(9) in Section VI]. In this analysis, however, power correlations are considered as they are easier to handle than envelope correlations. This is analytically justified by a close relation in value between the two [7]. This can also be seen experimentally in Fig. 10, where a very strong correlation (*cc*) of 0.99 is found between power and envelope correlation coefficients for the data presented in this paper. This relation has also been experimentally displayed in [8] for two vertically separated antennas with a slightly poorer fit, although $|\rho_{12}|^2$ was used instead of ρ_{pow}. While theoretically $|\rho_{12}|^2 = \rho_{pow}$ [7] for Rayleigh fading signals, it follows from Fig. 13 that experimentally this equality is slightly poorer than the approximate relation between the envelope and power correlation as shown in Fig. 10. Higher variances for the

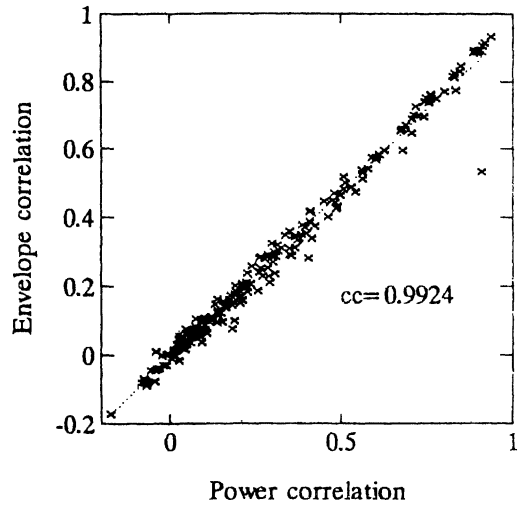


Fig. 10. Envelope correlation coefficient versus power correlation coefficient.

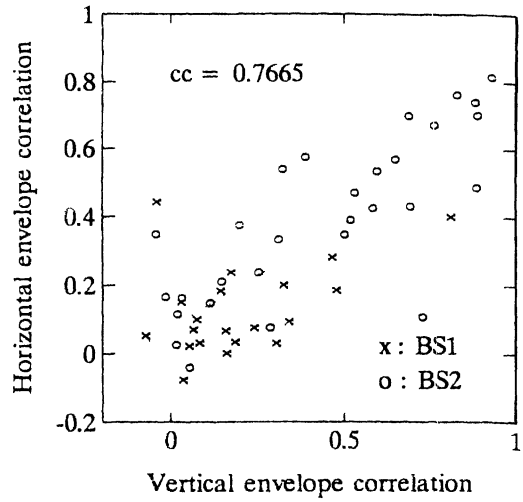


Fig. 11. Envelope correlation coefficient for decoupled (horizontal) polarization versus envelope correlation coefficient for the co-polarized (vertical) signal.

$|\rho_{12}|^2 = \rho_{pow}$ relation can be explained by two factors. For the relation to hold, the signals must be truly Rayleigh fading while both power and envelope correlations have proven to be fairly insensitive to changes in envelope dynamics (not all experimental results being Rayleigh distributed). Secondly, the experimental results allow for negative correlations but $|\rho_{12}|^2$ does not. The field components are assumed to have Rayleigh-distributed envelopes (r) with random uniformly distributed phases (θ) and Gaussian-distributed quadrature components. The field components of the two polarizations at both antennas are assumed uncorrelated in envelope and phase. For the envelope, this assumption is supported by [9], [12], and Table III.

Following Vaughan's [12] procedure, we find each antenna has vertical and horizontal Rayleigh fading field components as

$$E_{V1,2} = r_{V1,2}\cos(\omega t + \theta_{V1,2});$$
$$E_{H1,2} = r_{H1,2}\cos(\omega t + \theta_{H1,2}) \qquad \text{(A-1)}$$

with subscript 1 and 2 denoting the field at antennas 1 and 2, respectively.

Assuming the two antennas are horizontally and/or vertically displaced (space diversity), we have two power correlation coefficients ($\rho_{pow\,V}$ and $\rho_{pow\,H}$ for the vertical and horizontal components, respectively). From [10] and Fig. 11 (showing envelope correlations), it follows that:

$$\rho_{pow\,V} \approx \rho_{pow\,H} \qquad \text{(A-2)}$$

is a reasonable assumption. The correlation (cc) between the two is larger for macrocells (cc = 0.88) [10] than for the small and microcells considered here (cc = 0.77). In practice, this relation implies that the antennas have approximately equal radiation patterns, and that the radiation patterns in the vertical and horizontal plane are approximately equal, i.e., rotation symmetry of the radiation pattern around the axis of the main lobe. Fig. 12 shows the 1% cumulative envelope level for the cross and copolarized signals. It is seen that the envelope dynamics are also approximately equal for the two

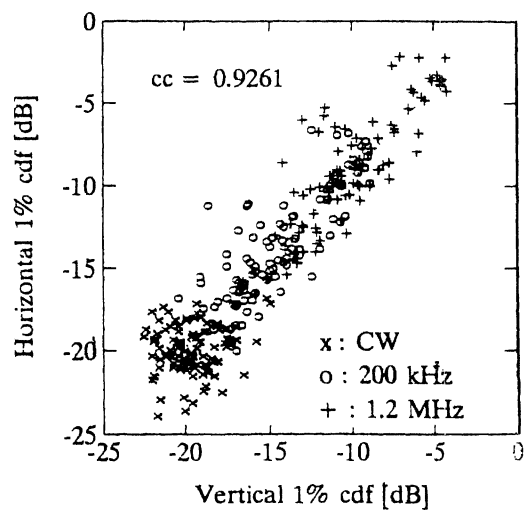


Fig. 12. 1% cumulative envelope levels for decoupled (horizontal) polarization versus co-polarized (vertical) signal.

