Antenna Diversity in Mobile Communications

RODNEY G. VAUGHAN, MEMBER, IEEE, AND J. BACH ANDERSEN, SENIOR MEMBER, IEEE

Abstract—**The conditions for antenna diversity action are investigated. In terms of the fields, a condition is shown to be that the incident field and the far field of the diversity antenna should obey (or nearly obey) an orthogonality relationship. The role of mutual coupling is central, and it is different from that in a conventional array antenna. In terms of antenna parameters, a sufficient condition for diversity action for a certain class of high gain antennas at the mobile, which approximates most practical mobile antennas, is shown to be zero (or low) mutual resistance between elements. This is not the case at the base station, where the condition is necessary only. The mutual resistance condition offers a powerful design tool, and examples of new mobile diversity antennas are discussed along with some existing designs.**

I. Introduction

THE DEMAND for better spectrum efficiency in narrow-band cellular frequency reuse systems can be eased by the application of antenna diversity. The possible improvements from diversity are well known for reduction of fading, but there are other advantages potentially available in the case of mobile communications. These are the suppression of both the random FM, which limits BER improvement in angle modulated systems, and cochannel interference, which limits frequency reuse base station density.

The signal conveyed through a narrow-band mobile channel becomes impaired by long-term (shadow) fading, short-term (Rayleigh-like) fading, random FM (including click noise), and especially in cellular systems, cochannel interference. Perhaps the most serious of these is the Rayleigh-like fading caused by the multipath environment. The random FM is caused by the Doppler shifts of the multipath signals, and the click noise component is associated with the deeper fades. The shadow fading is caused by a lack of power density, and this problem cannot be solved by diversity action at the mobile alone. The macrodiversity action required, if necessary, to overcome shadow fading is accomplished by strategically sited base stations. Macrodiversity will not be addressed here.

The simplest technique to maintain acceptable channel capacity (relative to the nonfading channel) is to increase the transmitted power. However, in doing so, the overall spectrum efficiency is reduced because the distance between frequency reuse transmitters must be greater to maintain acceptable cochannel interference levels. Moreover, the random FM cannot be suppressed by simply increasing the transmitted power. Alternative techniques to maintain channel capacity employ some kind of diversity scheme. Both antenna and signaling based diversity systems are well known (e.g., Jakes [13]).

With antenna diversity, the problems of the mobile channel are attacked directly. Higher orders of diversity are readily available in principle. An existing mobile antenna can be replaced by a diversity antenna with combiner so that existing systems can be improved without the need for implementing a signaling diversity scheme. The random FM is suppressed according to the order of diversity and the combining technique.

There are well-known schemes other than antenna diversity for improving the mobile channel capacity. Proponents of antenna diversity view the inherent advantages as follows. While covering "system" and "overall" spectrum efficiencies requires much discussion, it is sufficient here to note that

1) antenna diversity improves the channel capacity at the expense of adding extra equipment (antenna, combiner) to the receive end of the link (no extra spectrum is consumed); and
2) all other schemes consume extra spectrum to improve the channel capacity.

Regarding the first point, it is worth adding that adaptive retransmission with feedback allows the diversity antenna to be at the transmitting end of the link. The price paid is the required coding and housekeeping functions at both ends of the link with a corresponding slightly degraded channel message capacity compared to the receive antenna diversity case. A possible exception to the second point is delay diversity, in which uncorrelated signals arriving at different delay times are aligned (in time) for combination (cf. Rake and Drake schemes). There is no guarantee, however, that the natural delay distribution is suitable in the general case and so the scheme is not deemed appropriate.

The traditional disadvantage of antenna diversity is the cost and inconvenience of the extra equipment. There is much concern regarding efficient use of the spectrum, so it seems a matter of time until this concern forces greater use of antenna diversity. Much recent effort has been toward data coding to improve the information bit error rate (BER). Considerable progress has been made using *a priori* knowledge of the channel. Specifically, the Rayleigh-like fading gives rise to bursts of errors during the deeper fades. The channel is often treated as having "good" and "bad" states of transmission in a scheme known as the Gilbert–Elliot model (e.g., Ahlin, [1]). Most coding schemes rely on the channel signal-to-noise ratio (SNR) being exactly Rayleigh distributed, so the calculated

Manuscript received May, 9, 1986; revised May 10, 1987.

R. G. Vaughan is with the Department of Scientific and Industrial Research, Physics and Engineering Laboratory, Gracefield Road, Gracefield, Private Bag, Lower Hutt, New Zealand.

J. Bach Andersen is with the Institute of Electronic Systems, Aalborg University, Fr. Bajers Vej 7, 9220 Aalborg Ø, Denmark.

IEEE Log Number 8718834.

0018-9545/87/1100-0149$01.00

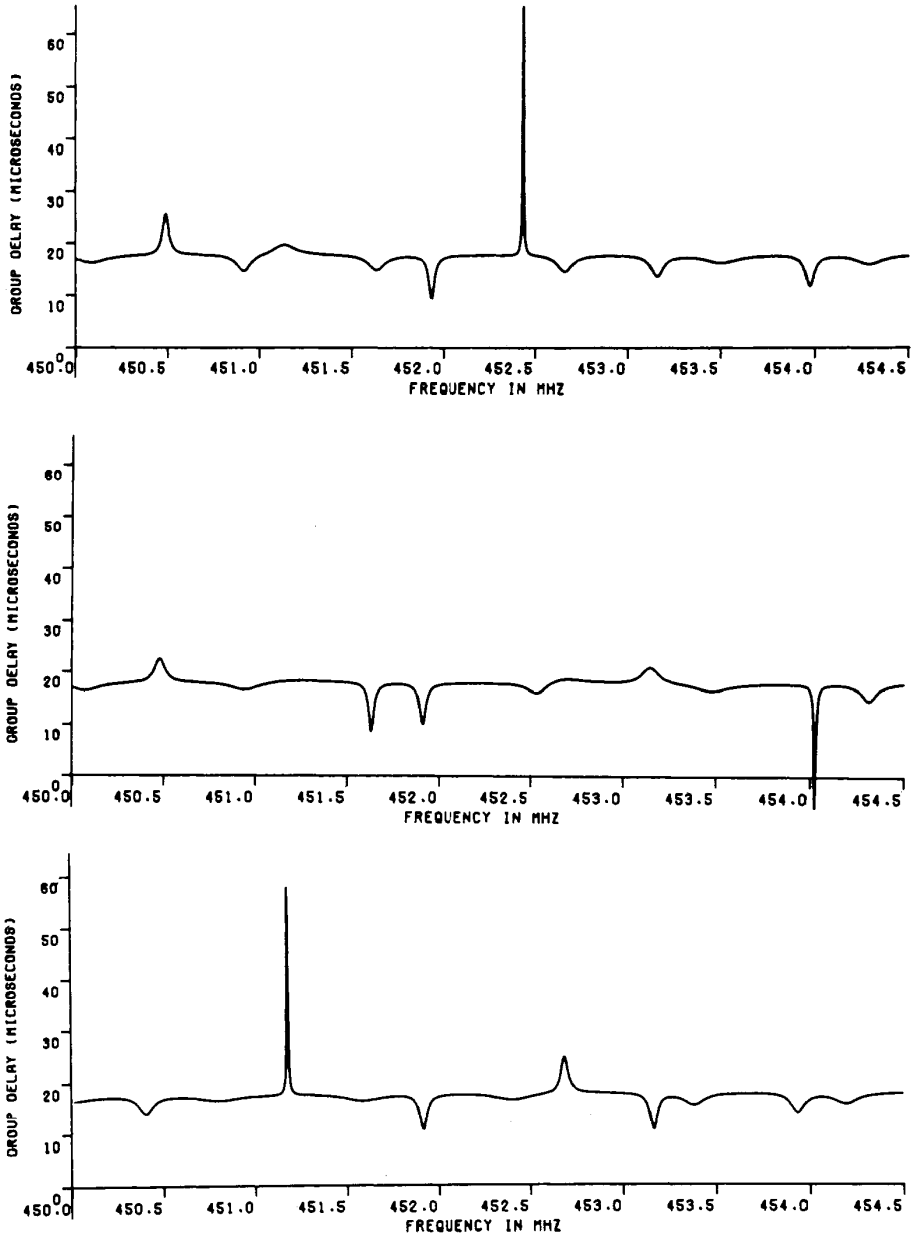


Fig. 1. Illustration that antenna diversity can also work for wide-band (frequency hopping) systems. The three figures are group delays from simulation of three diversity antenna elements. Dispersive (bad) channels are independent for each element. Average group delay is about 17 μs, which is exaggerated for clarity on the scale (it is typically less than 0.5 μs).

performance may well be quite different from actual performance. To the authors' knowledge, detailed investigation of the coding gain from a diversity antenna signal have not been reported. This should be a rather straightforward step, since the model with diversity would involve only a modification to the Rayleigh distribution term (maximum ratio combination could be assumed).

Much effort has also been expended on wide-band systems. The spread spectrum approach seems to be necessary for implementation of optimum combining, which is discussed by Winters [34]. Frequency hopping schemes (often referred to as frequency diversity) do not seem to have been implemented in public systems to date. It is worth noting that antenna diversity offers potential channel improvement for wide-band systems also. The scheme is illustrated by simulation results in Fig. 1, which shows that the group delays are uncorrelated between branches, so that a highly dispersive channel in one branch will be well behaved in another. The group delay characteristics in a wide-band system are analogous to the random FM in the narrow-band case. There is an "irreducible" BER effect for wide-band systems with single-port antennas, which is caused by the group delay characteristic. This irreducible BER is thus analogous to that in narrow-band systems caused by the random FM. The spikes of high dispersion in Fig. 1

correspond to the deep fades of the Rayleigh-like envelope. By avoiding the deep fades, depicted here in the frequency domain, the highly dispersive channels (where low channel bandwidths occur) are also avoided.

A note on terminology is in order, since the multidisciplinary nature of mobile communcations results inevitably in inconsistent nomenclature. Most terms used here follow from original articles or by convention according to the pertinent discipline. An example is the use of Γ for both polarization matrices (e.g., (4)) and signal-to-noise-ratio (e.g., (17)). Some inconsistencies also arise from historical "misuse." For example, covariances and (complex) correlations are considered the same, despite their mathematical distinction, and the terms carrier-to-noise ratio (CNR) and signal-to-noise ratio are also interchangeable, although this is not generally true. Strictly speaking, the CNR is the quantity of interest since the signals under consideration are RF (or IF) carriers, yet to be demodulated (predetection combining is assumed). SNR should only be applied to a signal after detection and will not, in general, be the same as the CNR. From here on, however, the term SNR is used, following Jakes' principal convention [13]. The time average is denoted $\langle \cdot \rangle$ and is interchangeable with the ensemble expectation since all processes are assumed ergodic. For matrix operations, the following superscripts apply: T means transpose, the asterisk means complex conjugate, and H means Hermitian transpose. When discussing the mobile communications scenario (see Section II), the word *source* refers to each point in space that can be considered to supply energy to the mobile antenna. The word *signal* refers to the intelligence conveyed by the energy from the sources. (Many sources convey the same signal.) When discussing antenna diversity, the *diversity gain* differs from the *diversity return* in that the latter includes the effects of mutual coupling. Strictly speaking, the diversity gain should include mutual coupling effects, but traditionally, this has not been the case. In referring to mobile antennas, the term *high gain* is used for antennas whose receiving patterns are confined (or almost so) to the directions of the sources.

Section II covers some basic aspects of antenna diversity and gives a fleeting mention of other methods for improving the mobile channel. Stein [28] and Jakes [13] discuss diversity in great depth, and the basics are indeed well covered. Some aspects are clarified in Section II. Not a great deal has been reported about the scenario of sources incident on an urban-based mobile or base station. For diversity antenna pattern considerations, a convenient distributed souce model is used to describe the (ensemble) average scenario, despite the fact that the instantaneous scenario may contain only a few sources. Energy considerations demonstrate the potential of multiple port antennas without resorting to space diversity. A figure of merit for a diversity antenna, the diversity gain, and its behavior in the presence of mutual coupling receives attention. It is shown that when correlated branches undergo nonswitched combining (or when the diversity antenna elements are always terminated), more care than that displayed in the literature is required to interpret the diversity gain. A fundamental difference exists between high-gain antennas at the mobile and base station antennas in this regard. A short discussion on the effect of different levels of branch mean SNR's concludes the section.

Section III presents several new ideas and viewpoints regarding antenna diversity. The conditions for diversity action are investigated. It is shown that under certain idealized conditions, the correlation coefficient between branch signals of a diversity antenna for the mobile can be equated with the mutual resistance between the antenna elements. This result is new, fundamental, and useful. It means that the performance of a class of diversity antenna designs for urban applications can be ascertained in the laboratory. The alternative is to measure correlations between branch signals in the field, normally an expensive and time-consuming exercise. The textbooks (see Stein [28], Jakes [13], Lee [22]) divide antenna diversity techniques into classes such as angle, polarization, space, field component, etc. These techniques are unified into pattern diversity. The condition for diversity action is found to be orthogonal element patterns over the sources. This is also a new and rather fundamental result. The formulation is given, and the situations at both the base station and the mobile are discussed.

Section IV (and the remainder of the article) concentrates on antenna diversity at the mobile. An element figure of merit (the element directivity toward the distributed sources scenario) is used to find useful design information. An array figure of merit (the diversity return) can also be applied to find useful and optimum diversity antenna configurations. The role of mutual coupling is investigated in detail, and ideas are fixed by considering rotationally symmetric two- and three-element array designs.

Section V discusses specific examples of diversity antennas for the mobile in terms of the pattern orthogonality. Both existing and new designs are included. It is noted that space diversity from concentric horizontal ring elements will not work well at the mobile. A circular array of three outward sloping monopoles is also discussed. The advantage is that the feedpoint spacings can be arbitrarily close. A sinusoidal current distribution is assumed for all configurations. As the antennas become closely spaced, a moment method solution would be better. However, is seems unnecessary to solve the problem exactly since both the infinite ground plane and source distribution are only approximations. Experimental values of the envelope correlation are in excellent agreement with the theory for a three-element example. The two-element case is mentioned and some remarks are offered for the many-branched circular array. Section VI concludes the paper, and the Appendix details the cumulative probability distribution of the combined signal from a circularly symmetric three-element array.

II. Antenna Diversity : Some Basic Aspects

Source Scenario at the Base Station [30]

Models are required for the scenario of sources producing the fields at the mobile and base station. At the base station, the incident fields due to a single mobile in an urban area occupy a very small portion of the base station field-of-

coverage. In fact, the incident signal is often well represented by a single direction when the antenna is clear of obstructions. Define the directions and extent of the sources (from a single mobile) by upper and lower limits θ_u, ϕ_u; θ_L, ϕ_L where the origin is the base station. The incident electric field is denoted

$$\boldsymbol{h}(\theta, \phi, t) = h_\theta(\theta, \phi, t)\hat{\theta} + h_\phi(\theta, \phi, t)\hat{\phi} \tag{1}$$

where the units of h, h_θ, and h_ϕ are volts/meter/steradian. The polarization matrix for the incident fields is defined as

$$\Gamma'(\theta_1, \phi_1; \theta_2, \phi_2) = \begin{bmatrix} \Gamma'_{\theta\theta} & \Gamma'_{\theta\phi} \\ \Gamma'_{\phi\theta} & \Gamma'_{\phi\phi} \end{bmatrix} \tag{2}$$

where the elements are of the form

$$\Gamma'_{\theta\phi}(\theta_1, \phi_1; \theta_2, \phi_2) = \langle h_\theta(\theta_1, \phi_1, t) h^*_\phi(\theta_2, \phi_2, t) \rangle. \tag{3}$$

If the polarizations are considered uncorrelated and each polarization considered spatially uncorrelated, then

$$\Gamma'(\theta_1, \phi_1; \theta_2, \phi_2) = P(\theta, \phi)\delta(\theta_1 - \theta_2)\delta(\phi_1 - \phi_2) \cdot \begin{bmatrix} \mathrm{XPD}(D) & 0 \\ 0 & 1 \end{bmatrix} \tag{4}$$

where

$$P(\theta, \phi) = P, \quad \theta_L \leqslant \theta_1 \leqslant \theta_u, \quad \phi_L \leqslant \phi \leqslant \phi_u$$
$$= 0, \quad \text{elsewhere} \tag{5}$$

is the (constant) power density per steradic square distribution and

$$\mathrm{XPD} = \frac{\Gamma'_{\theta\theta}}{\Gamma'_{\phi\phi}} \tag{6}$$

is the cross polar discrimination (XPD). For vertically polarized antennas in urban areas, the XPD is given by Kozono *et al.* [17] as a weak empirical function of the distance D between the mobile and base station. However, it is also a function of the polarization of the mobile antenna and the type of terrain along the path. For a vertically polarized base station and a vertically polarized urban based mobile antenna, XPD $\approx$ 6 dB (Lee and Yeh [21]). For a horizontally polarized base station, the value is ≈ -6 dB [21]. Most existing mobile antennas are principally vertically polarized. At the base station, then, we choose an average value XPD = 6 dB, but note that "instantaneous" values between -6 dB and 18 dB can occur (Kozono *et al.* [17]).

