On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment

JACK H. WINTERS, MEMBER, IEEE

Abstract—**In this paper, we study the fundamental limits on the data rate of multiple antenna systems in a Rayleigh fading environment. With M transmit and M receive antennas, up to M independent channels can be established in the same bandwidth. We study the distribution of the maximum data rate at a given error rate in the channels between up to M transmit antennas and M receive antennas and determine the outage probability for systems that use various signal processing techniques. We analyze the performance of the optimum linear and nonlinear receiver processor and the optimum linear transmitter/receiver processor pair, and the capacity of these channels. Results show that with optimum linear processing at the receiver, up to $M/2$ channels can be established with approximately the same maximum data rate as a single channel. With either nonlinear processing at the receiver or optimum linear transmitter/receiver processing, up to M channels can be established with approximately the same maximum data rate as a single channel. Results show the potential for large capacity in systems with limited bandwidth.**

I. INTRODUCTION

IN a radio communication system in a multipath environment, such as a mobile radio or indoor wireless system, the communication channels between multiple transmit and/or receive antennas can have low cross correlation even when the transmit or receive antennas are closely spaced. Thus, communication systems, with appropriate signal processing techniques, can use antenna diversity (e.g., space, direction, or polarization) to establish multiple independent channels within the same bandwidth between the transmitters and receivers, thereby achieving large capacity despite the multipath.

One signal processing technique that can be used to permit multiple simultaneous signals in the same bandwidth is optimum combining at the receiver [1].[1] With optimum combining, the signals received by the antennas are combined to enhance desired signal reception and suppress interfering signals, and thereby maximize the signal-to-noise-plus-interference power in the output. With optimum combining using M antennas, up to $M - 1$ interfering signals can be nulled with desired signal reception, thus permitting up to M simultaneous signals in the same bandwidth. However, optimum combining is only the best *linear* processing technique for the receiver, and other techniques can be used to further improve performance. In particular, since all signals are detected at the receiver, interference cancellers can be used to eliminate the interference in the output signals, i.e., performance can be improved through nonlinear techniques. In addition, coding can further improve performance. Finally, performance can also be improved by cooperation between transmit antennas, i.e., appropriate combining of the signals prior to transmission (with multiple transmit antennas at the remotes). In this paper, we study two basic systems, 1) communication between multiple remotes and a base station with multiple antennas, and 2) communication between two users, each with multiple antennas. For these systems, we determine the information-theoretic capacity and the efficiency index (maximum data rate for a given error rate) in bits/cycle (bits/s/Hz) for different processing techniques. Note that since the multipath changes with position, the capacity (and efficiency index) is a random variable. Therefore, we study the distribution of the capacity and present results for given outage probability. Efficiency index results are given for a 10^{-3} outage probability at a 10^{-4} error rate. We assume independent flat (nondispersive) Rayleigh fading between antennas and constrain the total transmit power per user (remote or base station). With M transmit and M receive antennas, we note that there can be up to M independent channels between the transmitter(s) and receiver. Therefore, in case 1), for a base station with M antennas, we study the maximum data rate per remote as up to M remotes access the system. In case 2), for a receiver with M antennas, we study the maximum total data rate for the channels between the receiver and a transmitter with up to M antennas.

For communication between remotes and a base station with M antennas, we first study the efficiency index per remote (at a 10^{-3} outage probability) with optimum linear processing at the base station receiver. Results show that the efficiency index per remote decreases only slightly as up to $M/2$ remotes access the system. However, with M remotes, the efficiency index per remote is dramatically lower. We then study the efficiency index per remote with optimum nonlinear processing (maximum likelihood detection) at the base station receiver. Results show that the efficiency index per remote decreases only slightly as up to M remotes access the system. For example, with four

Manuscript received August 14, 1986; revised January 30, 1987.
The author is with AT&T Bell Laboratories, Holmdel, NJ 07733.
IEEE Log Number 8714182.

[1] In [1], optimum combining at the receiver is studied in a cellular mobile radio system. The signals received by the antennas at the base station are combined to enhance desired signal reception and suppress interfering signals from other cells, thereby allowing for frequency reuse in adjacent cells. In this paper, we consider using optimum combining, plus other techniques, to permit frequency reuse within the same cell, thereby achieving even greater capacity. Specific details on the implementation of optimum combining with frequency reuse within the same cell are presented in [2] for indoor radio.