polarizations. Furthermore, it follows from [10] and Table III that it is reasonable to assume approximately equal mean power at the two antennas for each polarization.

$$E[P_{V1}] \approx E[P_{V2}]; \qquad E[P_{H1}] \approx E[P_{H2}]. \qquad \text{(A-3)}$$

Thus, the XPD ($\Gamma = E[P_{V1}]/E[P_{H1}] \approx E[P_{V2}]/E[P_{H2}]$) is assumed approximately equal at the two antennas. The powers are $P_{V1,2} = r_{V1,2}^2$, $P_{H1,2} = r_{H1,2}^2$. As the envelopes of both the polarizations at each antenna ($r_{V,H1,2}$) are assumed Rayleigh distributed, the following relationship holds:

$$E[P_{V,H1,2}^2] = 2E[P_{V,H1,2}]^2. \qquad \text{(A-4)}$$

Using (A-4) and (A-3), we get the correlations

$$\rho_{pow\,V,H} \approx \frac{E[P_{V,H1}P_{V,H2}] - E[P_{V,H1}]^2}{E[P_{V,H1}]^2}. \qquad \text{(A-5)}$$

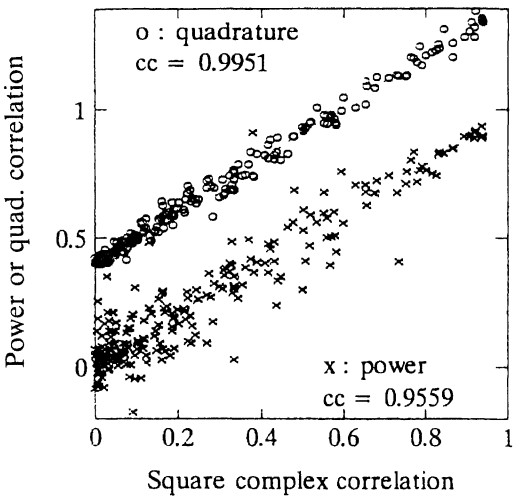


Fig. 13. Square quadrature component correlation coefficients ($\rho_{II}^2 + \rho_{IQ}^2$) and power correlation coefficient versus square module of complex correlation coefficient ($|\rho_{12}|^2$).

Comparing the two polarization correlations by substituting (A-5) into (A-2), we have

$$E[P_{V1}P_{V2}] \approx \Gamma^2 E[P_{H1}P_{H2}]. \tag{A-6}$$

Now consider the two antennas being rotated to an angle α to the vertical polarization axis. Then, the voltages received at antennas 1 and 2 are proportional to

$$U_1 = aE_{V1} + bE_{H1}: \qquad U_2 = bE_{V2} - aE_{H2} \tag{A-7}$$

with $a = \cos(\alpha)$ and $b = (\sin(\alpha)$ as the antenna signal quadrature weights. Following Vaughan's [12] procedure and substituting (A-1) into (A-7) and squaring, the antenna branch powers can be found as

$$\begin{aligned} P_1 &= [ar_{V1}\cos(\theta_{V1}) + br_{H1}\cos(\theta_{H1})]^2 \\ &\quad + [ar_{V1}\sin(\theta_{V1}) + br_{H1}\sin(\theta_{H1})]^2 \\ &= a^2P_{V1} + b^2P_{H1} + 2abr_{V1}r_{H1}\cos(\theta_{V1} - \theta_{H1}) \\ P_2 &= b^2P_{V2} + a^2P_{H2} - 2abr_{V2}r_{H2}\cos(\theta_{V2} - \theta_{H2}). \end{aligned} \tag{A-8}$$

The antenna branch voltages in (A-7) are a sum of two Rayleigh fading signals which produce a new Rayleigh fading signal as the quadrature components are Gaussian distributed). Then, the power relationship of (A-4) applies for P_1 and P_2 as well, and the power correlation coefficient ρ_{pow12} for the two tilted antennas is

$$\rho_{pow12} = \frac{E[P_1P_2] - E[P_1]E[P_2]}{E[P_1]E[P_2]}. \tag{A-9}$$

The mean power moments are found as [and reduced via (A-3)]

$$\begin{aligned} E[P_1] &= a^2E[P_{V1}] + b^2E[P_{H1}] \\ &\quad + 2abE[r_{V1}]E[r_{H1}]E[\cos(\theta_{V1} - \theta_{H1})] \\ &= a^2E[P_{V1}] + b^2E[P_{H1}] = E[P_{V1}]\left(a^2 + \frac{b^2}{\Gamma}\right) \\ E[P_2] &= b^2E[P_{V2}] + a^2E[P_{H2}] \approx E[P_{V1}]\left(b^2 + \frac{a^2}{\Gamma}\right) \\ &= E[P_1]\frac{\Gamma b^2 + a^2}{\Gamma a^2 + b^2} \end{aligned} \tag{A-10}$$

due to the phases being random, uniformly distributed, and uncorrelated between the polarizations. The cross-power moment is found as [and reduced via (A-3) and (A-6)]

$$\begin{aligned} E[P_1P_2] &= a^2b^2(E[P_{V1}P_{V2}] + E[P_{H1}P_{H2}]) \\ &\quad + a^4E[P_{V1}]E[P_{H2}] \\ &\quad + b^4E[P_{V2}]E[P_{H1}] + R(U_1, U_2) \\ &\approx a^2b^2\left(1 + \frac{1}{\Gamma^2}\right)E[P_{V1}P_{V2}] \\ &\quad + \frac{(a^4 + b^4)}{\Gamma}E[P_{V1}]^2 + R(U_1, U_2). \end{aligned} \tag{A-11}$$