Source Scenario at the Urban Based Mobile [30]

At the mobile, the model is that the distributed sources occupy the far field evenly in the directions $0° \leqslant \phi < 360°$, $60° \leqslant \theta \leqslant 90°$, where θ and ϕ are now the spherical coordinates with the mobile at the origin. Both polarizations are uncorrelated and equally likely, the latter property implying that the base station receives equal powers in both polarizations. Each polarization is assumed spatially uncorrelated. The polarization matrix at the mobile is thus

$$\Gamma'(\theta_1, \phi_1; \theta_2, \phi_2) = S(\theta, \phi)\delta(\theta_1 - \theta_2)\delta(\phi_1 - \phi_2) \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7}$$

where

$$S(\theta, \phi) = \begin{cases} S, & 60° \leqslant \theta \leqslant 90°,\ 0° < \phi < 360° \\ 0, & \text{elsewhere} \end{cases} \tag{8}$$

is the constant power density per steradic square distribution around the mobile. It is emphasized that the XPD at the mobile has been assumed to be unity, a case corresponding to equal powers in the vertical and horizontal polarizations at the base station. This scenario is referred to as the mobile communications scenario (MCS).

Energy Considerations at the Mobile and Base Station

The energy density at a point (or in a small volume, strictly speaking) in space is proportional to

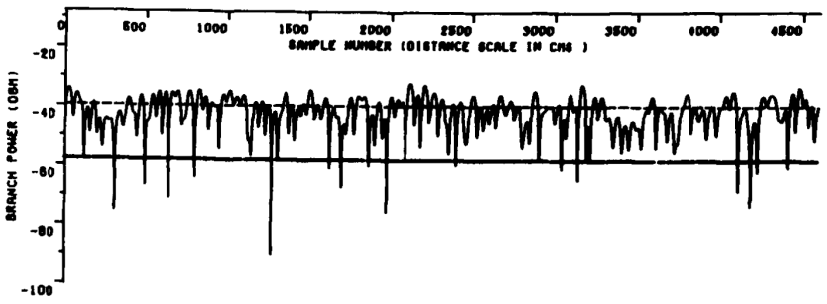
$$\text{energy} = |\boldsymbol{E}|^2 + |Z_0 \boldsymbol{H}|^2 \tag{9}$$

which is a six-component sum in the MCS (no earth plane is assumed present). The envelopes of the $|E_z|^2$ component and the total energy are plotted as a function of position in Fig. 2 along with their Rayleigh curves. Very little fading of the total energy occurs, and in principle, if an antenna could be designed to gather the energy coherently, there would be no need to resort to space diversity. Obviously, this antenna cannot have just a single port (a combiner is required as in space diversity). The presence of an earth plane close to the antenna reduces the number of field components to three. Pierce's energy density antenna (Gilbert [9]) was designed to receive these three components, and the technique is often called field component diversity. The antenna is mentioned in Section V. The reason it works well is that the three field components are uncorrelated at a point in an omnidirectional scenario (see Jakes [13, p. 38]).

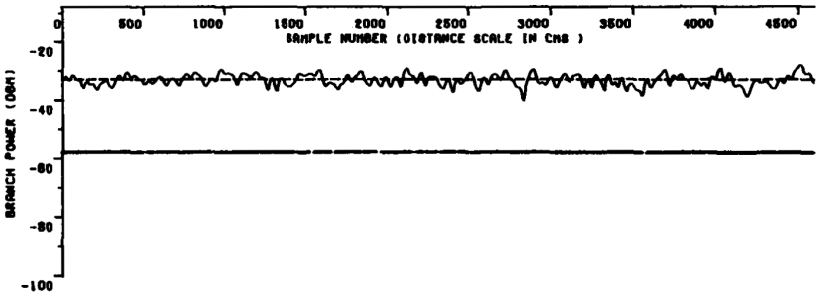
One interpretation of Fig. 2 is that the Rayleigh-like fading of the mobile channel is a result of using a single port antenna. At (or rather above) the mobile, the total energy is relatively constant so that compact diversity antennas are possible, at least in principle.

At the base station, it is not unreasonable to assume that the incident signal from a single mobile is from a single direction. This means that the incident energy is restricted to the two orthogonal polarizations in this direction. The maximum theoretical performance without resorting to space diversity (as far as the fading is concered) can thus be realized by polarization diversity (Vaughan and Bach Andersen [31]).

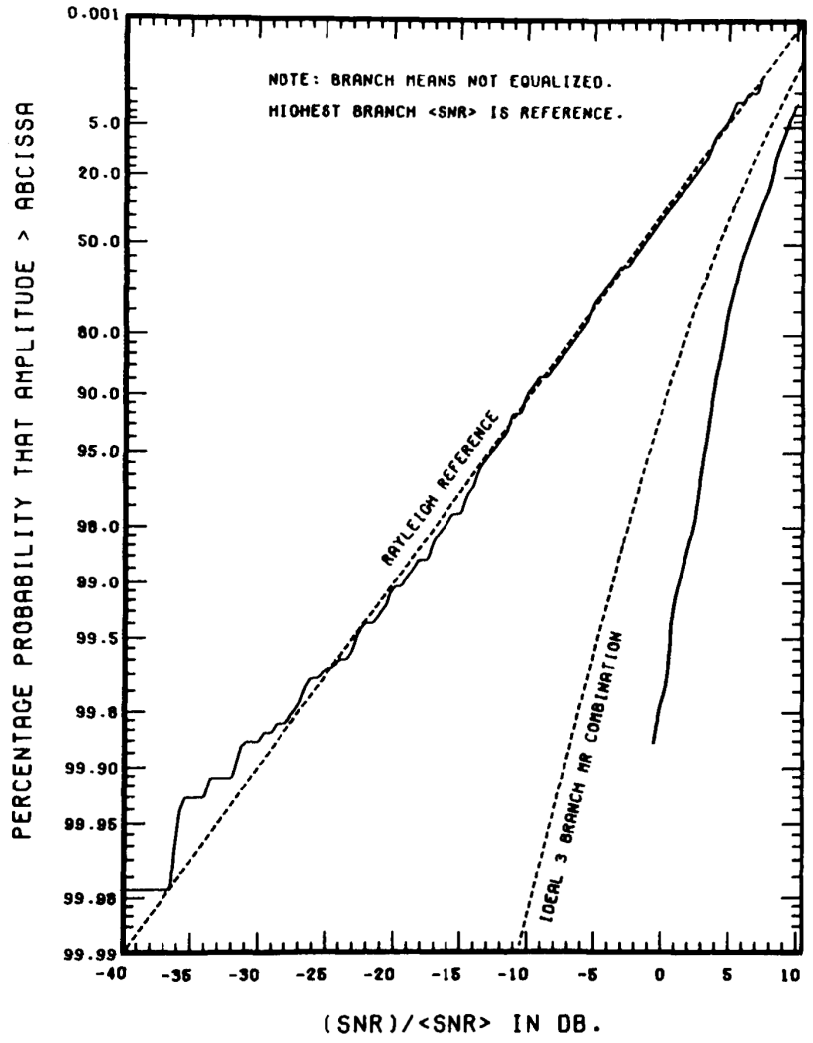
There is an important difference between the fading of energy at the mobile and at the base station. The energy at a point above a mobile in the MCS corresponds closely to a maximum ratio combination of five uncorrelated branches of equal mean SNR's (cf. Fig. 2(c)). The energy at a point at the base station has a theoretical limit of only two combined branches.



(a)



(b)



(c)

Fig. 2. (a) Energy in $\hat{z}$ component of electric field along path within MCS. (b) Total energy at point moving within MCS. No ground plane is assumed present. See [30] for the dB definition of standard deviation used here. (c) Rayleigh curves for fading of (a) and (b).

Signal Combination

In this section, predetection maximum ratio combining is of principal interest. There is little difference in diversity gains between equal gain, selection, and maximum ratio combinations. The relative performance returns for each of these schemes are well known (e.g., Jakes, [13, ch. 5]).

Switched diversity offers economical and practical schemes which are usually the type implemented. The local mean level of the signal can be measured so the threshhold can be floating, but relative to the local signal mean. Arnold and Bodtmann [2] give an example with wide-band simulation results of this technique. An interesting result is that the performance is rather insensitive to the threshhold value, over a range of several (~ 5) dB [2, fig. 7, p. 159]. Their simulation used four uncorrelated signals, and the switching rule was just sequential commutation, which surprisingly gives significantly better results than the three-branch selection case.

While switched schemes offer practical advantages, the maximum ratio combining is mainly of theoretical use and as a performance benchmark. More recently, the more complicated optimum combining (Bogachev and Kiselev [6], Winters [34]) has been discussed, although implementation details are lacking. The advantage of optimum combining is the possibility of improving strong interference suppression (over other combining schemes), an issue which will also become of increasing importance as the demands on spectrum efficiency in cellular systems increase. The degree of interference suppression is related to the number of branches, so optimum combining motivates many-branch systems. For interferers of similar or less power than the wanted signal, conventional combining gives quite good interference suppression. Miki and Hata [21] give some examples for two-branch switched combining which include the amount of interference suppression.

In maximum ratio combining (Kahn [15]), the weights are proportional to the conjugate of the signal voltage and the inverse of the branch noise power. Implementation of a maximum ratio combiner is expensive since the weights have both amplitude and phase, and measurement of the channel (instantaneous) SNR is required for each weight update. The technique is the best linear combination in the sense that it yields the largest output SNR, which turns out to be the sum of the branch SNR's. The latter property makes maximum ratio combining very attractive for finding theoretical characteristics of the combined signal.

If uncorrelated Rayleigh distributions and identical mean SNR's are assumed for each input channel, then the cumulative probability of the SNR of the maximum ratio combined signal is (e.g., Jakes [13, p. 319])

$$P_M(\gamma) = 1 - e^{-\gamma/\Gamma} \sum_{k=1}^{M} \frac{\left(\frac{\gamma}{\Gamma}\right)^{k-1}}{(k-1)!} \tag{10}$$

where M is the number of input channels and Γ is the mean SNR of each channel. Setting the number of branches M to 1 in (10) leads to the Rayleigh distribution.

The diversity gain is defined as the decrease in SNR compared to a nondiversity receiver for a given performance factor. The performance factor usually used with antenna diversity is related to $P_M(\gamma)$. For example, two-branch antenna diversity with maximum ratio combining gives a diversity gain of about 16 dB for $P_2(\gamma) = 0.001$. After three-branch diversity, diminishing returns from adding extra branches sets in for this measure of diversity gain.

Rather lax application of the term diversity gain has led to some misconceptions regarding actual diversity returns. Specifically, when branches become correlated, it is incorrect to read the diversity gain off a Rayleigh diagram without taking proper account of the mutual coupling. Before elaborating on this point, some discussion is in order regarding the correlation coefficient.

Correlated Branch Signals

The correlation coefficient ρ of two narrow-band signals whose envelopes follow a Rayleigh distribution is known (Pierce and Stein [27]) to obey

$$|\rho|^2 \approx \rho_e \tag{11}$$

where ρ_e is the correlation coefficient of the envelopes. It follows that the square root of the envelope correlation gives the signal correlation to within an arbitrary angle. This angle is usually considered as zero for practical purposes, and the absolute value sign in (11) is correspondingly dropped.

The property that the correlation coefficient is never negative for Rayleigh distributed signals is interesting. Measurements by the authors of envelope correlations obtained in urban environments have often been negative. Kozono *et al.* [17] also report negative correlation coefficients from their base station measurements. This is one way to demonstrate that the signal envelope of the mobile channel does not have a truly Rayleigh distribution. For diversity considerations, signals with a negative envelope correlation coefficient can offer better diversity gain than signals with zero correlation, such as those indicated in Fig. 2. Consider a two-source model in which the sources are directly in front of and behind the mobile. If two space diversity antennas were mounted such that the envelopes were

$$r_1 = |\sin x| \tag{12}$$

and

$$r_2 = |\cos x| \tag{13}$$

then the envelope correlation is readily established to be -0.92. In this case, two-channel diversity is sufficient to eliminate the fading almost completely. The reason is that the correlation coefficient is nearly -1, which represents the ideal value. For the scenario which gives rise to Rayleigh fading, the best value for envelope correlations between diversity antenna element signals, as far as curing the fading is concerned, is zero.

When the branch signals become correlated, it becomes very difficult to find $P_M(\gamma)$ for combinations other than maximum ratio. $P_2(\gamma)$ for a finite branch correlation is well

known and $P_3(\gamma)$ for a circular array (identical correlations for all branches in the MCS or any rotationally symmetric source scenario) is established in the Appendix. The Rayleigh curves for $P_2(\gamma)$ and $P_3(\gamma)$ are displayed in Figs. 3(a), 3(b). The curves for $P_2(\gamma)$ are well known (e.g., Jakes [13, p. 327]). Note the SNR is that of the combined signal and the reference $\langle \text{SNR} \rangle$ is that of a single branch. It is common practice to read the diversity gain off these curves for a given correlation coefficient. This is correct only if the mutual impedance has no effect. At the base station, this is not completely unreasonable because the mutual impedance decreases much more rapidly than the signal correlation as similar antennas are spaced apart. Space diversity, for example, at the base station requires distances of tens of wavelengths between elements (e.g., Lee [22, p. 201]), which for conventional antennas means that the mutual coupling is very low. Stated in another way, the correlation coefficient between base station elements can be very close to unity while the mutual coupling is negligible.