0733-8716/87/0600-0871$01.00

antennas at a base station, for a 10^{-3} outage probability at a 10^{-4} error rate with binary phase-shift keying, the required average received signal-to-noise ratio is 13 dB with a single remote and 15 dB (only 2 dB higher) with four cochannel remotes. Finally, we study the Shannon capacity of this system.

For communication between two users, one with M antennas and the other with up to M antennas (with a total transmit power constraint on each user), we study the efficiency index with optimum linear processing at one receiver, the efficiency index with the optimum linear transmitter and receiver processing pair, and the system capacity. The efficiency index and capacity per channel (transmit antenna) for these three cases are similar to the efficiency index and capacity per remote for the three cases with the multiple remote system, considering the power constraint. For example, with the optimum linear transmitter/receiver processing pair and four antennas at the receiver, for a 10^{-3} outage probability at a 10^{-4} error rate, the required average received signal-to-noise ratio is 17 dB with one quaternary phase-shift keyed signal transmitted by one antenna and 18 dB (only 1 dB higher) with four signals transmitted by four antennas.

In Section II, we study the capacity of the system with a base station and multiple remotes. The capacity of the system with two users is analyzed in Section III. A summary and conclusions are presented in Section IV.

II. Base Station with Remotes

A. System Description

Fig. 1 shows a radio system consisting of a base station with M antennas and N ($N \leq M$) remotes, each with one antenna. We assume 1) a transmit power constraint on each remote and the base station, 2) independent Rayleigh (nondispersive) fading between each remote and base station antenna, 3) no direct communication between remotes (i.e., except through the base station), and 4) independent additive Gaussian noise at each base station receive antenna.

The system can also be represented in matrix form as in Fig. 2. The N independent input data streams can be expressed in vector form, with the nth input vector given by

$$A_n = \begin{bmatrix} a_{1,n} \\ \vdots \\ a_{N,n} \end{bmatrix} \tag{1}$$

where the $a_{i,n}$ are complex. We assume L-level quadrature amplitude modulation (QAM) such that the real and imaginary parts of $a_{i,n}$ take on values of $[\pm 1, \pm 3, \cdots, \pm (L-1)]/\sigma_d$ where $\sigma_d^2 = E[|a_{i,n}|^2] = 2(L^2-1)/3$ (for signals with unity average power). The input vector is multiplied by the transmitter matrix $\boldsymbol{P}$ to generate the transmitted vector $\boldsymbol{PA}_n$. Since we assume no direct communication between remotes, $\boldsymbol{P}$ is a diagonal matrix, i.e.,

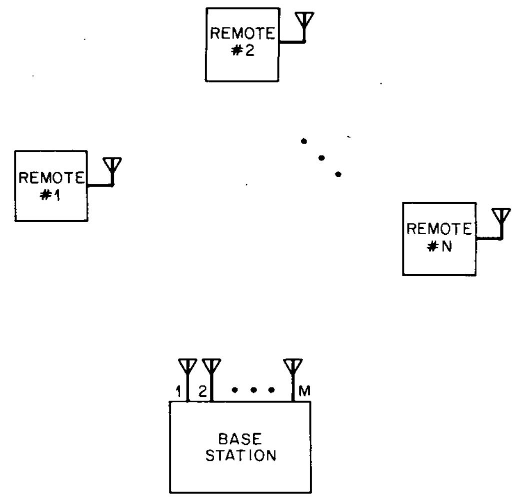

Fig. 1. Radio system consisting of a base station with M antennas and N remotes, each with one antenna.

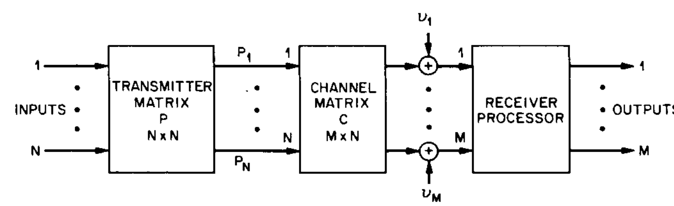

Fig. 2. System represented in matrix form.

$$\boldsymbol{P} = \begin{bmatrix} \sqrt{P_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{P_N} \end{bmatrix}. \tag{2}$$