$R(U_1, U_2)$ can be expressed directly by the quadrature components of field components given in (A-1) as

$$\begin{aligned} R(U_1, U_2) &= -4a^2b^2(E[I_{V1}I_{V2}]E[I_{H1}I_{H2}] \\ &\quad + E[Q_{V1}Q_{V2}]E[Q_{H1}Q_{H2}] \\ &\quad + E[I_{V1}Q_{V2}]E[I_{H1}Q_{H2}] \\ &\quad + E[I_{V2}Q_{V1}]E[I_{H2}Q_{H1}]) \end{aligned} \tag{A-12}$$

where the in-phase components are $I_{V,H1,2} = r_{V,H1,2}\cos(\theta_{V,H1,2})$ and the quadrature components $Q_{V,H1,2} = r_{V,H1,2}\sin(\theta_{V,H1,2})$. As for each polarization, the signals are assumed Rayleigh fading and the in-phase and quadrature components are Gaussian distributed with zero mean and equal variances ($\sigma_I^2 = \sigma_Q^2$). Using (A-3), the ratio between the variances for each polarization is

$$\begin{aligned} E[P] &= \sigma_I^2 + \sigma_Q^2 = 2\sigma^2 \\ \sigma_{IV,H1}^2 &= \sigma_{QV,H1}^2 \approx \sigma_{IV,H2}^2 = \sigma_{QV,H2}^2 \\ \Gamma &= \frac{\sigma_V^2}{\sigma_H^2}. \end{aligned} \tag{A-13}$$

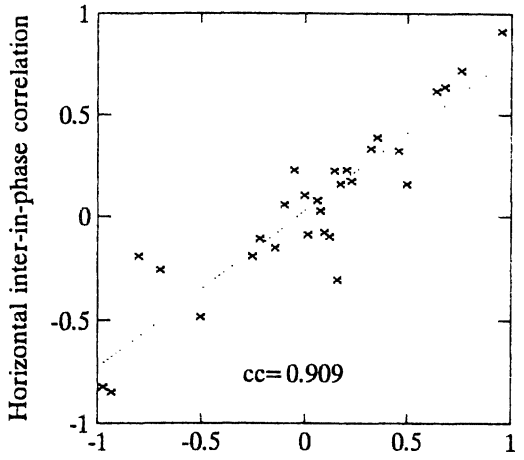
The quadrature component correlation coefficient is then found as

$$\begin{aligned} \rho_{IIV,H} &= \frac{E[I_{V,H1}I_{V,H2}]}{E[I_{V,H1}^2]} = \frac{E[Q_{V,H1}Q_{V,H2}]}{E[Q_{V,H1}^2]} \\ &= \frac{2E[I_{V,H1}I_{V,H2}]}{E[P_{V,H1}]}. \end{aligned} \tag{A-14}$$

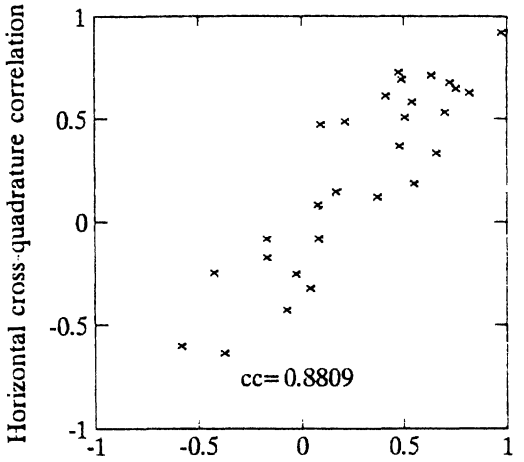
Similarly, the cross quadrature expectation becomes

$$E[I_{V,H1}Q_{V,H2}] = -E[Q_{V,H1}I_{V,H2}] = \rho_{IQV,H}\frac{E[P_{V,H}]}{2}. \tag{A-15}$$

For Rayleigh fading signals, it can be found that $\rho_{IQ} = -\rho_{QI}$ and $\rho_{II} = \rho_{QQ}$ [13]. Furthermore, the complex correlation is related as $\rho_{12} = 1/2(\rho_{II} + \rho_{QQ}) + j1/2(\rho_{IQ} - \rho_{QI} = \rho_{II} + j\rho_{IQ}$. From [7], it follows that $\rho_{II}^2 + \rho_{IQ}^2 = |\rho_{12}|^2 = \rho_{powV}$. This is shown in Fig. 13 for the data presented in this paper. The quadrature component correlations (offset by 0.4 for clarity) have an extremely high correlation ($cc = 0.995$) with the squared complex correlation coefficients. The power



(a)



(b)

Fig. 14. Horizontal versus vertical (a) inter-in-phase (ρ_{II}) and (b) cross-quadrature (ρ_{IQ}) correlation coefficients for BS2.

correlation coefficients have a slightly poorer correlation (cc = 0.96), but are still very high.

With the previous discussion, we can reduce (A-12) (with the use of (A-14) and (A-15) and rearranging) to

$$R(U_1 U_2) = -2\frac{a^2 b^2}{\Gamma} E[P_{V1}]^2(\rho_{IIV}\rho_{IIH} + \rho_{IQV}\rho_{IQH}). \tag{A-16}$$

It follows from Fig. 14, that it is reasonable to assume that the quadrature component correlations are approximately equal for the two polarizations, though the shown data only represent BS2. At BS1, the antenna cables were changed between the vertical and horizontal polarization experiment. Thus, the phase offset on the complex correlations will be changed from the vertical to horizontal experiment. Consequently, the ρ_{II} and ρ_{IQ} relations between the two polarizations at BS1 are altered and not usable. With this discussion and inserting (A-16) into (A-11), the power correlation coefficient in (A-9) reduces to the expression (9) shown in Section VI.