At the mobile, this cannot be the case. Consider again space diversity, but now at the mobile. The spatial correlation coefficient in the MCS which lies between $J_0(kx)$ and $\text{sinc}(kx)$ (Vaughan [30]) shows that for finite correlations (appreciable values, greater than, say 0.5), the antennas must be closer than a fraction of a wavelength. (In space diversity at the mobile, there is seldom interest in having a larger spacing than the first zero of the correlation function.) Now, in the limit as $\rho \rightarrow 1$, the spacing approaches zero and the elements merge into one. Nevertheless, the curves of Fig. 3(a), (b) indicate a 3-dB and 4.77 dB (power factors of 2 and 3, respectively) diversity gain for this case! Evidently, the diversity gain has to be defined in these cases as having a reference $\langle \text{SNR} \rangle$ from a single element *in the presence of the other elements of the diversity antenna while it is operating as a diversity antenna.* This definition can only be properly corrected by accounting for the mutual coupling. In Section III, it is shown that, for certain high-gain mobile antenna elements, the open circuit signal correlation coefficient ρ_0 is closely related to the normalized mutual resistance r,

$$\rho_0 \approx r. \tag{14}$$

For many antennas, the open circuit and terminated circuit correlation coefficients are reasonably close (cf. for example, Figs. 12 and 13 for sloping monopoles discussed below) and so to a reasonable approximation,

$$r^2 \approx \rho_e. \tag{15}$$

With these results, the approximate effect of mutual coupling can be included in the Rayleigh diagrams. The abscissa is modified by the multiplicative factor (additive, for dB quantities)

$$F = \frac{\langle \text{SNR(1 branch, mutual coupling ignored)} \rangle}{\langle \text{SNR(1 branch, mutual coupling accounted for)} \rangle}, \tag{16}$$

which is investigated in Section V. The form of the curves will be the same as those of Fig. 3, but they become shifted along the abscissa. For small and medium values of envelope correlation, the shift is quite small. For very large values of correlatin coefficient, the shift is large. For example, in the two-branch case, the curves for $\rho_e = 0$, 0.1, 0.5, 0.9, and 1 are shifted to the left by about 0, 0.2, 0.9, 2.4, and 3 dB, respectively. For the three-branch case, the corresponding shifts are about 0.2, 0.9, 1.6, and 4.77 dB, respectively. These shifted curves would then have the effect of mutual coupling fully included, albeit approximately, and can be used to read off the true diversity gain (now identical to the diversity return).

An explicit relation between ρ_L (the loaded circuit correlation), ρ_0 and r is available in Section IV, so that for a given antenna, the curves can be derived exactly. The above approximations are good for high-gain antennas at the mobile and the curves will not change much for all such antennas. Note also that the factor of (16) does not affect switched antenna diversity systems, where mutual coupling does not play an important role for this defintion of diversity gain (the unused elements are assumed to be open circuit and to obey the approximation of (14)).

The diversity gain available from Fig. 3 is not particularly sensitive to the envelope correlation coefficient ρ_e, as long as ρ_e is less than about 0.7. Indeed, $\rho_e \approx 0.7$ is quoted almost universally to be acceptable for diverstiy considerations. For maximum ratio combining at the mobile, this figure corresponds in a given diversity gain sense, to about 0.5 when the mutual coupling is accounted for. A condition for good diversity action using maximum ratio combining is that the correlation coefficient should be "low," which can be taken as $\rho_e < \sim 0.7$ at the base station or $\rho_e < \sim 0.5$ at the mobile.

Mean SNR Differences

It has been assumed that all branches have the same mean SNR's. When these become different, a combiner will, of course, favor the branch with the highest mean SNR, and the diversity returns will be reduced. In terms of the diversity gain, the degradation is similar to that caused by finite correlations. In the case of two branches, it is clear that the condition of one branch having much higher mean SNR than the other will result in the combined signal having the fading characteristics of a single channel independent of the branch signal correlation. This same effect occurs for correlated branchs ($\rho \rightarrow 1$), where the combined signal fades as a single channel, independent of the difference in the branch mean SNR's. The trade-off indicates that the branch mean SNR's should be "similar" for diversity action. Stein [28, p. 438] notes that for selection and maximum ratio combining of uncorrelated signals with unequal mean SNRs, the geometric mean of the branch mean SNR's gives an effective common branch SNR (here, Γ denotes mean SNR):

$$\Gamma_{\text{eff}} = \left[\prod_{k=1}^{N} \Gamma_k \right]^{1/N}, \tag{17}$$

which can be interpreted as the principal parameter for diversity performance. Stein [28, pp. 474, 480], also claims that for both selection and maximum ratio combining of two

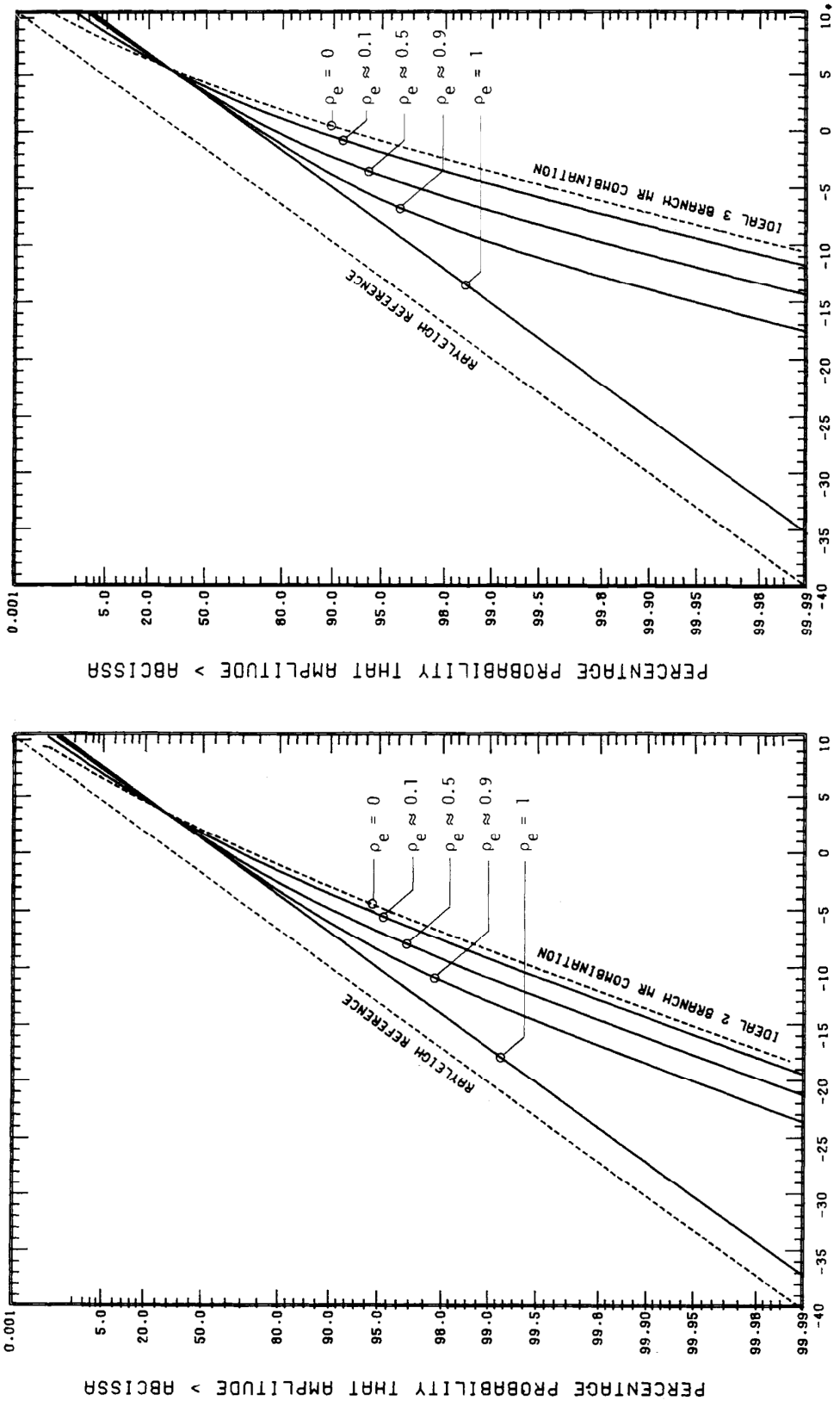


Fig. 3. Rayleigh curves for (a) two- and (b) three-branch maximum ratio combination with envelope correlations ρ_e. All branch SNR's are assumed to be Rayleigh distributed with identical ⟨SNR⟩'s, and in (b) identical correlations between each branch. The effect of mutual coupling is ignored. The curves of (a) are well known (e.g., Jakes [13, p. 327]).

branches having a finite envelope correlation coefficient ρ_e (where ρ_e is not too close to unity), the effective uncorrelated common branch SNR is further reduced by a factor of $\sqrt{1-\rho_e}$.

For two-branch maximum ratio combining, the requirement of "similar" SNR's (uncorrelated branches) turns out to be a difference of less then 10 dB, which yields a diversity gain at $P_2(\gamma) \approx 0.001$ similar to two branches of equal mean SNR and a correlation of $\rho_e \leqslant 0.7$.

III. Antenna Diversity: A New Approach

Conditions for Diversity

The conditions for achieving good diversity gain are already mentioned in Section II, the correlation coefficient between branch signals should be zero, or "low," and mean SNR's in each branch should be "similar." It is of interest in diversity antenna design to see how these signal requirements are related to, or indeed if they can be expressed in terms of, the antenna parameters.

It is assumed that the antenna noise is negligible compared to the receiver noise so that the SNR depends only on the gain of the antenna element. The gain is not considered towards a single direction as in the conventional definition, but is rather considered toward an area, viz., the MCS. An antenna element figure of merit is established along these lines in Section IV. If the element pattern is confined to the MCS, the gain of a lossless antenna becomes independent of the pattern shape, since the source distribution is assumed uniform. Furthermore, it is assumed that similar elements of an array antenna will provide similar mean SNR's. The discussion is from now on limited to the correlation coefficient.

The following treatment uses a similar notation to Collin and Zucker [8, ch. 4] except that the $\boldsymbol{E}$ and $\boldsymbol{h}$ vectors are interchanged. Here $\boldsymbol{h}$ is the source vector (see (1)) and $\boldsymbol{E}$ is the antenna pattern vector. The open circuit voltage (here, for the kth element) is given by (Collin and Zucker [8, p. 115])

$$V_{0k}(t) = \int\int \boldsymbol{E}_k(\Omega) \cdot \boldsymbol{h}(\Omega, t)\, d\Omega \tag{18}$$

where Ω is the solid angle (θ, ϕ) and the appropriate time dependence has been introduced. The kth antenna element has been assumed to be at the origin. A two-element antenna polarization matrix can be defined in an analogous way to the source polarization matrix. For the jth and kth elements,

$$\boldsymbol{\Gamma}_{jk}(\Omega_1, \Omega_2) = \begin{bmatrix} \Gamma_{\theta\theta jk} & \Gamma_{\theta\phi jk} \\ \Gamma_{\phi\theta jk} & \Gamma_{\phi\phi jk} \end{bmatrix} \tag{19}$$

in which a matrix element is defined

$$\Gamma_{\theta\phi jk}(\Omega_1, \Omega_2) = E_{\theta j}(\Omega_1) E^*_{\phi k}(\Omega_2) \tag{20}$$

where

$$\boldsymbol{E}_k = E_{\theta k}\hat{\theta} + E_{\phi k}\hat{\phi}. \tag{21}$$

Note that the polarization matrix is different from the usual form (e.g., Collin and Zucker [8]) for single elements. $\boldsymbol{\Gamma}$ is known from the element far fields. The units of $\boldsymbol{E}$ are V/m, in contrast to those of the source vector $\boldsymbol{h}$.

The correlation coefficient for the open circuit jth and kth signals is

$$\rho_{0jk} = \frac{E\{(V_{0j} - \bar{V}_{0j})(V_{0k} - \bar{V}_{0k})^*\}}{[E\{(V_{0j} - \bar{V}_{0j})^2\} E\{(V_{0k} - \bar{V}_{0k})^{2*}\}]^{1/2}} \tag{22}$$

where the expectation is taken over the same interval as in the biased time average, which is denoted by the swung dash. For conciseness, the demeaning and normalization processes are from now on understood, so (22) is written

$$\rho_{0jk} = E\{V_{0j} V^*_{0k}\}. \tag{23}$$

Using (18) and (23),

$$\rho_{0jk} = E\left\{ \int \boldsymbol{E}_j(\Omega_1) \cdot \boldsymbol{h}(\Omega_1, t)\, d\Omega_1 \int \boldsymbol{E}^*_k(\Omega_2) \cdot \boldsymbol{h}^*(\Omega_2, t)\, d\Omega_2 \right\} \tag{24}$$

and interchanging the order of expectation and integration, as well as dropping the element dependences of the polarization matrix elements,

$$\begin{aligned} \rho_{0jk} &= \int\int (\Gamma_{\theta\theta}\Gamma'_{\theta\theta} + \Gamma_{\phi\phi}\Gamma'_{\phi\phi} + \Gamma_{\theta\phi}\Gamma'_{\theta\phi} + \Gamma_{\phi\theta}\Gamma'_{\phi\theta})\, d\Omega_1\, d\Omega_2 \\ &= \int\int \operatorname{tr}(\boldsymbol{\Gamma}\boldsymbol{\Gamma}')\, d\Omega_1\, d\Omega_2, \end{aligned} \tag{25}$$

where $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}'$ are defined in (19), (20) and (2), (3), respectively. The correlation coefficient is thus expressible explicitly in terms of the source and antenna element polarization matrices. This result is general.

In the presence of Rayleigh fading, the lowest correlation coefficient is zero (as noted in Section II, a correlation coefficient of -1 is the ideal general value). Thus a condition (uncorrelated signals) for ideal diversity action to combat Rayleigh fading, is that the source and antenna polarization matrices be orthogonal over the sources in the sense of the inner product defined by (25), i.e.,

$$\int\int \operatorname{tr}(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}')\, d\Omega_1\, d\Omega_2 = 0. \tag{26}$$

Signals at the Mobile

In the MCS, the expression for the open circuit correlation coefficient (25) simplifies in a particularly interesting way. Applying the assumption that the orthogonal polarizations are uncorrelated, then

$$\rho_{0jk} = \int\int (\Gamma_{\theta\theta}\Gamma'_{\theta\theta} + \Gamma_{\phi\phi}\Gamma'_{\phi\phi})\, d\Omega_1\, d\Omega_2 \tag{27}$$

and that each polarization is uncorrelated in space, there results

$$\begin{aligned} \rho_{0jk} &= \int\int (E_{\theta j}(\Omega_1) E^*_{\theta k}(\Omega_2) + E_{\phi j}(\Omega_1) E^*_{\phi k}(\Omega_2)) \\ &\quad \cdot \delta(\Omega_1) S(\Omega_1 - \Omega_2)\, d\Omega_1\, d\Omega_2 \qquad (28) \\ &= \int\int \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}^*_k(\Omega) S(\Omega)\, d\Omega, \qquad (29) \end{aligned}$$

and finally that the source power density distributions are constant, then

$$\rho_{0jk} = \int_{\text{MCS}} \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}^*_k(\Omega)\, d\Omega. \tag{30}$$

The normalization process accounts for the cancellation of constants. For the inner products defined in (29) and (30), it is emphasized that the integration is over the sources. The integration could be extended over all space since the inner product weighting ($S(\Omega)$) will be zero in this region.

The contributions of each polarization to the correlation coefficient can be separated (the open circuit notation is dropped),

$$\rho_{jk} = \rho_{\theta jk} + \rho_{\phi jk} \tag{31}$$

where

$$\rho_{\theta jk} = \int_{\text{MCS}} E_{\theta j}(\Omega) E^*_{\theta k}(\Omega)\, d\Omega \tag{32}$$

and similarly for $\rho_{\phi jk}$.

A derivation for conventional space diversity (Clarke's model scenario [7]) makes a good example. Two monopoles are spaced distance d apart on an infinite groundplane (the groundplane could be removed and dipoles used). Equation (29) for the correlation becomes

$$\rho_{0jk} = \int\int P_\theta(\theta) e^{jk\cdot d} S(\theta, \phi) \sin\theta\, d\theta\, d\phi \tag{33}$$

where $P_0(\theta)$ is the power pattern of one of the elements. Insertion of Clarke's model

$$S(\theta) = S\delta(\theta) \tag{34}$$

collapses the integral in (33) to

$$\rho_{0jk} = \int_0^{2\pi} e^{jkd\cos\theta}\, d\phi$$

$$= J_0(kd). \tag{35}$$

It is interesting that ρ_{0jk} is real in this example. Any identical patterns which are circularly symmetric will give this result. In fact, it is apparent that whenever the element patterns are identical, the imaginary part of the coordinate translation term is zero, leaving a real correlation coefficient.