Thus, the transmitted power of the ith signal (data stream) is given by

$$P_i = [\boldsymbol{P}^\dagger \boldsymbol{P}]_{ii}, \qquad i = 1, N \tag{3}$$

where $[\boldsymbol{P}^\dagger \boldsymbol{P}]_{ii}$ is the ith diagonal element of $\boldsymbol{P}^\dagger \boldsymbol{P}$ and the superscript $\dagger$ denotes complex conjugate transpose. We constrain the transmit power of each remote such that

$$P_i \leq 1, \qquad i = 1, N. \tag{4}$$

The signal vector received by the M antennas at the base station is the transmitted signal vector multiplied by the channel matrix $\boldsymbol{C}$, $\boldsymbol{CPA}_n$, plus additive Gaussian noise. Under the assumption of independent Rayleigh fading, the elements of the $M \times N$ channel matrix C_{ij} are complex Gaussian random variables, i.e., the real and imaginary parts of C_{ij} are Gaussian random variables with zero mean and a variance of $\frac{1}{2}$ (for equal average transmit and receive powers). The noise vector v_n is given by

$$v_n = \begin{bmatrix} v_{1,n} \\ \vdots \\ v_{M,n} \end{bmatrix} \tag{5}$$

where $v_{i,n}$ are independent, complex Gaussian random variables, i.e., the real and imaginary parts are Gaussian

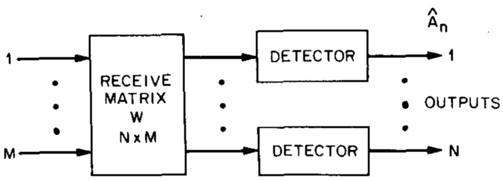

Fig. 3. A linear receiver processor.

random variables with zero mean and variance $\sigma_n^2/2$. The received signal vector S_n is then given by

$$S_n = CPA_n + v_n. \tag{6}$$

The M received signals are processed at the receiver to generate the output signal vector $\hat{A}_n$ (an estimate of the transmitted signal vector).

The system shown in Fig. 2 is similar to that studied for digital radio [3]–[6]. In [3], a single-channel ($N = M = 1$) digital radio system with frequency selective fading (dispersion) is studied. The performance of equalization and maximum likelihood sequence estimation is analyzed, using the probability distribution of the efficiency index as the performance measure. The efficiency index is an estimate of the maximum number of bits per cycle of bandwidth that can be achieved in a given system and, in [3], it is obtained by using a Chernoff bound on the error rate for given signal-to-noise ratio, channel bandwidth, and signaling rate with quadrature amplitude modulation. In [4], the analysis of [3] is extended to digital radio with dual polarization ($N = M = 2$). The structure of the optimum transmitter and receiver matrix filters is studied, and the information-theoretic capacity and efficiency index with these filters is analyzed. In [5], the results of [4] are extended to the $N \times N$ channel matrix, and in [6], the results are extended to the $M \times N$ channel matrix. For our system, one major difference from digital radio is that we do not have dispersive fading. Therefore, in our study, we use the extensive analysis of [3]–[6], simplifying the results for nondispersive fading.

B. Optimum Linear Processing

We first consider linear processing at the base station receiver. The linear processor is shown in Fig. 3 where the received signals are combined using the receiver $N \times M$ matrix W to generate the output signals for detection. Thus, at the output of the linear processor, the nth output vector is given by

$$Z_n = W[CPA_n + v_n]. \tag{7}$$

The detected symbols $\hat{A}_n$ are determined from the real and imaginary parts of Z_n using a decision rule with decision levels at $[0, \pm 2, \cdots, \pm(L-2)]/\sigma_d$ (for L-level QAM).

We consider an optimum linear processor as optimum in the sense of minimizing the mean-squared error MSE of each of the N output signals. Note that the receive matrix W that minimizes the MSE is the same matrix as that which maximizes the signal-to-interference-plus-noise ratio (optimum combining, as studied in [1]) [7]. Furthermore, this optimum W also minimizes an upper bound on the symbol error rate [3].