Acknowledgment

The authors acknowledge the participation of C. Jensen in the experimental part of this work during his employment at Aalborg University, and H. Fredskild from Telecom Denmark Telelaboratoriet for providing a second matching base-station antenna. The comments of the reviewers, who helped enhance the quality of this paper, are greatly appreciated.

References

[1] B. Glance and L. J. Greenstein, "Frequency-selective fading effects in digital mobile radio with diversity combining" *IEEE Trans. Commun.*, vol. 31, no. 2, pp. 1085–1094, Sept. 1983.

[2] J. B. Andersen, P. C. F. Eggers, and B. L. Andersen, "Propagation aspects of datacommunications over the radio channel—A tutorial," in *Proc. EUROCON '88,* Stockholm, Sweden, pp. 301–307.

[3] B. L. Andersen, P. C. F. Eggers, and J. B. Andersen, "Time and phase variations in the mobile channel," in *Proc. Nordic Radio Symp. '89,* Saltsjöbaden, Sweden, pp. 297–304.

[4] P. E. Mogensen, P. C. F. Eggers, C. Jensen, and J. B. Andersen, "Urban area radio propagation measurements at 955 and 1845 MHz for small and micro cells," in *Proc. IEEE GLOBECOM'91,* pp. 1297–1302.

[5] P. E. Mogensen, "Preliminary results from short-term measurements in urban area," COST 231 TD(90)-88, Paris, France, Oct. 8–11, 1990.

[6] A. A. M. Saleh and R. A. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Select. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[7] J. N. Pierce and S. Stein, "Multipath diversity with nonindependent fading," *IRE Proc.*, pp. 89–104, Jan. 1960.

[8] M. T. Feeney and J. P. Parsons, "Cross-correlation between 900 MHz signals received on vertically separated antennas in small-cell mobile radio systems," *IEE Proc.*, vol. 138, no. 2, pp. 81–86, Apr. 1991.

[9] S. Kozono, T. Tsuruhara, and M. Sakamoto, "Base station polarization diversity reception for mobile radio," *IEEE Trans. Vehic. Technol.*, vol. 33, no. 4, pp. 301–306, Nov. 1984.

[10] P. C. F. Eggers and J. B. Andersen, "Base station diversity for NMT900," in *Proc. Nordic Radio Symp.*, Saltsjöbaden, Sweden, pp. 77–85.

[11] A. M. D. Turkmani and J. P. Parsons, "Characterisation of mobile radio signals: base station crosscorrelation," *IEE Proc.*, vol. 138, no. 6, pp. 557–565, Dec. 1991.

[12] R. G. Vaughan, "Polarization diversity in mobile communications," *IEEE Trans. Vehic. Technol.*, vol. 39, no. 3, pp. 177–186, Aug. 1990.

[13] D. E. Kerr, *Propagation of Short Radio Waves.* Boston, MA: Boston Tech. Pub. Inc., 1964.

A Statistical Model for Angle of Arrival in Indoor Multipath Propagation

Quentin Spencer, Michael Rice, Brian Jeffs, and Michael Jensen
Department of Electrical & Computer Engineering
Brigham Young University
Provo, Utah 84602

Abstract— **Multiple antenna systems are a useful way of overcoming the effects of multipath interference, and can allow more efficient use of spectrum. In order to test the effectiveness of various algorithms such as diversity combining, phased array processing, and adaptive array processing in an indoor environment, a channel model is needed which models both the time and angle of arrival in indoor environments. Some data has been collected indoors and some temporal models have been proposed, but no existing model accounts for both time and angle of arrival. This paper discusses existing models for the time of arrival, experimental data that were collected indoors, and a proposed extension of the Saleh-Valenzuela model [1], which accounts for the angle of arrival. Model parameters measured in two different buildings are compared with the parameters presented in the paper by Saleh and Valenzuela, and some statistical validation of the model is presented.**

I. Introduction

There have been many different approaches for overcoming the problem of multipath interference, both in outdoor and indoor applications. Some of them include channel equalization, directional antennas, and multiple antenna systems, each being more particularly suited to different applications. The use of multiple antenna systems can be particularly useful for indoor applications such as local area networks, because they allow the possibility of communicating with multiple users simultaneously over a single frequency band, increasing throughput and making efficient use of frequency spectrum. The signals from different antennas can be combined in various ways, including diversity combining, phased array processing, and adaptive array algorithms. Adaptive array sytems are becoming increasingly feasible for high bandwidth applications with continuing improvements in digital signal processors. In addition, the availability of new, higher frequency bands has made wireless networks an increasinly attractive and feasible option. The effects of multipath interference have been studied extensively in various outdoor scenarios. However, the study of the indoor multipath channel is relatively new. In order to be able to predict the performance of indoor communications systems, models are needed that accurately model the behavior of radio transmissions in indoor environments.

Several other researchers have already collected various types of data on indoor mulipath propagation. The foundation for much of today's work was by Turin, et al [2], which was a study of outdoor multipath propagation in an urban environment. The first model for indoor multipath propagation was proposed by Saleh and Valenzuela [1], whose work was based on the work of Turin. Their work consisted of collecting temporal data on indoor propagation, from which they proposed a time domain model for indoor propagation.