A conclusion for mobile antennas results from (30). For zero correlation between diversity branches at the mobile, the antenna element patterns should be orthogonal over the sources (here, the MCS), i.e.,

$$\int_{\text{MCS}} \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}^*_k(\Omega)\, d\Omega = 0. \tag{36}$$

This is a requirement for an ideal diversity antenna at the mobile.

The Ideal Diversity Antenna for the Mobile

For a maximum gain, the far-field pattern of the diversity antenna is confined to the source region, the MCS. Under this condition, the open circuit correlation coefficient (or, as shown for this case in Section IV, the terminated circuit correlation coefficient) is the same as the normalized mutual resistance. The ideal antenna has zero mutual resistance between elements. Each element should provide the same mean SNR. Ideally, the element patterns should span the space of the sources, although this is not a necessary condition for the case of the MCS because of the assumption that the sources all convey the same information.

For the mobile antenna to work outside of the urban environment, the composite pattern should be omnidirectional. Here, the omnidirectionality can be restricted to the vertically polarized component since the vertically polarized waves will be normally dominant close to ground level (assuming a vertically polarized base station). This is a much more practical restraint than having the element patterns spanning the space of the MCS. The presence of a ground plane in the MCS (and on the vehicle roof, for roof-mounted antennas) suppresses the horizontally polarized component of the wave. Other important practical properties are closely spaced (ideally adjacent) element feed points, compactness, ruggedness, and operation over the band of interest.

It would be unlikely that an ideal diversity antenna as already defined could be realized. However, the diversity returns are rather insensitive to quite large deviations from many of the ideal diversity antenna parameters, so the ideal parameter should be considered only as a direction in which to aim. For example, the envelope correlation coefficient

$$\rho_{ejk} \approx \rho_{jk}\rho^*_{jk} \approx (\text{Re}\ \{\rho_{jk}\})^2 \approx r^2_{jk} \tag{37}$$

where r_{jk} is the normalized mutual resistance between the jth and kth elements can be less than about 0.7 instead of being zero.

Mutual Impedances

Mutual impedances can be expressed in terms of the far-field patterns if the antenna elements are minimum scattering antennas. This idealized class of lossless antennas has the property that when terminated by an appropriate reactance, their scattered fields are identical to their radiated fields (Dicke, in Montgomery *et al.* [24]; Kahn and Kurss [16].

If the terminating reactance is zero for this property, the antennas are called canonical minimum scattering antennas. Some corollaries are as follows: the antenna is rendered invisible when it is terminated in the appropriate reactance; when a matched load is introduced, the absorbed and scattered powers are identical; and for reciprocal antennas, the patterns are symmetrical in any plane through the origin.

Wasylkiwskyj and Kahn [33, p. 212] give the mutual impedance between two minimum scattering antennas with normalized far fields (in the sense that total radiated power from each element is unity) $\boldsymbol{E}_j$ and $\boldsymbol{E}_k$, as

$$Z_{jk} = 2\int_{\phi=0}^{2\pi} \int_C \boldsymbol{E}_j(\theta, \phi) \cdot \boldsymbol{E}^*_k(\theta^*, \phi) e^{j\boldsymbol{k}\cdot\boldsymbol{d}} \sin\theta\, d\theta\, d\phi \tag{38}$$

where the path of integration c for θ is from $-\pi/2 - j\infty$ to $\pi/2 + j\infty$ in the complex θ plane. $\boldsymbol{k}$ is the incident propagation vector and $\boldsymbol{d}$ is the vector from antenna element k to antenna element j. The real part of the integration over real space (the far field) will be the sole real contribution to Z_{jk} in (38). This is not obvious from inspection of the integral but can be deduced by noting that, in a lossless environment, power

transfer can occur only in real space. (The antennas are already assumed lossless.) The mutual resistance can therefore be written as the real part of the real space integration of (37), i.e.,

$$R_{jk} = 2 \text{ Re} \left\{ \int_{\phi=0}^{2\pi} \int_{\theta=-\pi/2}^{\pi/2} \boldsymbol{E}_j(\theta, \phi) \cdot \boldsymbol{E}_k^*(\theta^*, \phi) e^{jk\cdot d} \sin\theta \, d\theta \, d\phi \right\}. \quad (39)$$

The mutual reactance arises primarily from integration over the invisible region, although some contributions may arise from real space for certain antenna patterns. In principle, the mutual reactance can be tuned out using a lossless network between antenna ports. In practice, it sometimes turns out that this is not necessary, since the mutual reactance is negligible. Most single-mode antennas are considered by Bach Andersen *et al.* [3] to satisfy (39) approximately. The antennas considered are here operated as dominantly single mode so that the relation (39) is assumed to be basically valid (but probably not exact) from here on.

The form for normalized resistance will be of particular interest. The coordinate translation term in (38) is assumed to be included in the far-field pattern of one of the antenna elements. The normalized mutual resistance becomes

$$r_{jk} = \text{Re} \left\{ \int \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}_k^*(\Omega) \, d\Omega \right\} \quad (40)$$

in which the constants have dropped out in the normalization process. The integral of (40) is for all real space. It is noteworthy that the condition for zero mutual resistance between antenna elements is that the diversity antenna element patterns should be orthogonal over real space, i.e.,

$$\text{Re} \left\{ \int \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}_k^*(\Omega) \, d\Omega \right\} = 0. \quad (41)$$

For identical antenna element patterns, the dot product in the integrand becomes the power pattern of one of the anntenna elements multiplied by a coordinate translation term. Furthermore, when this term is integrated in azimuth, the imaginary contribution is zero. The Re $\{\cdot\}$ thus becomes unnecessary in (40), i.e., for identical antenna element patterns with circular symmetry, the normalized mutual resistance is

$$r_{rk} = \int P(\theta) e^{jk\cdot d} \sin\theta \, d\phi \quad (42)$$

where $P(\theta)$ is the power pattern of a single element.

Relation Between Correlation and Mutual Resistance

The similarity between the real part of the normalized correlation coefficient

$$\text{Re} \{\rho_{0jk}\} = \text{Re} \left\{ \int_{\text{MCS}} \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}_k^*(\Omega) S(\Omega) \, d\Omega \right\}, \quad (43)$$

and the normalized mutual resistance of (40) is striking. The imaginary part of the correlation coefficient is usually not of interest. The Re $\{\cdot\}$ symbol therefore becomes unnecessary in (43) and, since the inner product weighting function $S(\Omega)$ is zero over those portions of real space where no sources exist, the (real) correlation coefficient becomes

$$\rho_{0jk} = \int \boldsymbol{E}_j(\Omega) \cdot \boldsymbol{E}_k^*(\Omega) S(\Omega) \, d\Omega \quad (44)$$

where the integration is now over all real space.

Equations (40) and (44) show that the correlation coefficient can be couched as the same inner product as the mutual resistance but with a different weighting function. However, a converse procedure is more revealing. If it can be assumed that the element radiation patterns are confined to the source region, then the normalized mutual resistance and real correlation coefficient generated in the presence of the MCS are identical. The assumption is not impractical. A good mobile antenna (high gain) will have the majority of its pattern confined to the MCS. For example, a quarter wavelength monopole on an infinite ground plane has 73 percent of its power pattern within the MCS. For an 0.6-wavelength vertical monpole, the figure is about 96 percent (see Fig. 4). For practical antennas operating in the MCS, it is reasonable that

$$r_{jk} \approx \rho_{0jk}. \quad (45)$$

This is an important result. The zero correlation condition can now be restated: for ideal diversity action, the mobile antenna elements should have zero mutual resistance.

Of course, any active array should have "low" mutual resistance. However, here a quantitative, albeit approximate, value is available for diversity antennas at the mobile. Furthermore, it allows antenna design and laboratory measurements (the mutual resistance) to establish essentially how well a mobile diversity antenna will perform; without having to test it in the field.

The ideal requirement of having zero correlation/mutual resistance is a rather insensitive and approximate one, as noted earlier. In practice, an optimum antenna may well have quite large correlations and mutual resistances, especially at the mobile, where antenna compactness is important.

Signals at the Base Station

The polarization matrix for signals at the base station is given in (4). Using this result in (25), the open circuit correlation for the base station antenna becomes

$$\rho_{0jk} = \iint (\Gamma_{\theta\theta}\Gamma'_{\theta\theta} + \Gamma_{\phi\phi}\Gamma'_{\phi\phi}) \, d\Omega_1 \, d\Omega_2 \quad (46)$$

$$= \iint P(\Omega_1)\delta(\Omega_1 - \Omega_2) \, (\text{XPD}(D)\Gamma_{\theta\theta}(\Omega_1, \Omega_2) + \Gamma_{\phi\phi}(\Omega_1, \Omega_2)) \, d\Omega_1 \, d\Omega_2 \quad (47)$$

$$= \int P(\Omega) \, (\text{XPD}(D)E_{\theta j}(\Omega)E_{\theta k}^*(\Omega) + E_{\phi j}(\Omega)E_{\phi k}^*(\Omega)) \, d\Omega \quad (48)$$

$$= \int_{\Omega_L}^{\Omega_u} (\text{XPD}(D)E_{\theta j}(\Omega)E_{\theta k}^*(\Omega)$$

$$+E_{\theta j}(\Omega)E^*_{\phi k}(\Omega)\,d\Omega \tag{49}$$

$$=\int_{\Omega_L}^{\Omega_u} E_j(\Omega)\cdot E^*_k(\Omega)\,d\Omega \tag{50}$$

only if the XPD is unity (i.e., 0 dB). The various constants drop out under normalization. Equation (49) is the general result under the assumptions about the sources made in Section II. These assumptions (orthogonal polarizations uncorrelated and each polarization spatially uncorrelated) must be kept in mind in interpreting (49) as the condition for diversity action between the jth and kth base station antennas. Consider, for example, polarization diversity. A crossed dipole configuration could be oriented so that

$$E_{\theta j}=0, \qquad E_{\phi k}=0; \tag{51}$$

then

$$E_{\theta j}E^*_{\theta k}=E_{\phi j}E^*_{\phi k}=0 \tag{52}$$

and ρ_{jk} is identically zero from the general result of (49), independent of the illumination! This is correct, but it must be remembered that the illumination is already assumed to be uncorrelated for the base station polarizations used for diversity. Base station antennas polarized at $\pm 45°$ from the vertical, illuminated by waves with dominantly vertical polarization will, of course, bear correlated signals because the components in the $\pm 45°$ planes are correlated.

Finally, it is noted that for elements of the base station array, the mutual resistance cannot be equated with the open circuit correlation coefficient as it could be for the mobile antennas. The reason is that the XPD is usually not unity and that the incident source region ($\Omega_u - \Omega_L$) is generally much smaller than the field of view of the base station antenna.

IV. Antenna Diversity at the Mobile

Element Figure of Merit

An element figure of merit can be defined by the proportion of its far-field (power) pattern illuminated by the source region, i.e., the MCS. Denoting the element figure of merit F_e and invoking reciprocity,

$$F_e=\frac{\text{power radiated into MCS by element}}{\text{total power radiated by element}} \tag{53}$$

$$=\frac{\int_{\text{MCS}}(\boldsymbol{E}\times\boldsymbol{H}^*)\cdot\hat{n}\,dS}{\int_{\text{real space}}(\boldsymbol{E}\times\boldsymbol{H}^*)\cdot\hat{n}\,dS} \tag{54}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic fields radiated by the element and $\hat{n}$ is the outward unit normal on the enclosing surface. For lossless antennas, F_e is the gain toward the MCS. A couple of examples illustrate the design information available from the element figure of merit. In Fig. 4, F_e for a sloping monopole above an infinite groundplane is plotted in dB, with the sloping (elevation) angle as argument and monopole length as parameter. The reference level is a vertical quarter wave monopole. A sinusoidal current distribution is assumed in all cases. It is seen that increasing the monopole length improves the figure of merit only for element elevation angles greater than about 57°. For the vertical element, a 0.6 λ monopole has a figure of merit more than one dB better than the quarter wave monpole. At very low elevation angles, where the antenna approaches the ground plane, F_e tends to a limit. F_e tends to its maximum when the monopole becomes vertical.

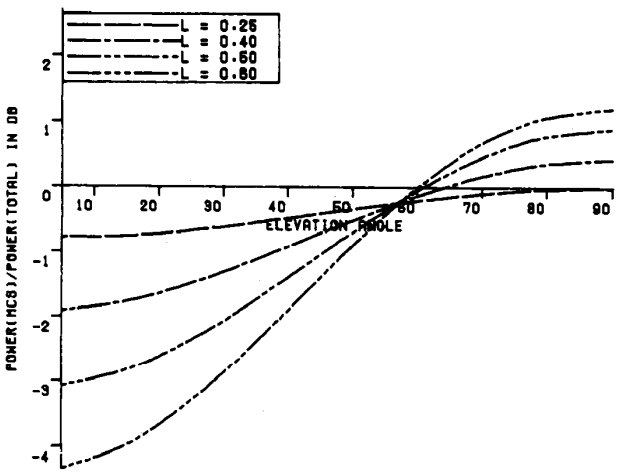


Fig. 4. Element figure of merit (power radiated into MCS/total power radiated) for sloping monopole with sinusoidal current distribution on infinite ground plane. Element length is L wavelengths and element elevation angle in degrees is argument. Reference level is vertical quarter wave monopole which corresponds to 73 percent. Vertical 0.6 λ monopole figure of merit is about 1.2 dB higher, corresponding to about 96 percent.

F_e for a single mode circular patch antenna on an infinite ground plane is plotted as a percentage, in Fig. 5 (Vaughan and Bach Andersen [31]). The patch radius is the argument and the parmeter n is the mode of the azimuthal form of the fields cos $n\phi$. Higher order modes are seen to provide higher gain antennas, with the $n = 0$ mode ("shortened monopole") falling between the $n = 2$ and $n = 3$ modes. An interesting feature is that as the radius decreases, and the antenna becomes more compact, F_e increases toward a limit. After the $n = 7$ mode (not shown), the figure limits (at 100 percent) and there is no advantage, from an element pattern viewpoint, of going to still higher modes. (In practice, however, a finite ground plane causes a "lifting" of the radiation pattern, and modes higher than $n = 7$ may be required before limiting of F_e occurs.)

Diversity Gain

The central parameter of merit for a diversity antenna is the diversity gain. The result sought in this section is the cumulative probability density (CPD) of the SNR of the combined signal. By assuming that each branch supplies a Rayleigh fading signal and maximum ratio combining is employed, considerable progress is possible in establishing the CPD and thence the diversity gain. Here, the key property of the Rayleigh distribution is that is can be expressed as a zero mean complex Gaussian variable. Likewise, the property that maximum ratio combining yields an SNR which is the sum of the branch SNR's means that the characteristic function of the combined signal SNR is just a product of the branch SNR

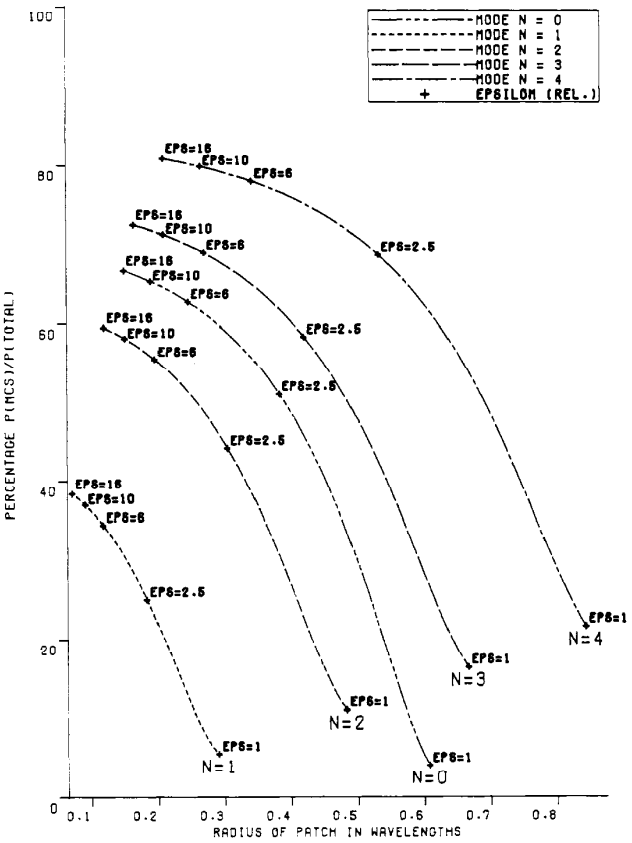


Fig. 5. Element figure of merit (power radiated into MCS/total power radiated) for resonant circular patch (single cos $n\phi$ mode ring source on infinite ground plane). Different curves are for different azimuthal modes, and labeled points are substrate relative permittivity required for single mode cavity model patch antenna [31].

characteristic functions. The approach of Pierce and Stein [27] is followed here.