For a QAM signal with optimum linear processing at the receiver, the MSE in each of the output signals is given by [4]

$$\text{MSE}_i = \frac{\overline{\text{MSE}_i}}{\sigma_d^2} = [I + P^\dagger C^\dagger C P \rho]_{ii}^{-1}, \qquad i = 1, N \tag{8}$$

where $\rho\,(=\sigma_d^2/\sigma_n^2)$ is the received signal-to-noise ratio at each antenna without fading (i.e., the signal-to-noise ratio averaged over the Rayleigh fading). For each of the N signals, the maximum data rate that can be supported at a given symbol error rate P_e (the efficiency index) is then given by [4]

$$I_i = \log_2\left(1 + 1.5\,\mu_i \Big/ \left|\ln\frac{P_e}{2}\right|\right), \qquad i = 1, N \tag{9}$$

where

$$\mu_i = (1 - \text{MSE}_i)\text{MSE}_i. \tag{10}$$

For $N = 1$, the efficiency index is the highest when $P_1 = 1$ [for the power constraint of (4)]. Thus, from (8)–(10),

$$I_1 = \log_2\left(1 + 1.5\rho\, C^\dagger C \Big/ \left|\ln\frac{P_e}{2}\right|\right). \tag{11}$$

Since $C^\dagger C$ is the sum of the magnitude squared of complex Gaussian random variables with zero mean and a variance of $\frac{1}{2}$, $C^\dagger C$ is a chi-squared random variable with $2M$ degrees of freedom. Thus, the probability density function of $C^\dagger C$ is given by

$$p(C^\dagger C) = \frac{(C^\dagger C)^{M-1}\, e^{-C^\dagger C}}{(M-1)!}. \tag{12}$$

From (11) and (12), the probability density function of I_1 can be shown to be given by

$$p(I_1) = \frac{(2^{I_1} - 1)^{M-1}\, e^{-(2^{I_1}-1)/\alpha}\, 2^{I_1} \ln 2}{\alpha^M (M-1)!} \tag{13}$$

where

$$\alpha = 1.5\rho \Big/ \left|\ln\frac{P_e}{2}\right|. \tag{14}$$

Thus, the probability distribution of I_1 can be calculated as

$$P(I_1) = 1 - e^{-(2^{I_1}-1)/\alpha} \sum_{k=0}^{M-1} \frac{\left(\frac{2^{I_1}-1}{\alpha}\right)^k}{k!}. \tag{15}$$

Since with $N = 1$ optimum linear processing is the same as maximal ratio combining, (15) can also be determined from the maximal ratio combining equations in [8].

For $N \geq 2$, the values of the P_i's ($i = 1, N$) for the maximum efficiency index per remote could not be easily determined. Therefore, implementation of a system that optimizes the P_i's does not appear to be practical. Since

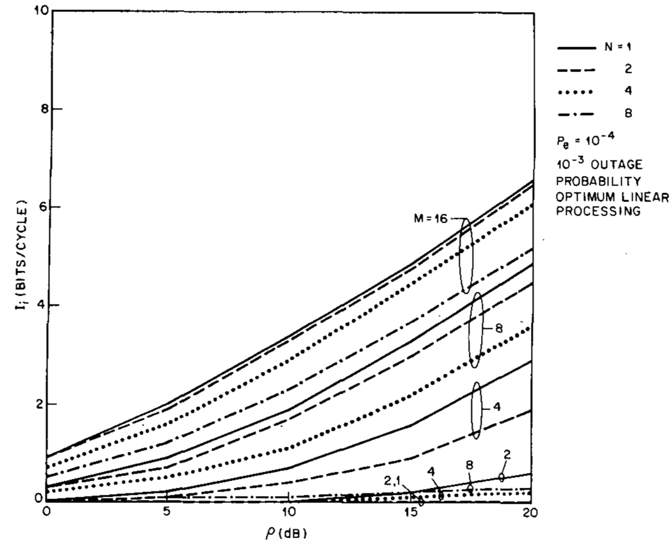

Fig. 4. The efficiency index per remote versus ρ for a 10^{-3} outage probability at a 10^{-4} error rate for optimum linear processing at the base station receiver; N remotes and M antennas at the base station.

one practical technique is to adjust the transmit powers such that the received signal power for each signal is constant, we first examined this case. For this system,

$$P_i = \frac{\min_j \, [C^{\dagger}C]_{jj}}{[C^{\dagger}C]_{ii}}, \qquad i = 1, N. \tag{16}$$

Results, however, show that the efficiency index is slightly higher when $P_i = 1$, $i = 1, N$. Therefore, in the analysis below, we only present results for $P_i = 1$ $(i = 1, N)$, noting that the results do not appear to be significantly affected by small variations in P_i.