Most indoor propagation research has dealt with the time of arrival and paid little attention to the angle of arrival. In order to predict the performance of adaptive array systems, the angle of arrival is very important information. Some recent papers have begun to address the angle of arrival. Lo and Litva [3] found that multipath arrivals tend to occur at varying angles indoors, but were not able to arrive at any conclusions based on their limited data. Guerin [4] collected angular and temporal data separately, but did not correlate the two. Wang, et al [5], used a rectangular array to estimate both the elevation and azimuth angles of arrival for major multipaths, but did not measure the corresponding time of arrival. Litva, et al, [6] collected simultaneous time and angle of arrival data, similar to the format of the data used in this paper. They came to the preliminary conclusion that it is possible to make accurate measurements of this type and learn more about what is happening in the indoor multipath channel. However, their experiment was not extensive enough to make any conclusions about the channel.

This paper presents an extension to the Saleh-Valenzuela model which accounts for the angle of arrival. This is based on data that includes information about both the time and angle of arrival, presented in [7]. The Saleh-Valenzuela model is explained, and the new data is discussed. Model parameters based on the new data are derived and compared to the parameters found by Saleh and Valenzuela at a lower frequency.

II. The Saleh-Valenzuela Model

The model proposed by Saleh and Valenzuela is based on a clustering phenomenon observed in their experimental data. In all of their observations, the arrivals came in one or two large groups within a 200 ns observation window. It was observed that the second clusters were attenuated in amplitude, and that rays, or arrivals within a single cluster, also decayed with time. Their model proposes that both of these decaying patterns are exponential with time, and are controlled by two time constants: Γ, the cluster arrival decay time constant, and γ, the ray arrival decay time constant. Fig. 1 illustrates this, showing the mean envelope of a three cluster channel.

Reprinted from *IEEE Vehicular Technology Conference*, pp. 1415–1419, May 1997.

The impulse response of the channel is given by:

$$h(t) = \sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\beta_{kl}\delta(t - T_l - \tau_{kl}), \quad (1)$$

where the sum over l represents the clusters, and the sum over k represents the arrivals within each cluster. The amplitude of each arrival is given by β_{kl}, which is a Rayleigh distributed random variable, whose mean square value is described by the double-exponential decay illustrated in Fig. 1. Mathematically it is given by:

$$\overline{\beta_{kl}^2} = \overline{\beta^2(T_l, \tau_{kl})} \quad (2)$$

$$= \overline{\beta^2(0,0)}e^{-T_l/\Gamma}e^{-\tau_{kl}/\gamma}, \quad (3)$$

where $\overline{\beta^2(0,0)}$ is the average power of the first arrival of the first cluster. This average power is determined by the separation distance of transmitter and receiver.

The time of arrival is described by two Poisson processes which model the arrival times of clusters and the arrival times of rays within clusters. The time of arrival of each cluster is an exponentially distributed random variable conditioned on the time of arrival of the previous cluster. The case is the same for each ray, or arrival within a cluster. Following the terminology used by Saleh and Valenzuela, rays shall refer to arrivals within clusters, so that the cluster arrival rate implies the parameter for the intercluster arrival times and the ray arrival rate refers to the parameter for the intracluster arrival times. The distributions of these arrival times are shown in equations 4 and 5:

$$p(T_l|T_{l-1}) = \Lambda e^{-\Lambda(T_l - T_{l-1})} \quad (4)$$

$$p(\tau_{kl}|\tau_{(k-1)l}) = \lambda e^{-\lambda(\tau_{kl} - \tau_{(k-1)l})}, \quad (5)$$

where Λ is the cluster arrival rate, and λ is the ray arrival rate. In their data, Saleh and Valenzuela did not have any information on angle of arrival, and assumed that the angles of arrival were uniformly distributed over the interval $[0, 2\pi)$.

Other indoor multipath models have been proposed, such as the model proposed by Ganesh and Pahlavan [8], but they will not be discussed here. The data used in this paper fit the Saleh-Valenzuela model well, and as a result the model was chosen as the basis for the extended model presented here.

III. Experimental Data

In order to analyze and model the indoor multipath channel, a data gathering apparatus was designed which was able to take simultaneous measurements of the time and angle of arrival. The frequency band was from 6.75 to 7.25 GHz. Using the system, a total of 65 data sets were collected in two buildings on the Brigham Young University Campus. In the Clyde building, a reinforced concrete and cinder block building, 55 data sets were collected. For comparison, ten additional data sets were collected in the Crabtree Building, constructed mostly of steel and gypsum board. Each data set can be viewed as an image plot, with angle as one axis, and time as the second axis. A typical data set is pictured in Fig. 2. The images were processed to remove blurring effects so that the precise time, angle and amplitude of each major multipath arrival is known. The data collection and processing is discussed in greater detail in [7].

Visual observation of the data showed that clustering like that observed by Saleh and Valenzuela was present in the data. The nature of the clustering tended to follow the model of Saleh and Valenzuela quite well. In general, the strength of clusters tended to decay with increasing delay times, and arrivals within each cluster showed a similar pattern of decay. One difference from the Saleh-Valenzuela data is the higher average number of clusters per data set.

IV. A Proposed Time/Angle Model for Indoor Multipath Propagation

In this section we propose a statistical model for the indoor multipath channel that includes a modified version of the Saleh-Valenzuela model, and incorporates an angle-of-arrival model. In addition, methods of estimating parameters from the data are discussed.

A. Time of Arrival

The time and amplitude of arrival portion of the combined model is represented by $h(t)$ in equation (1), where, as before, $\overline{\beta_{kl}^2}$ is the mean square value of the kth arrival of the lth cluster. This mean square value is described by the exponential decay given in equation (3) and illustrated in Fig. 1.