Each antenna port is assumed to provide an SNR of w so that, for maximum ratio combining, the output SNR is

$$\gamma = \frac{1}{2} \boldsymbol{w}^H \boldsymbol{w} \tag{55}$$

where $\boldsymbol{w}$ is a column vector of the w_k.

The (Laplace) characteristic function for the probability density distribution of γ is

$$G(s) = \frac{1}{\prod_{m=1}^{M} (1 + js\lambda_m)} = \frac{1}{|I + sL|} \tag{56}$$

where the latter result is from Stein [28, p. 476]; after Kac and Siegert [14]. The probability density distribution is thus given by the Laplace transform of $G(s)$. The ensuing integration contour can be closed around the left-hand plane so that the residue theorem can be invoked. In general, the poles of $G(s)$ must be found numerically. However, there are forms of $\boldsymbol{L}$ with known eigenvalue solutions which correspond to realizable antenna array configurations. These include the tridiagonal form which corresponds to coupling between adjacent linear array elements only, and the circulant. The circulant is of particular interest here since the covariance matrix of a circular array gives rise to a special case of this form. For the circulant

$$\boldsymbol{C} = \begin{bmatrix} C_0 & C_1 & \cdots & C_{N-1} \\ C_{N-1} & C_0 & \cdots & C_{N-2} \\ \vdots & & & \\ C_1 & \cdots & & C_0 \end{bmatrix}, \tag{58}$$

Bellman [5, p. 242] gives the general eigenvalue solution

$$\lambda_k = C_0 + C_1 r_k + C_2 r_k^2 + \cdots + C_{N-1} r_k^{N-1} \tag{59}$$

where r_k is the kth root of unity from $r^N = 1$.

The eigenvectors are also given by Bellman as

$$U_k^T = (1\, r_k\, r_k^2 \cdots r_k^{N-1}). \tag{60}$$

It is evident that the eigenvectors are linearly independent and exist in complex conjugate pairs.

A covariance matrix resulting from a circular array will also be Hermitian, i.e., $c_n = c_{N-n}^*$ in (58). If the signals are Rayleigh distributed, the correlation matrix will be real, from the results of Section III. Since the eigenvalues must then be real, the eigenvalue equations can be written in the form

$$\boldsymbol{L}U_k = \lambda_k U_k \tag{61}$$

$$\boldsymbol{L}U_k^* = \lambda_k U_k^* \tag{62}$$

so that each eigenvalue is associated with two conjugate eigenvectors and the system is always degenerate for more than two branches.

The cumulative distribution for γ is obtained by integrating the probability density distribution in the normal manner. Lee [22, p. 308] supplies the result

$$P(\gamma > x) = \sum_{j=1}^{m} \frac{(\lambda_j)^{M-1} e^{-x/\lambda_j}}{\prod_{k \neq j}^{M} (\lambda_j - \lambda_k)} \tag{63}$$

which, however, is valid only for the nondegenerate case. The eigenvalues for (63) are actually for a product of the correlation matrix and an impedance matrix. The inclusion of the diagonal impedance matrix does not alter the degeneracy property, however. As an example, consider a set of four elements arranged in a square. This is a four-element circular array with a covariance matrix of the form (Rayleigh fading is assumed)

$$\boldsymbol{L} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{12} \\ \rho_{12} & 1 & \rho_{12} & \rho_{13} \\ \rho_{13} & \rho_{12} & 1 & \rho_{12} \\ \rho_{12} & \rho_{13} & \rho_{12} & 1 \end{bmatrix} \tag{64}$$

where ρ_{jk} is the branch signal correlation between the jth and kth elements. $\boldsymbol{L}$ is degenerate with the eigenvalues available

from (59) as

$$\lambda_1 = 1 + 2\rho_{12} + \rho_{13}$$
$$\lambda_2 = 1 - 2\rho_{12} + \rho_{13}$$
$$\lambda_3 = \lambda_4 = 1 - \rho_{13}. \quad (65)$$

For a two-element array, the cumulative probability can be found explicitly (see, e.g., Jakes [13, p. 325]). For more elements than two in a general configuration, no such formulas for the probability density distribution seem to be available. However, it is evident that formulas can be found for circular arrays. It follows that the diversity gain can be calculated for any circular array. A useful example is the three-element circular array. From the Appendix, the cumulative probability distribution is

$$P_3(\gamma) = 1 - \frac{1}{(3\rho)^2}\left\{(1+2\rho)^2 e^{-\gamma/\Gamma(1+2\rho)} - (3\rho\left(\frac{\gamma}{\Gamma} + 1 - \rho\right) + (1+2\rho)(1-\rho))e^{-\gamma/\Gamma(1-\rho)}\right\}. \quad (66)$$

The diversity gain for a three-element circular array (neglecting mutual coupling) is then available as

$$\left(\frac{\gamma_3}{\Gamma}\ (\text{in dB}) - \frac{\gamma_1}{\Gamma}\ (\text{in dB})\right)\Big|_{\text{probability given}} \quad (67)$$

where γ_m is the SNR of the m-branch combined signal for the given probability.

Mutual Coupling

Previous work on mutual coupling in antennas for the mobile seems to be confined to that of Lee [19], [20] which is summarized in Jakes [13, sec. 5.3] and some results reproduced in Lee [22, sec. 10.11].

Lee's treatment is restricted to linear and planar arrays of quarter wavelength monopoles. The original publications [19], [20] present many numerical results for the behavior of the second-order statistics and the expected reduction in average received power as the element spacing decreases, but do not cover the points to be discussed here.

The role of mutual coupling in a mobile diversity antenna is different to that of conventional array antenna. The statistical nature of the currents means that the effect of the mutual impedances becomes statistical. A simple physical situation lends insight: for a pair of minimum scattering antenna elements with even quite large correlation (say $\rho_{12} \approx R_{12}/R_{11} \approx 0.7$), the deep fades are rarely simultaneous. When one element is in a deep fade, its current is small and it behaves almost as if it were open circuit, and therefore almost invisible, as far as the other element is concerned. The effect of the relatively high mutual coupling is thus not strongly felt with regard to the second-order statistics. The decrease in average received power due to mutual coupling in noted by Lee [19] and his results [19, table I, p. 784] are similar to the appropriate special case configurations of the sloping monopole array which is mentioned below.

An equivalent circuit for an array antenna with mutual couplings and terminations is given in Fig. 6. A further appreciation of what is going on can be derived from the equivalent circuit for a single element in the same figure.

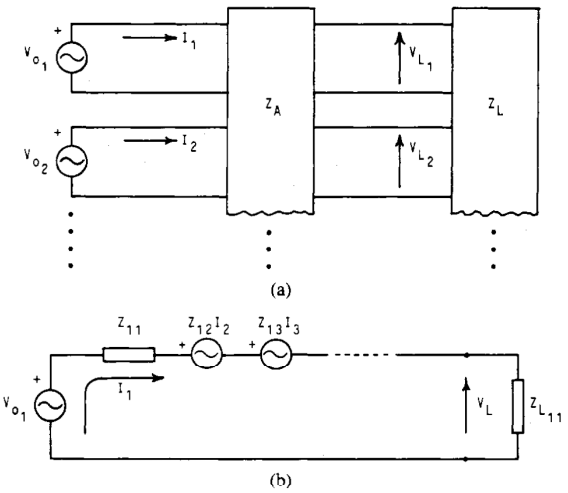


Fig. 6. (a) Equivalent circuit for terminated array. Z_A is antenna impedance matrix, Z_L is load matrix, and V_{oi} is open circuit voltage. (b) Equivalent circuit for single element (denoted 1) with diagonal load impedance matrix.

Denoting a vector of open circuit element voltages

$$V_0^T = (V_{01}, V_{02}, \cdots) \quad (68)$$

and similarly for the loaded circuit voltages V_L and currents I, the circuit relations are

$$V_0 = (Z_A + Z_L)I \quad (69)$$
$$V_L = Z_L I \quad (70)$$

where Z_A is the antenna impedance matrix and Z_L is the load impedance matrix. The correlation matrix of open circuit voltages is

$$L_0 = E\{V_0 V_0^H\} \quad (71)$$

and using the above circuit relations, the loaded circuit correlation matrix

$$L_L = E\{V_L V_L^H\} \quad (72)$$

can be expressed in terms of the open circuit correlations and impedances, viz.

$$L_L = E\{Z_L(Z_A+Z_L)^{-1}V_0(Z_L(Z_A+Z_L)^{-1}V_0)^H\}$$
$$= Z_L(Z_A+Z_L)^{-1}L_0(Z_A+Z_L)^{-1H}Z_L^H. \quad (73)$$

Because L_0 is related to the available power and L_L to the received power, L_L will in general be unnormalized for normalized L_0. The relationship between L_L and L_0 is a complicated function of the antenna impedances. The maximum power transfer theorem shows that for maximum received power,

$$Z_L = Z_A^* \quad (74)$$

(Jakes [13] after Haus and Adler [11]. Realizing Z_A^* can be

difficult, and in most practical situations the load network is resistive and Z_L becomes diagonal and real. Bach Andersen and Rasmussen [4] have shown that the reactive coupling can be removed (tuned out) and provide a practical example. Often, the reactive coupling turns out to be negligible. In the examples to follow, all reactances are ignored so that Z_A and Z_L become real and Z_L is diagonal for similar elements.

The total average received power in the load is proportional to

$$P_{\text{rec}} = \langle V_L^H I \rangle \tag{75}$$

$$= \langle V_L^H Z_L^{-1} V_L \rangle \tag{76}$$

$$= \langle V_0^H Y V_0 \rangle \tag{77}$$

where

$$Y = (Z_A + Z_L)^{-1H} Z_L^H (\bar{Z}_A + Z_L)^{-1} V_0. \tag{78}$$

If the signals are all normalized in the sense

$$\langle V_{0n} \rangle = 1, \quad \text{all } n \tag{79}$$

then the power is expressed in terms of the correlations as

$$P_{\text{rec}} = \sum_{m,n}^{N} Y_{mn} \rho_{0mn} \tag{80}$$

$$\approx \sum_{m,n}^{N} Y_{mn} \left(\frac{R_{mn}}{R_{11}} \right) \tag{81}$$

if the conditions discussed in Section II apply.

If the mutual coupling is neglected (Z_A, Z_L become diagonal), then

$$P_{\text{rec}}^{(\text{no mc})} = \sum_{n}^{N} (Y)_{nn}. \tag{82}$$

The received power is dependent on the load resistance R_L. It is worthwhile optimizing R_L for maximum received power. The optimum load can be found from

$$\frac{\partial P_{\text{rec}}}{\partial R_L} = 0, \tag{83}$$

i.e.,

$$\sum_{m,n}^{N} \frac{\partial Y_{mn}}{\partial R_L} \rho_{0mn} = 0 \tag{84}$$

in conjunction with some physical conditions of the problem. The maximum received power is $N/4$ using this formulation.

The special cases of symmetric two- and three-element circular arrays provide interesting and useful examples. The symmetric three-element array allows the simplification that the mutual impedances and correlations are common between all branches. This cannot be the case for arrays with more than three elements with nonzero mutual impedances and correlations. Denote the real Z_A by

$$Z_A = \begin{bmatrix} R_{11} & R_{12} & \cdots \\ R_{12} & R_{11} & \cdots \\ \vdots & & \ddots \end{bmatrix} \tag{85}$$

and the elements of the real and diagonal Z_L to be R_L.

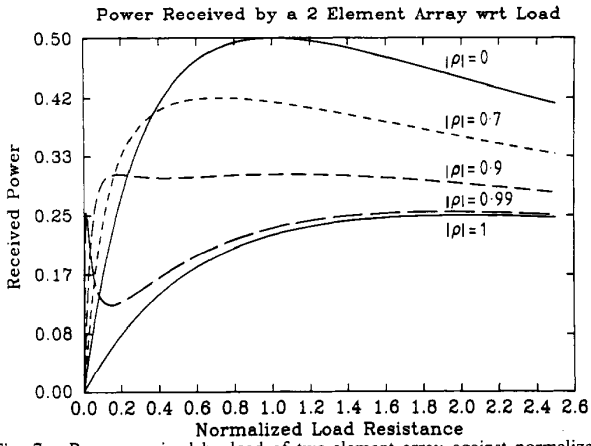


Fig. 7. Power received by load of two-element array against normalized load resistance of each element. Parameter is signal correlation coefficient, which is assumed to be equal to normalized mutual resistance.

For the two-element array,

$$Y^{(2)} = R_L \begin{bmatrix} R_{11} + R_L & R_{12} \\ R_{12} & R_{11} + R_L \end{bmatrix}^{-2} \tag{86}$$

and so

$$P_{\text{rec}}^{(2)} = \frac{2R_L/R_{11}}{((1 + R_L/R_{11})^2 - (R_{12}/R_{11})^2)^2} \cdot \{(1 + R_L/R_{11})^2 + (R_{12}/R_{11})^2 - 2\rho_{0_{12}} R_{12}/R_{11}(1 + R_L/R_{11})\} \tag{87}$$

$$= \frac{2R_L/R_{11}}{((1 + R_L/R_{11})^2 - (R_{12}/R_{11})^2)^2} \cdot \{(1 + R_L/R_{11})^2 + (R_{12}/R_{11})^2(1 - 2(1 + R_L/R_{11}))\} \tag{88}$$

if the normalized mutual resistance is taken to be the same as the open circuit correlation coefficient. Fig. 7 shows $P_{\text{rec}}^{(2)}$ plotted against the normalized load resistance with the correlation coefficient (or equivalently, the normalized mutual resistance) as a parameter. The development of a second optimum load resistance is apparent for very high correlations. The optimum load $R^{(2)}_{L_{\text{opt}}}$ is found from (84), i.e.,

$$0 = [R_{11}^2 - R_{12}^2 - R_{L_{\text{opt}}}^2][R_{11} R_{L_{\text{opt}}}^2 + 2(R_{11}^2 - 2R_{12}^2) R_{L_{\text{opt}}} + R_{11}(R_{11}^2 - R_{12}^2)] \tag{89}$$

and is plotted in Fig. 8(a) as a function of the mutual resistance. The first factor is a circle in the (R_{12}/R_{11}, $R_{L_{\text{opt}}}/R_{11}$) plane but appears as an ellipse in the digram. The second factor is cubic in R_{12}/R_{11} and R_L/R_{11} considered together but can be interpreted as a hyperbola in the $((R_{12}/R_{11})^2, R_{L_{\text{opt}}}/R_{11})$ plane. (In this plane, the first factor is a hyperbola.) The graph is single valued from (0, 1) to (0.866, 0.5) where it splits, with a locus of minima (dotted) in between. One branch goes steeply to (1, 2) while the other heads towards (1, 0). The values of $P_{\text{rec}}^{(2)}$ on these two branches seem to be equal for equal R_{12}, except right at (1, 2) and (1, 0). The branch leading to the point (1, 2) is the physically more sensible case since it corresponds to two separate antennas merging into a single antenna.

$P_{\text{rec}}^{(2)}$ for the optimum load is plotted in Fig. 8(b). Note the relative insensitivity to the mutual resistance of both the optimum load (as long as the branch leading to (1, 0) is avoided) and the received power.

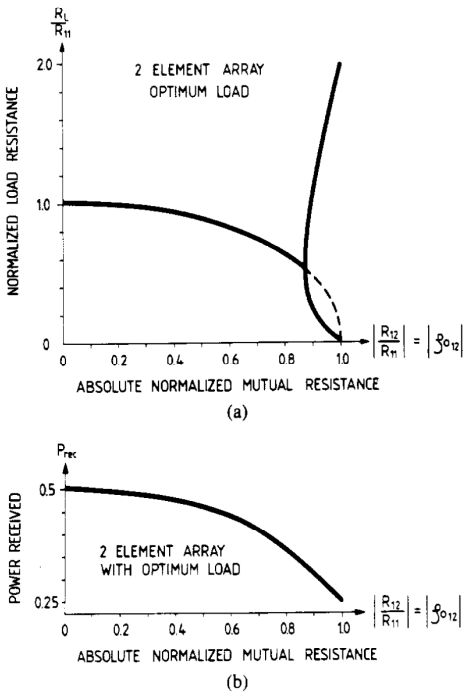


Fig. 8. (a) Locus of optimum (for maximum received power) load resistance against normalized mutual resistance (assumed equal to correlation coefficient) for two-element array. (b) Maximum received power for two-element array (with an optimum load), against normalized mutual resistance.

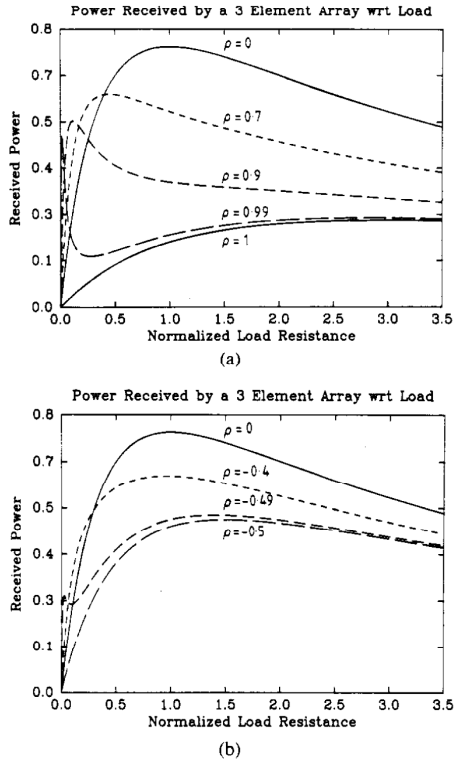


Fig. 9. Power received by load of rotationally symmetric three-element array against normalized load resistance of each element. (a) Parameter is positive correlation coefficient (assumed equal to mutual resistance). (b) Negative correlation coefficient.

For the three-element array, the expressions are of similar order to the two-element case, because the factor $(R_{11} + R_L - R_{12})^2$ cancels. The expressions are

$$P_{\text{rec}}^{(3)} = \frac{3R_L\{R_{11}[(R_{11}+R_L+R_{12})^2+2R_{12}^2-2R_{12}^2[2(R_{11}+R_L)+R_{12}]\}}{(R_{11}+R_L-R_{12})^2(R_{11}+R_L+2R_{12})} \tag{90}$$

and $R_{L_{\text{opt}}}^{(3)}$ is found from

$$\begin{aligned}0 = {} & (R_{11}+2R_{12})^2(R_{11}-R_{12})^2(R_{11}+R_{12} \\ & + R_{L_{\text{opt}}}[(R_{11}+2R_{12})(R_{11}-R_{12}) \\ & (2R_{11}^2+R_{11}R_{12}-9R_{12}^2)] \\ & + R_{L_{\text{opt}}}^2[-3R_{12}(R_{11}-R_{12})(R_{11}+2R_{12})] \\ & + R_{L_{\text{opt}}}^3[-2R_{11}^2-3R_{11}R_{12}+8R_{12}^2] \\ & + R_{L_{\text{opt}}}^4[-R_{11}]\end{aligned} \tag{91}$$

which are plotted in Figs. 9 and 10(a), respectively. The range of negative correlations, for the three-branch case, extends only to -0.5 which is a physical limit (for positive received power).

In Fig. 10(a), the locus of $R_{L_{\text{opt}}}^{(3)}$ runs from (-0.5, 1.5) through (0, 1) toward (1, 0) with a broad maximum near (0, 1). A second maximum for R_L forms near (-0.5, 1.5), which is evident in Fig. 9(a), but this maximum is seen to have a lower value of $P_{\text{rec}}^{(3)}$ and is omitted from Fig. 10(a). Near (0.94, 1.3), the locus of $R_{L_{\text{opt}}}$ splits and a branch heads steeply toward the physically sensible (1, 3). However, the values of $P_{\text{rec}}^{(3)}$ on this upper locus are not as high as on the lower dotted section.

$P_{\text{rec}}^{(3)}$ for optimum load is plotted in Fig. 10(b) against the mutual resistance. The value as the array merges to a single antenna is dotted in, since the exact curve depends on just where on the locus of $R_{L_{\text{opt}}}^{(3)}$, the "hop" from the lower locus (worse received power values, except at $R_{12}/R_{11} = 1$) to the upper locus (physically sensible at $R_{12}/R_{11} = 1$) occurs.

Near $(R_{12}/R_{11}, R_L) = (1, 0)$ in both the $N = 2$ and $N = 3$ cases, there seems to be a pathological condition in the sense that there are good and bad points within the neighborhood.

The factor F, in (16), may be obtained by dividing the right-hand side of (79), in which the Y_{mn} all contain the optimal load resistance, by the right-hand side of (80).

V. Some Examples of Mobile Diversity Antennas

Traditional Designs

All good mobile diversity antennas will have element patterns which are orthogonal, or almost orthogonal, over the MCS (see (36)). There are three degrees of freedom available

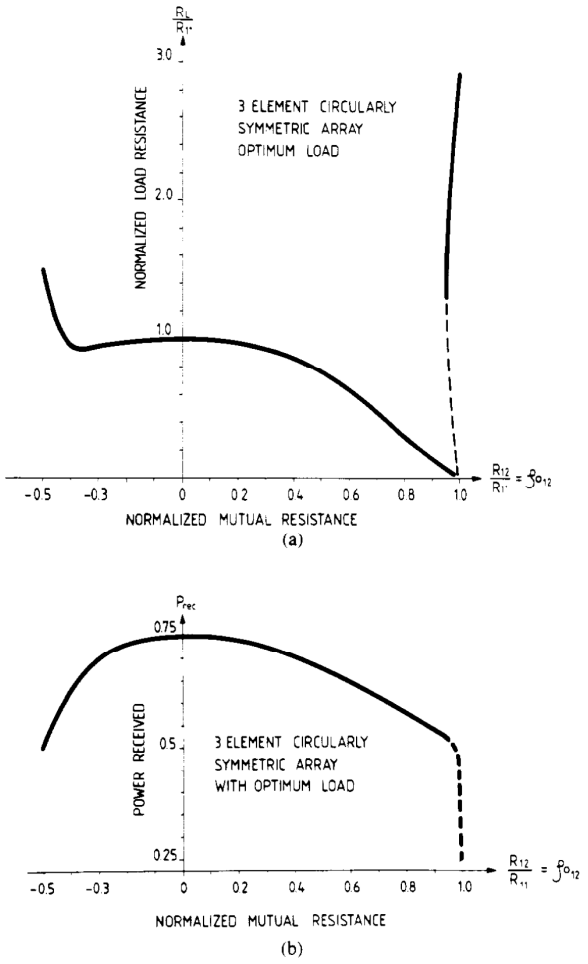


Fig. 10. (a) Locus of optimum (for received power) load resistance against normalized mutual resistance between any elements (assumed equal to correlation coefficient) for symmetric three-element array. (b) Power received by load of three-element array against normalized mutual resistance. Dotted line region is only approximate.

for obtaining the orthogonality: amplitude, phase, and polarization. Many diversity antennas employ two or all three of these possibilities. Orthogonal amplitude functions are often called angle diversity, orthogonality by phase functions accounts for space diversity antennas and polarization seems to have been rather unexploited as a degree of freedom. In antennas described below (see Figs. 11–14), all three degrees of freedom are used to some extent by seeking low values of mutual resistance/terminated circuit signal correlation. The element pattern orthogonality results without specific reference to any one of the degrees of freedom.

Pierce's energy density antenna (also known as field component diversity) reported by Gilbert [9] uses element patterns of the azimuthal form: 1, $\sin\phi$, $\cos\phi$; which are obviously orthogonal over the MCS. However, the magnetic dipole elements (giving the $\sin\phi$ and $\cos\phi$ forms) receive only half of the power compared with the vertical electric dipole in the omnidirectional urban scenario (Clarke's model). Nevertheless, this is a small price to pay for three truly uncorrelated branches occupying a common spatial position. The antenna offers a compact three-branch design concept, is wide-band, provides omnidirectional coverage, and can have closely spaced feed points. Since the E_z field can be detected by either or both of the loops, the antenna can be realized with two loops and a beamformer as reported by Lee [18]. Similarly, a single loop can be used as a two channel diversity antenna through the use of a beamformer. The beamformer is not increasing the number of ports from the loop antennas—both ends of the loop are inputs to the beamformer. The magnetic dipoles can be implemented as slots (cf. Itoh and Cheng [12], Parsons [25]) which simplifies the matching, but the beauty of a compact design is lost. A version of the single slot with monopole has been investigated more recently by Halpern and Mayes [10].

Angle diversity antennas are perhaps the simplest form of pattern orthogonality. They have not been widely applied for other than experimental work. The orthogonal magnetic dipole components of the energy density antenna are really a form of angle diversity. Single narrow beams require large antenna cross sections and for more than two branches (for example the two loops of the energy density antenna), the antennas become complicated.

Space diversity forms a special class of its own; the element patterns are identical, but the elements are appropriately spaced apart. It is not so obvious that the patterns are orthogonal over the MCS (see (36)). The multitude of sources that surround the mobile allows the phase degree of freedom to be employed with great effectiveness. Equations (33)–(35) show the development for Clarke's scenario and numerical integration for the MCS yields a result close to (35). Space diversity antennas have the disadvantage of having inherently separate feed points for each branch, and the necessary spacing becomes excessive for low-frequency operation. However, it allows the use simple omnidirectional monopoles which makes it a particularly attractive diversity technique for mobiles operating at higher frequencies. The monopoles can be spaced surprisingly close depending on the number of monopoles and their layout (cf. Lee [19]). For example, for two monopoles giving two-channel diversity, a spacing of 0.15 wave engths (5 cm at 900 MHz) gives an "acceptable" envelope correlation coefficient of less than 0.7 in a two-dimensional scenario. For three in-line monopoles, however, the spacing increases to more than 0.25 wavelengths to maintain a corresponding three-channel diversity advantage. A typical modern space diversity system is two-branch, with a spacing of 0.2 wavelengths (e.g., Miki and Hata [23]) which from the analysis leading to Fig. 12 and 13 (see below), have an open circuit *envelope* correlation coefficient of about 0.45 and a terminated circuit envelope correlation coefficient of 0.1. (The open circuit correlation is relevant for switched or selection combining and the terminated circuit correlation for nonswitched combining.)

As the base station, space diversity is considerably less practical than at the mobile because the narrow angle of incident fields calls for large spacings of the antennas. Lee [22, p. 201] notes that for an azimuthal angle width of incident sources of 0.4° and envelope correlation coefficient of 0.7, base station antennas must be spaced by 25 wavelengths for the

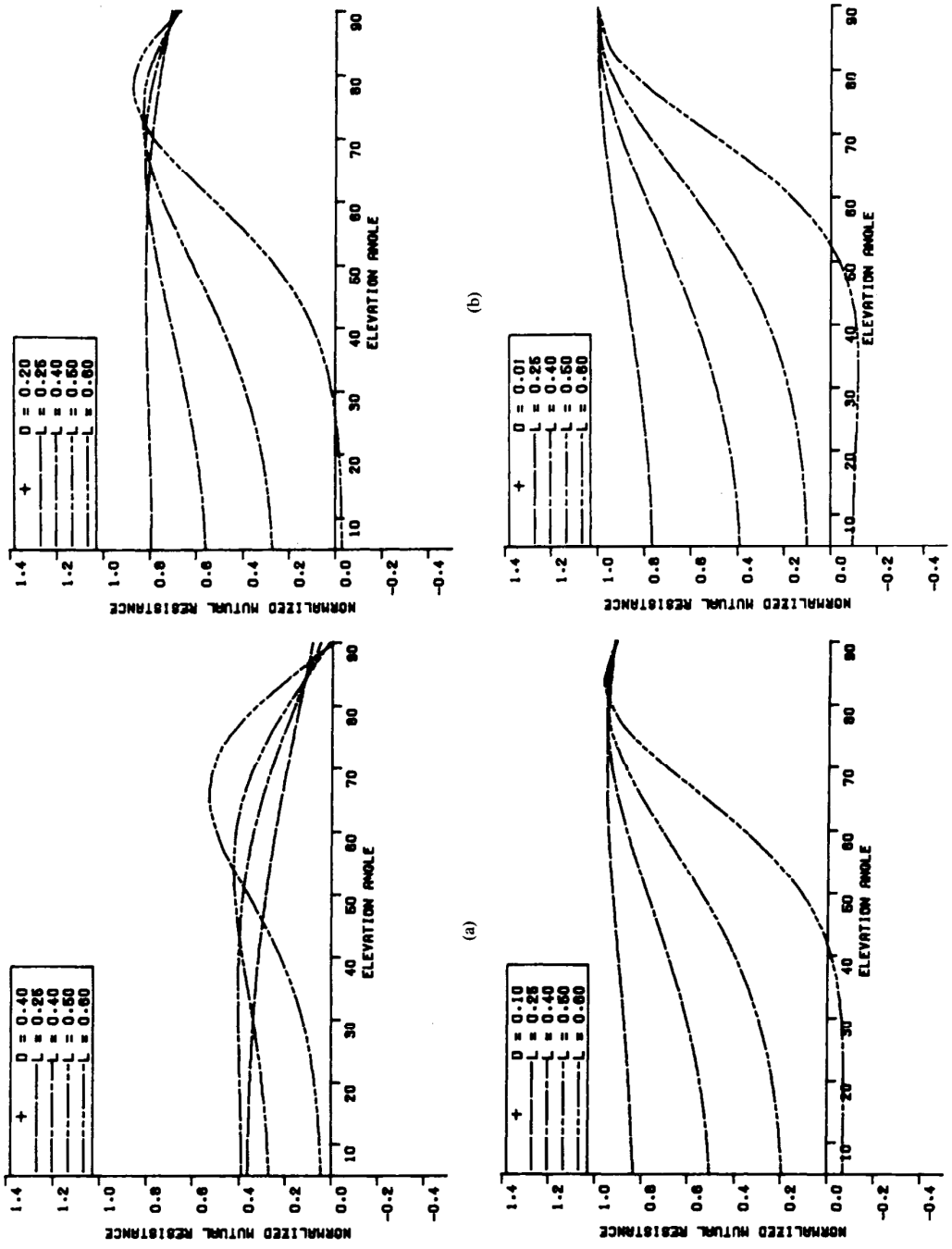


Fig. 11. Normalized mutual resistance for circular array of three outward sloping monopoles with sinusoidal current distributions on an infinite ground plane. Feedpoint spacing is D wavelengths, element length L wavelengths, and element elevation angle is argument.