As before, the ray arrival time within a cluster is given by the Poisson distribution of equation (5), and the first arrival of each cluster is given by T_l, described by the Poisson distribution of (4). The inter-ray arrival times, τ_{kl}, are dependent on the time of the first arrival in the cluster T_l. In the Saleh-Valenzuela model, the first cluster time T_1 was dependent on T_0 which was assumed to be zero. With the estimated parameter in [1] of $1/\Lambda \approx 300$ ns, the first arrival time will typically be in the range of 200 to 300 ns, which is a reasonable figure. However, a problem with this was found when the Λ parameter in the new data was discovered to be very low, but the delay time to the first arrival was often still on the order of 200 ns. Under the Saleh-Valenzuela model, this would make any long delays which would occur at larger separation distances between transmitter and receiver highly improbable. To remedy this problem, it is proposed that T_0 be the line of sight propagation time:

$$T_0 = \frac{r}{c}, \quad (6)$$

where c is the speed of light, and r is the separation distance. This allows for the time of the first arrival to be more directly dependent on the separation distance.

B. Angle of Arrival

It will be assumed that time and angle are statistically independent. If there were a correlation, it would be expected that a longer time delay would correspond to a larger angular variance from the mean of a cluster. This was not observed in the data, so at this point an assumption of independence is reasonable, but further study of the correlation structure may be warranted. The consequence of this independence is that the complete impulse response with respect to both time and angle, which we will call $h(t,\theta)$, becomes a separable function:

$$h(t,\theta) \approx h(t)h(\theta). \quad (7)$$

As a result, $h(\theta)$ can be be addressed separately from $h(t)$.

We propose an independent angular impulse response of the system, similar to the time impulse response of the channel given in 1:

$$h(\theta) = \sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\beta_{kl}\delta(\theta - \Theta_l - \omega_{kl}), \quad (8)$$

where, as before, β_{kl} is the ray amplitude for the kth arrival in the lth cluster, given in equations (2) and (3). Θ_l is the mean angle of each cluster, which is distributed uniformly on the interval $[0, 2\pi)$. We propose that the ray angle within a cluster, ω_{kl}, be modeled as a zero mean Laplacian distribution with standard deviation σ:

$$p(\theta) = \frac{1}{\sqrt{2}\sigma}e^{-|\sqrt{2}\theta/\sigma|}. \quad (9)$$

The correlation of these distributions to the data is shown in the next section.

C. Parameter Estimation

This section outlines methods of deriving the distributions and estimating the parameter σ given in the previous section. The distribution parameters of the cluster means, Θ_l, is found by identifying each of the clusters in a given data set. The mean angle of arrival for each cluster is calculated. In order to remove the specific room geometry and orientation, the first arrival (in time) for each data set is taken as the reference. The relative cluster means are calculated by subtracting the mean of the reference cluster from all other cluster means. To estimate the distribution of cluster means over the ensemble of all data sets, a histogram can be generated of all relative cluster means, disregarding the first clusters (since their relative mean is always 0).

The procedure to estimate σ is similar. The cluster mean is subtracted from the absolute angle of each ray in the cluster to give a relative arrival angle with respect to the cluster mean. The relative arrivals are collected over the ensemble of all data sets, and a histogram can be generated. Using a least mean square algorithm, the histogram is fit to the closest Laplacian distribution, which gives the value for σ.

D. Using the Model

The extended model for $h(t,\theta)$ is useful for analysis or simulation of array processing algorithms that might be used in an indoor environment. In order, for example, to conduct a Monte Carlo simulation of an array antenna processor, it is necessary to generate a random channel using the statistical model. This section outlines the procedure for doing so.

The first step is to choose the transmitter/receiver separation distance r, which can be chosen either randomly or arbitrarily. Knowing r, the next step is to determine $\overline{\beta^2(0,0)}$, the mean power of the first arrival, which is given by

$$\overline{\beta^2(0,0)} \approx \frac{G(1m)r^{-\alpha}}{\gamma\lambda}, \quad (10)$$

where $G(1\text{m})$ is the channel gain at $r = 1$ meter, and α is a channel loss parameter. γ and θ are respectively the ray decay parameter and ray arrival rate in the model for $h(t)$. Equation (10) is derived and the characteristics of α in the indoor environment are discussed in greater detail in [1].

After $\overline{\beta^2(0,0)}$ is determined, the next step is to determine the cluster and ray arrival times. The corresponding distributions are given in equations (4) and (5), where $T_0 = r/c$. After the times are determined, the mean amplitudes $\overline{\beta_{kl}}$ are determined by equation 3. The actual amplitudes for each arrival, β_{kl}, are determined by sampling a Rayleigh distribution whose mean is $\overline{\beta_{kl}}$.

The angles are determined by first randomly choosing the cluster angles, which are uniformly distributed from 0 to 2π. Relative ray angles are then determined by sampling a Laplacian distribution as given in equation (9).

V. Model Parameters from the Data

The intercluster time decay constant, Γ, was estimated by normalizing the cluster amplitudes (the amplitude of the first arrival) so that the first one had an amplitude of 1 and a time delay of 0. All of the cluster amplitudes were superimposed as shown in Fig. 3. The estimate for Γ was found by curve fitting the line (representing an exponential curve) to minimize the mean squared error. The values for Γ and γ were estimated for both buildings in a similar manner. In this particular example, the fit is less than ideal, but it was better in the other cases, especially when there were more data points. In their data, Saleh and Valenzuela did not have exact amplitudes available, and as a result were not able to use curve fitting or generate plots as in Fig. 3. Their parameters were as a result very rough estimates, but they did observe the same general decay trend as in this data, which supports the exponential decay model.