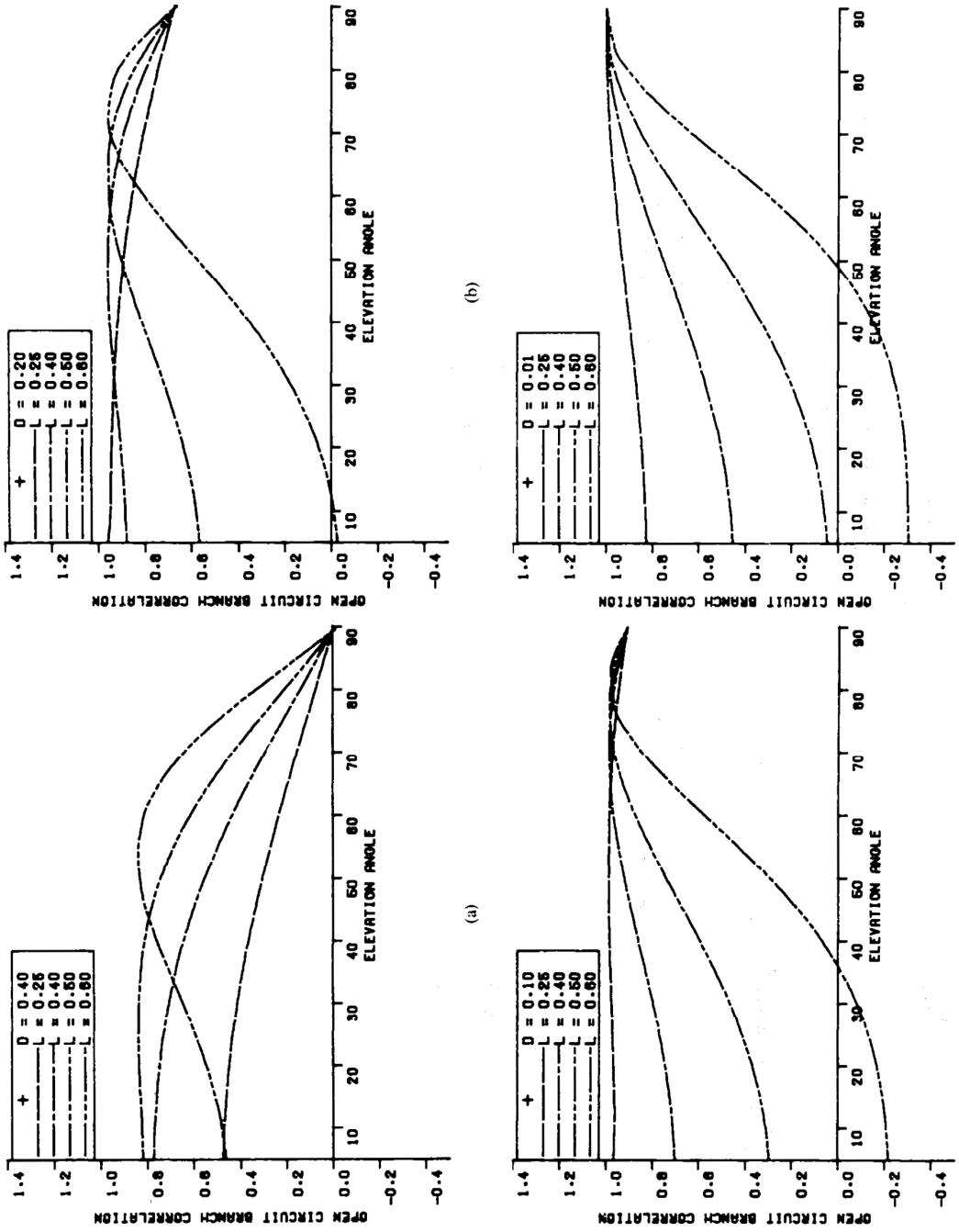


Fig. 12. Open circuit correlation coefficient for circular array of three outward sloping monopoles. Feedpoint spacing is D wavelengths, element length L wavelengths, and element elevation angle is argument.

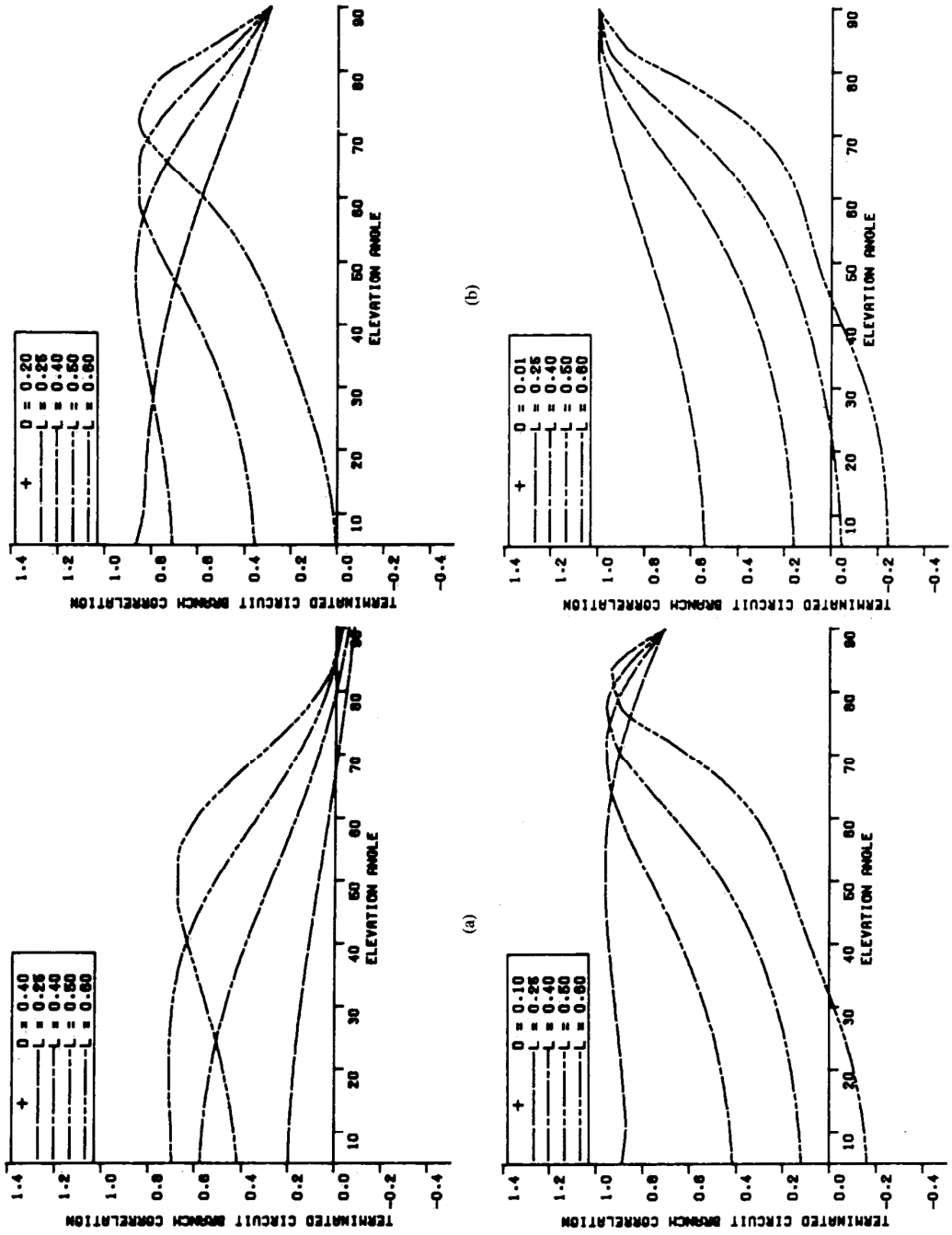


Fig. 13. Terminated circuit correlation coefficient for circular array of three outward sloping monopoles. Feedpoint spacing is D wavelengths, element length L wavelengths, and element elevation angle is argument. Elements are terminated in matched resistances for all configurations. Mutual reactances are assumed to be zero.

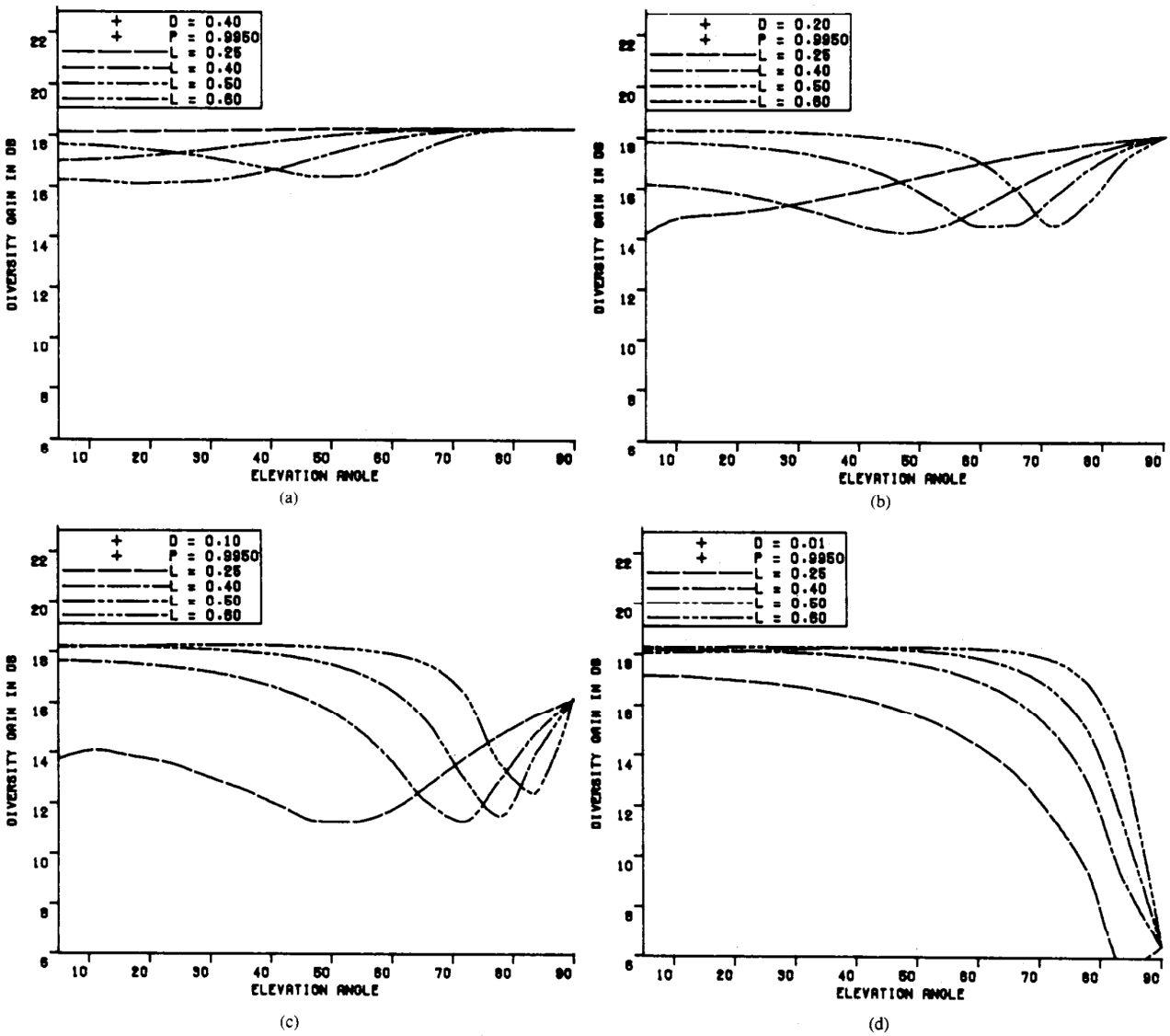


Fig. 14. Diversity gain (relative to one of array elements in presence of array) for maximally combined circular array of three outward sloping monopoles. Feedpoint spacing is D wavelengths, element length L wavelength, and element elevation angle is argument. Effects of mutual coupling are not fully included. For space diversity case of three quarter-wavelength monopoles, diversity gain is 18.25 dB, from (a).

broadside propagation case and well over 100 wavelengths for the in-line propagation case.

Polarization diversity seems to have received disproportionately little attention in the literature. For urban base stations, this technique looks promising (Lee and Yeh [21], Kozono *et al.* [17], Vaughan and Bach Andersen [31]. At the mobile, use of only the polarization degree of freedom has not been very successful. Lee and Yeh [21] measured the horizontal and vertical polarization components between mobile and base sites and reported them to be uncorrelated in the short term. Their mobile polarization diversity antenna consisted of colocated vertical electrical and magnetic (implemented as a horizontal wire loop) dipoles above the vehicle's conducting roof. The roof size was about 10 × 6 wavelengths. The resulting different array patterns from each element and its image introduce a considerable degree of (elevation) angle diversity. The presence of the vehicle roof ground plane thus complicates the mechanism of the diversity action at the mobile, it is not clear whether the decorrelation is via polarization (element pattern) or elevation angle (array pattern) diversity. Using the approach of Section III, the magnitude of the normalized scalar correlation coefficient in the case of an infinite groundplane can be found. This can be interpreted as the open circuit correlation between two antennas with far field patterns given by the array factors of electric and magnetic elements above an infinite groundplane, but of the same polarization. For any element height greater than about 0.5 wavelengths, the correlation coefficient turns out to be below 0.6 and the array patterns are effectively uncorrelated. Lee and Yeh's antenna height was 1.5 wave-

lengths, corresponding to a correlation coefficient of 0.2 (the envelope correlation is thus ~ 0.04) on an infinite ground plane.

The technique of finding the correlation coefficient from the array pattern can also be used for design of diversity antennas made from stacked monopoles (the lower section is coaxial, the outer conductor being the lower antenna and the inner conductor(s) leading to the upper monopole(s)). Stacked horizontal loops could be treated in the same way.

Diversity by Concentric Horizontal Ring Sources

The approach of Section III can be used to show that space diversity of ring antennas (cf. stacked horizontal circular microstrip antennas or circular slots), where the spacing is in the form of different radii of the circular antennas, is not particularly practical. The reason is that a very large difference in radii is necessary for pattern orthogonality. Fig. 5 indicates that the antenna with a large radius will have very low gain into the MCS, so that the gain (and thence mean SNR) of one element will be much greater than that of the other. For a scenario of the entire upper hemisphere, this type of space diversity would hold more promise. However, different modes (values of azimuthal field dependence n) or a pair of the same mode with one of the pair rotated by $\pi/(2n)$ (the patterns include cos $n\phi$ and sin $n\phi$) will give orthogonality over any rotationally symmetric scenario and practical mobile antennas result. When the same mode is used twice, the technique is really angle diversity. A microstrip antenna has been built using this approach (see Vaughan and Bach Andersen [31].

Circular Array of Three Outward Sloping Monopoles

The pattern of the vertical monopole antenna is restricted to vertical polarization. Recalling that the MCS is partially polarized, an available degree of freedom is not being utilized in seeking orthogonal element patterns, viz., reception of the horizontally polarized signals. In principle, the orthgonality condition (36) should then be more readily achieved (in terms of maintaining a close feedpoint spacing) with the extra degree of freedom of receiving both polarizations. It turns out, however, that the antenna pattern over the MCS for useful slope angles of the monopole is very strongly dominated by vertical polarization. The horizontally polarized pattern component is directed principally toward high elevation angles. However, the element orthogonality condition can still be met while imposing closely spaced feedpoints (relative to space diversity), by bending the antennas away from each other. In fact, the feedpoints can be made almost arbitrarily close if the monopoles are lengthened from the ubiquitous quarter wavelength. This result can be viewed as a type of space diversity, since the element centroids are spaced apart. Some comfort can be derived from such a viewpoint; the assumptions made in the analysis are considerable. The current distribution on all elements is assumed to be sinusoidal, even when the elements are very close. The groudplane is considered infinite, whereas a vehicle rooftop is usually only a few wavelengths across.