The Poisson parameters, Λ and λ, representing the intercluster and intracluster arrival rates were estimated by subtracting each arrival time from its predicessor to produce a set of conditional arrival times $p(\tau_{kl}|\tau_{(k-1)l})$. The

parameter	Clyde Building	Crabtree Building	Saleh-Valenzuela
Γ	33.6 ns	78.0 ns	60 ns
γ	28.6 ns	82.2 ns	20 ns
$1/\Lambda$	16.8 ns	17.3 ns	300 ns
$1/\lambda$	5.1 ns	6.6 ns	5 ns
σ	25.5°	21.5°	—

Table 1. A comparison of model parameters for the two buildings and from the Saleh-Valenzuela paper [1]

probability distribution of these with the best fitting pdf (for the Clyde Building) is shown in Fig. 4.

Fig. 5 shows a CDF of the relative cluster angles for the Clyde Building, illustrating the relatively uniform distribution of clusters in angle. The same was true in the Crabtree Building. The distribution of the ray arrivals with respect to the cluster mean is shown in Fig. 6. The sharp peak at the mean is characteristic of a Laplacian distribution. The superimposed curve is a Laplacian distribution that was fit by integrating a Laplacian PDF over each bin, and matching the curves using a least mean square goodness of fit measure. The Laplacian distribution turns out to be a very close fit in both buildings.

Table 1 shows a comparison of the model parameters estimated for the Clyde Building, the Crabtree Building, and those estimated by Saleh and Valenzuela from their data. The most obvious discrepancy is in the estimates for the value of Λ. This is due to the fact that there were significantly more clusters observed in both the Clyde and Crabtree buildings compared to an average of 1-2 clusters observed by Saleh and Valenzuela. This may be partly due to the higher RF frequency, but the more likely cause is the ability of our testbed to see clusters that were close together in time, but separated in angle. Another interesting phenomenon is that Γ is very low in the Clyde Building, and γ is larger than Γ in the Crabtree Building, meaning that the Clyde Building tends to attenuate more than the Crabtree Building. The values of σ were close in both buildings, and there is no precedent for comparison with other data.

VI. Conclusion

Many aspects of the model have plausible physical explanations. Because an absolute angular reference was maintained during the collection of the data, it was possible to compare the processed data with the geometry of each configuration. The strongest cluster was almost always associated with the direct line of sight, even when there were walls blocking the line of sight path. Apparent causes of weaker clusters were back wall reflections and doorway openings. It is likely that each cluster corresponds to a major path to the receiver, and the arrivals within each cluster are likely the result of smaller, closely associated objects that are part of a very similar group of paths to the receiver. These paths will take slightly longer to arrive than the first arrival in the cluster, and are usually attenuated relative to this first arrival.

The amplitudes of clusters and rays within clusters both follow the same pattern of exponential decay over time observed by Saleh and Valenzuela. The differences in model parameters are likely due to the difference in frequency (Saleh and Valenzuela used 1.5 GHz). The other discrepancy is in the markedly faster cluster arrival rate, which is most likely explained by the larger overall number of clusters resulting from a more sensitive data gathering apparatus. The model parameters for the Clyde and Crabtree Buildings were in general very similar. The most notable exception is the extremely slow amplitude decay of rays within a cluster in the Crabtree building. In general, the model seemed to be able to accurately describe the differing multipath characteristics in both buildings, regardless of their very different construction. This implies that the model could possibly provide a general representation for many different types of buildings, and model parameters could therefore be found for other types of buildings.

The angle-of-arrival model presented here, though yet unconfirmed, is a strong alternative to only previous option for simulation: random assignment of angles or guessing at the anglular properties of the channel. The most important area for continued research is applying the model for its intended purpose–comparison of array processing algorithms. This can be done either by mathematical analysis or Monte Carlo simulation. A mathematical analysis is likely intractible due to the large number of variables in the model, but the model can be a very useful tool for the generation of random multipath channels for simulation.

References

[1] Adel A. M. Saleh and Reinaldo A. Valenzuela. A statistical model for indoor multipath propagation. *IEEE Journal on Selected Areas of Communications*, SAC-5:128–13, February 1987.

[2] George L. Turin et al. A statistical model of urban multipath propagation. *IEEE Transactions on Vehicular Technology*, VT-21(1):1–9, February 1972.

[3] T. Lo and J. Litva. Angles of arrival of indoor multipath. *Electronics Letters*, 28(18):1687–1689, August 27 1992.

[4] Stephane Guerin. Indoor wideband and narrowband propagation measurements around 60.5 ghz. in an empty and furnished room. In *IEEE Vehicular Technology Conference*, pages 160–164, 1996.

[5] Jian-Guo Wang, Ananda S. Mohan, and Tim A Aubrey. Angles-of-arrival of multipath signals in indoor environments. In *IEEE Vehicular Technology Conference*, pages 155–159. IEEE, 1996.

[6] John Litva, Amir Ghaforian, and Vytas Kezys. High-resolution measurements of aoa and time-delay for characterizing indoor propagation environments. In *IEEE Antennas and Propagation Society International Symposium 1996 Digest*, volume 2, pages 1490–1493. IEEE, 1996.

[7] Quentin Spencer, Michael Rice, Brian Jeffs, and Michael Jensen. Indoor wideband time/angle of arrival multipath propagation results. In *IEEE Vehicular Technology Conference*. IEEE, 1997.

[8] R. Ganesh and K. Pahlavan. Statistical modeling and computer simulation of indoor radio channel. *IEE Proceedings-I*, 138(3):153–161, June 1991.