The sources generating the incident fields are assumed to be in the far field of the antenna. This allows element patterns to be identical, except for a rotation and translation term.

TABLE I
SUMMARY OF ENVELOPE CORRELATIONS FROM MEASUREMENTS IN AN URBAN AREA
The antenna is a circular array of monopoles of length 0.6λ, elevation angle 60° and feedpoints spacing 0.1λ.

Measured Loaded Circuit Envelope Correlations	
ρ_{L12}	0.09
ρ_{L13}	0.15
ρ_{L23}	0.12
Theoretical value	0.11

The normalized mutual resistance, open circuit correlation, and terminated circuit correlation are plotted in Figs. 11, 12, and 13, respectively, for a variety of antenna configurations. A configuration with element feedpoint spacing $D = 0.1\ \lambda$, element length $L = 0.6\ \lambda$ and element elevation angle $\alpha = 60°$ was built and tested in an urban area (4–7 story buildings either side of a narrow street) at 450 MHz. The terminated circuit *envelope* correlations are given in Table I, the agreement with the theory being excellent. Similar calculations for two-branch to eight-branch arrays yield configurations with low correlations and feedpoints as close as 0.05 λ.

Finally, Fig. 14 gives the diversity gain of the maximally combined (the calculation used terminated circuit correlations) three-branch antenna. For well-spaced feedpoints (cf. a spacing of 0.4 λ in Fig. 14(a), the diversity gain is rather insensitive to the length and elevation angle of the elements. It is of interest that for almost adjacent feedpoints (cf. 0.01λ in Fig. 14(d)), the diversity gain is not greatly reduced, until the element elevation angle approaches 20° from the vertical. It should be remembered, however, that the formulation (sinusoidal current distribution) becomes decreasingly accurate for the very closely spaced configurations. For switching rule combiners, the open circuit correlation and appropriate diversity gain (identical to the diversity return in this case) equation would be used. For nonswitched combining (or switched combiners in which the unused elements remain terminated), the diversity return is found by fully including the effect of mutual coupling. Note that if the elements are sufficiently high gain so that $r_{jk} \approx \rho 0jk$, the approximate effect can be found from the information in Figs. 10 and 12. In fact, a figure of merit for the antenna using maximum ratio combing can be defined by

$$F_A = (\text{diversity gain})\ (\text{received power loss factor due to mutual coupling})\ F_e$$

which would yield configurations for optimum performance. For switching rule combination, only the diversity gain and F_e terms are used in such a figure of merit. Optimum configurations for this type of antenna thus depend on the combiner type.

VI. Conclusion

The operating mechanism of diversity antennas for mobile communications has been examined in detail. A general relationship is given for the correlation coefficient between

antenna elements in terms of the incident fields and the element patterns.

In the presence of finite correlations (cf. compact antennas), the diversity gain is discussed and the role of mutual coupling investigated. Much progress can be made in the investigation if it can be assumed that the antenna patterns are confined to an idealized source region and the elements can be considered as minimum scattering antennas. Under such conditions, the real correlation coefficient and the mutual resistance are identical.

Practical antennas for the mobile can approximate these ideal conditions and several examples, including new designs, of mobile antennas are discussed. Measurements of the correlation coefficient for one antenna example agree well with predictions from an idealized model.

Appendix

The CPD for a Three-Element Circular Array

The cumulative probability distribution is found for a three-element circular array in which Rayleigh fading is present in each channel. The covariance matrix for a symmetrical three-element circular array is of the form

$$\boldsymbol{L} = \Gamma \begin{bmatrix} 1 & \rho & \rho \\ \rho^* & 1 & \rho \\ \rho^* & \rho^* & 1 \end{bmatrix} \tag{A1}$$

where it has been assumed that $\Gamma_1 = \Gamma_2 = \Gamma_3 = \Gamma$ and $\rho_{12} = \rho_{13} = \rho_{23} = \rho$. For the case of Rayleigh fading in each branch, the correlations become real and $\boldsymbol{L}$ reduces to a real symmetric matrix. Applying (59) (or, since the system is cubic and simple enough to manage, using the conventional technique of solving $|\boldsymbol{L} - \lambda \boldsymbol{I}| = 0$, where $\boldsymbol{I}$ is the identity matrix), the eigenvalues turn out to be

$$\lambda_1 = (1+2\rho)\Gamma$$
$$\lambda_2 = \lambda_3 = (1-\rho)\Gamma. \tag{A2}$$

Normalizing for the time being and using (56), the (Laplace) characteristic function is

$$G(s) = \frac{1}{(1+2\rho)(1-\rho)^2 \left(s + \frac{1}{(1+2\rho)}\right)\left(\frac{1}{(1-\rho)}\right)^2} \tag{A3}$$

and using the residue theorem, the probability density function is

$$p(\gamma) = a_1 + a_2 \tag{A4}$$

where a_1 is the residue of the simple pole at $s = -1/(1+2\rho)$ and a_2 is the residue of the double pole at $s = -1/(1-\rho)$. Now,

$$a_1 = \lim_{s \to -1/(1+s\rho)} \left[\frac{e^{\gamma s}}{(1+2\rho)(1-\rho)^2 \left(s + \frac{1}{1-\rho}\right)^2} \right]$$
$$= \frac{1+2\rho}{(3\rho)^2} e^{-\gamma/(1+2\rho)} \tag{A5}$$

and

$$a_2 = \lim_{s \to -1/(1+\rho)} \frac{d}{ds} \left[\frac{e^{\gamma s}}{(1+2\rho)(1-\rho)^2 \left(s + \frac{1}{1-2\rho}\right)} \right]$$
$$= -\frac{1+2\rho}{(3\rho)^2} \left[1 + \frac{3\rho\gamma}{(1+2\rho)(1-\rho)} \right] e^{-\gamma/(1-\rho)} \tag{A6}$$

so that

$$\rho_3(\gamma) = \frac{1+2\rho}{(3\rho)^2} \left[e^{-\gamma/(1+2\rho)} - \left(1 + \frac{3\rho\gamma}{(1+2\rho)(1-\rho)}\right) e^{-\gamma/(1-\rho)} \right]. \tag{A7}$$

Integrating and reintroducing Γ, the cumulative probability distribution is

$$P_3(\gamma) = 1 - \frac{1}{(3\rho)^2} \left[(1+2\rho)^2 e^{-\gamma/\Gamma(1+2\rho)} - \left[3\rho \left(\frac{\gamma}{\Gamma} + 1 - \rho \right) + (1+2\rho)(1-\rho) \right] e^{-\gamma/\Gamma(1-\rho)} \right]. \tag{A8}$$

It is worthwhile to check the limiting values when the correlation becomes 1 and 0. It is clear that

$$\lim_{\rho \to 1} P_3(\gamma) = 1 - e^{-\gamma/3\Gamma}, \tag{A9}$$

but the limit for zero correlation requires considerable manipulation. The terms $1/(1+2\rho)$ and $1/(1-\rho)$ are expanded in the Maclaurin series

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m \tag{A10}$$

followed by the corresponding expansion for the exponential terms, i.e.,

$$e^x = \sum_{m=0}^{x} \frac{x^m}{m!}. \tag{A11}$$

Powers of ρ larger than three can be discounted and after manipulation, the final result agrees with that of (10) as

$$\lim_{\rho \to 0} P_3(\gamma) = 1 - e^{-\gamma/\Gamma} \left(1 + \frac{\gamma}{\Gamma} + \frac{1}{2} \left(\frac{\gamma}{\Gamma} \right)^2 \right). \tag{A12}$$

Acknowledgment

One of the authors (RGV) acknowledges a private communication from D. Laws of the Physics and Engineering Laboratory regarding the factorization of (90) and (91). The Danish P and T, Storno A/S, and Aalborg University are thanked for their cooperation in obtaining the measurement data. Martin Langhorn is acknowledged for his contribution to the sloping monopole analysis.

REFERENCES

[1] L. Ahlin, "Coding methods for the mobile radio channel," in *Proc. Nordic Seminar on Digital Land Mobile Radiocommunications,* Espoo, Finland, Feb. 1985, pp. 185-194.

[2] H. W. Arnold and W. F. Bodtmann, "Switched-diversity FSK in frequency-selective Rayleigh fading," *IEEE Trans. Veh. Technol.,* VT-33, no. 3, pp. 156-163, 1984.

[3] J. Bach Andersen, H. A. Lessow, and H. Schjaer-Jacobsen, "Coupling between minimum scattering antennas," *IEEE Trans. Antennas Propagat.,* vol. AP-22, no. 6, pp. 832-835, 1974.

[4] J. Bach Andersen and H. Rasmussen, "Decoupling and descattering networks for antennas," *IEEE Trans. Antennas Propagat.,* vol. AP-24, pp. 841-846, 1976.

[5] R. Bellman, *Introduction to Matrix Analysis,* 2nd ed. New York: McGraw-Hill, 1970.

[6] V. M. Bogachev and I. G. Kiselev, "Optimum combining of signals in space diversity reception," *Telecommun. Radio Eng.,* vol. 34/35, no. 10, p. 83, 1980.

[7] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.,* vol. 47, pp. 957-1000, 1969.

[8] R. E. Collin and F. Zucker, *Antenna Theory, Part 1.* New York: McGraw-Hill, 1969.

[9] E. N. Gilbert, "Energy reception for mobile radio," *Bell Syst. Tech, J.,* vol. 44, pp. 1779-1803, 1965.

[10] B. M. Halpern and P. E. Mayes, "The monopole slot as a two-port diversity antenna for UHF land-mobile radio systems," *IEEE Trans. Veh. Technol.,* vol. VT-33, no. 2, pp. 76-83, 1984.

[11] H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noise Network.* New York: Wiley, 1959.

[12] K. Itoh and D. K. Cheng, "A slot-unipole energy-density mobile antenna," *IEEE Trans. Veh. Technol.,* vol. VT-21, pp. 59-62, 1972.

[13] W. C. Jakes, *Microwave Mobile Communications.* New York: Wiley, 1974.

[14] M. Kac and A. J. F. Siegert, "On the theory of noise in radio receivers with square law detectors," *J. Appl. Phys.,* vol. 18, p. 396, 1947.

[15] L. Kahn, "Ratio squarer," *Proc. IRE,* vol. 42, pp. 1704, 1954.

[16] W. K. Kahn and H. Kurss, "Minimum scattering antennas," *IEEE Trans. Antennas Propagat.,* vol. AP-13, pp. 671-675, 1965.

[17] S. Kozono, H. Tsuruhata, and M. Sakamoto, "Base station polarization diversity reception for mobile radio," *IEEE Trans. Veh. Technol.,* vol. VT-33, no. 4, pp. 301-306, 1984.

[18] W. C. Y. Lee, "An energy density antenna for independent measurement of the electric and magnetic field," *Bell Syst. Tech. J.,* vol. 46, no. 2, pp. 417-448, 1967.

[19] ——,"Mutual coupling on maximum-ratio diversity combiners and applications to mobile radio," *IEEE Trans. Commun.,* vol. COM-18, no. 6, pp. 779-791, 1970.

[20] ——,"Effect of mutual coupling on a mobile-radio maximum-ratio diversity combiners with a large number of branches," *IEEE Trans. Commun.,* vol. COM-20, pp. 1188-1193, 1972.

[21] W. C. Y. Lee and Y. S. Yeh, "Polarization diversity system for mobile radio," *IEEE Trans. Commun.,* vol. COM-20, no. 5, 1972.

[22] W. C. Y. Lee, *Mobile Communications Engineering.* New York: Wiley, 1982.

[23] T. Miki and M. Hata, "Performance of 16-GMSK transmission with postdetection selection diversity in land mobile radio," *IEEE Trans. Veh. Technol.,* vol. VT-33, no. 3, pp. 128-133, 1984.

[24] C. G. Montgomery, R. H. Dicke, and E. M. Purcell, *Principles of Microwave Circuits.* New York: McGraw-Hill, 1948.

[25] K. S. Packard, "Effect of correlation of combiner diversity," *Proc. IRE,* vol. 46, pp. 362-363, 1958.

[26] J. D. Parsons, "Field diversity antenna for UHF mobile radio," *Electron. Lett.,* vol. 10, no. 7, pp. 91-92, Apr. 1974.

[27] J. R. Pierce and S. Stein, "Multiple diversity with nonindependent fading," *Proc. IRE,* vol. 48, pp. 89-104, 1960.

[28] M. Schwartz, W. R. Bennet, and S. Stein, *Communication Systems and Techniques.* New York: McGraw-Hill, 1965.

[29] G. L. Turin, "The characteristics and function of a Hermitian quadratic form in a complex normal variable," *Biometrika,* vol 47, pp. 199-201, 1960.

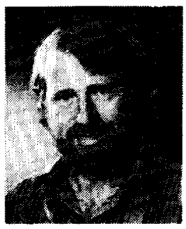
[30] R. G. Vaughan, "Signals in mobile communications, A review, *IEEE Trans. Veh. Technol.,* vol. VT-35, no. 4, pp. 133-145, 1986.

[31] R. G. Vaughan and J. Bach Andersen, "A multiport patch antenna for mobile communications," in *Proc. 14th European Microwave Conf.,* Sept. 1984, pp. 607-612.

[32] ——,"Antenna diversity in mobile communications," in *Proc. Nordic Seminar on Digital Land Mobile Radio Communications,* Espoo, Finland, Feb. 1985. pp. 87-96.

[33] W. Wasylkiwskyj and W. K. Kahn, "Theory of mutual coupling among minimum scattering antennas," *IEEE Trans. Antennas Propagat.,* vol. AP-18, no. 2, pp. 204-216, 1970.

[34] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE Trans. Veh. Technol.,* vol. VT-33, no. 3, pp. 144-155, 1984.



Rodney G. Vaughan (M'82) received the B.E. and M.E. degrees in electrical engineering from the University of Canterbury, New Zealand, in 1976 and 1977, respectively, and the Ph.D. degree from Aalborg University, Denmark, in 1985.

From 1977 to 1978 he was with the New Zealand Post Office working with toll traffic analysis and forecasting. Since 1979, he has been with the Physics and Engineering Laboratory, Department of Industrial and Scientific Research, New Zealand, developing a variety of computer based industrial and scientific equipment. He studied in Aalborg from 1982 to 1985, as a recipient of a New Zealand government study award. His current interests include mobile and satellite communications, antenna arrays, and signal processing.

Dr. Vaughan is a member of the IEEE Antennas and Propagation and Vehicular Technology Societies.

J. Bach Andersen (M'68-SM'78) received the M.Sc. and Ph.D. degrees from the Technical University of Denmark.

He was employed from 1961 to 1973 at the Electromagnetics Institute, University of Denmark. Since 1973 he has been a Professor at the Institute of Electronic Systems of Aalborg University in Jutland. His studies outside Denmark have included visits to Royal Technical University in Stockholm, Sweden, Polytechnic Institute of Brooklyn, New York, and the University of Arizona. His research interests include various aspects of applied electromagnetics, including applications in medicine such as hyperthermia and noninvasive thermometry, but also propagation of radio waves and antennas, recently with application to landmobile communications.

Dr. Andersen is on the editorial board for the *International Journal of Circuit Theory and Applications*, the *International Journal of Hyperthermia*, and the *International Journal of Electromagnetic Theory and Applications*. From 1984 to 1987 he was international chairman for Commission B of the International Union of Radio Science on Fields and Waves.