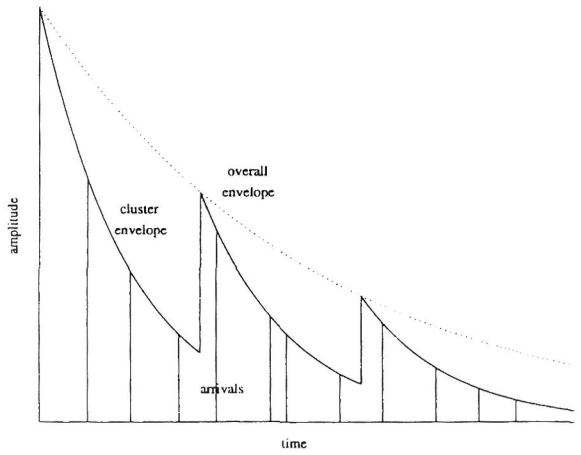


Fig. 1. An illustration of exponential decay of mean cluster power and ray power within clusters

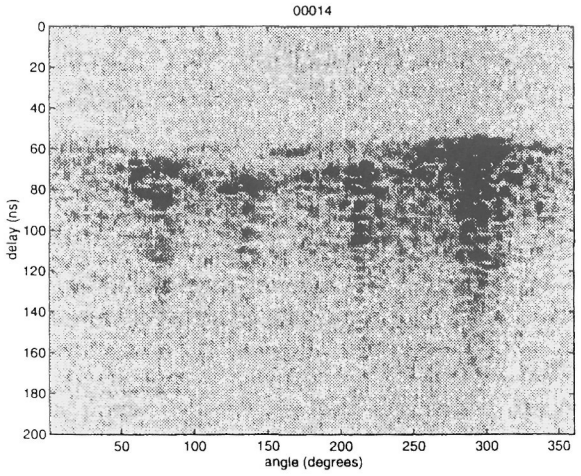


Fig. 2. A typical raw data set

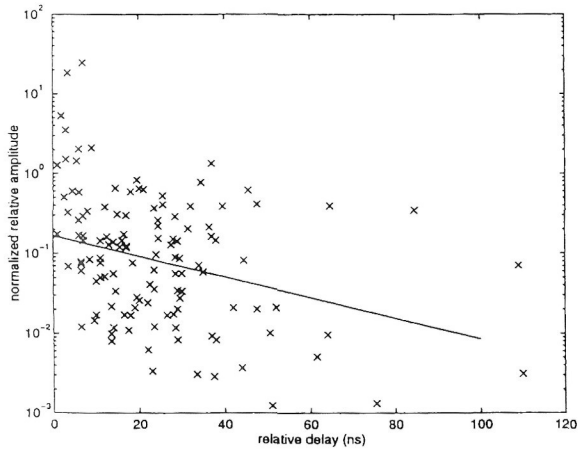


Fig. 3. Plot of normalized cluster amplitude vs. relative delay for the Clyde Building, with the curve for $\Gamma = 33.6$ ns superimposed.

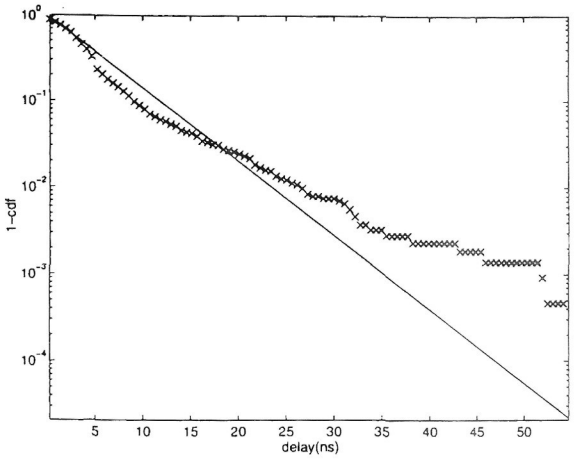


Fig. 4. CDF of Relative Arrival Times Within Clusters in the Clyde Building ($1/\lambda = 5.1$ns)

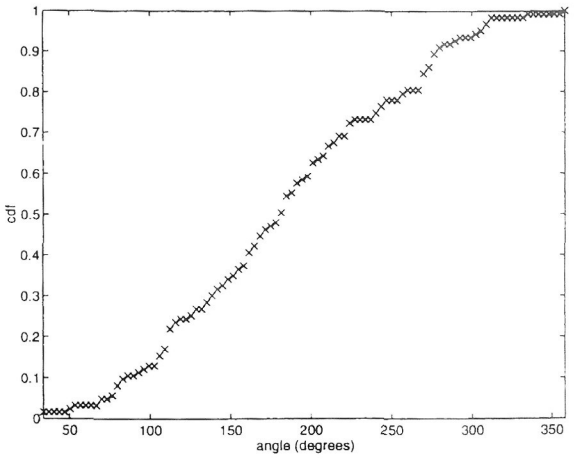


Fig. 5. CDF of relative mean cluster angles in the Clyde Building with respect to the first cluster in each set

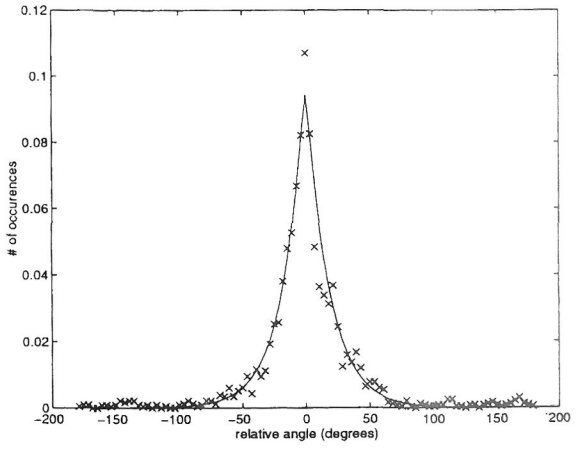


Fig. 6. Histogram of relative ray arrivals with respect to the cluster mean for the Clyde Building. Superimposed is the best fit Laplacian distribution ($\sigma = 25.5°$).