For $N \geq 2$, the distribution of I_i was determined by Monte Carlo simulation. For given N, M, ρ, and P_e, the I_i's $(i = 1, N)$ were calculated for 10^4 randomly generated C matrices. The distribution of these I_i's then determined the outage probability.

Fig. 4 shows the efficiency index versus ρ for a 10^{-3} outage probability at a 10^{-4} error rate, with $N = 1, 2, 4, 8$, and $M = 1, 2, 4, 8, 16$. Analytical results are shown for $N = 1$, and computer simulation results are shown for $N \geq 2$. For fixed M, the efficiency index per remote decreases as N increases (i.e., as more remotes access the system). Fig. 4 shows that the decrease in the efficiency index is small as N is increased from 1 to $M/2$, but for $N = M$, the efficiency index decreases dramatically. This is because, for $N << M$, N random signal vectors in an M-dimensional space (with M antennas) usually have low cross correlation (i.e., interference), and therefore, I_i is not significantly affected by the interference. However, as N approaches M (in particular, for $N > M/2$), N random signal vectors become increasingly likely to have high cross correlation (interference), and thus, I_i is greatly reduced by the interference.

Finally, considering implementation of optimum linear processing [1], [2], we note that the optimum W can be determined without knowledge of the C matrix by means of iterative techniques, such as the LMS algorithm [7]. Also, for base station to remote transmission, W^* (where * denotes complex conjugate) can be used as the transmitter matrix to obtain the same efficiency index as with remote to base station transmission. It should be emphasized that the major concern for implementation of the techniques in this paper is the fading rate. The processors must operate fast enough to track the fading [1], [2], and, if C must be known, the channel must be probed often enough so that the fading can be tracked. Thus, implementation may be much easier in indoor radio systems (where the users are stationary or walking) than in mobile radio systems (where the users can move at vehicular speeds). Also, it should be noted that our results are based on the assumption of nearly independent fading at each antenna. With multipath in buildings, the fading statistics of two antennas are (usually) nearly independent when the antennas are separated by more than a quarter wavelength. Furthermore, in addition to space diversity, direction [9], polarization [10], [11], and field [8] diversity can be used to achieve nearly independent fading at each antenna without increasing the physical size of the antenna array.

C. Optimum Nonlinear Processing

We now consider nonlinear processing at the base station receiver. We consider an optimum nonlinear processor as optimum in the sense of minimizing the probability of error in detecting A_n. Thus, the optimum nonlinear processor is a maximum likelihood detector [6], [12]. For a maximum likelihood detector, the error rate is approximately given by

$$P_e \approx e^{-d_{\min}^2 \rho / 4} \tag{17}$$

where

$$d_{\min}^2 = \min_{\substack{A_n, \hat{A}_n \\ A_n \neq \hat{A}_n}} \left| CP(A_n - \hat{A}_n) \right|^2. \tag{18}$$

Thus, given C, P, and L (number of levels for QAM), we can calculate the error rate at a given ρ by exhaustively searching all A_n and $\hat{A}_n$ for $d_{\min}^2$ and calculating P_e from (17). The efficiency index per remote is then $2 \log_2 L$. However, this procedure requires extensive computer time for large L and N. Therefore, we restrict our study to quarternary phase-shift keying QPSK $(L = 2)$ with $N = 1$ and 2, and binary phase-shift keying BPSK with $N = 1, 2$, and 4. Note that for BPSK, the data symbols $a_{i,n}$ in (1) are real, with values of ± 1.

We determined the ρ required for QPSK $(I_i = 2)$ and BPSK $(I_i = 1)$ by the following method. Using Monte Carlo simulation, for given N and M, 10^4 random C matrices were generated. For each matrix, $d_{\min}^2$ was determined by exhaustive search of the possible combinations of data symbols for either QPSK or BPSK, and the ρ required for a given P_e was calculated from (17). From the distribution of required ρ, we determined the ρ for a given outage probability.

TABLE I
AVERAGE SIGNAL-TO-NOISE RATIO REQUIRED FOR A 10^{-3} OUTAGE PROBABILITY AT A 10^{-4} ERROR RATE, WITH OPTIMUM NONLINEAR AND OPTIMUM LINEAR PROCESSING

M	N	Required ρ (dB)			
		BPSK		QPSK	
		nonlinear	linear	nonlinear	linear
1	1	39.6	39.6	43.0	43.0
2	1	23.0	23.0	26.4	26.4
	2	24.1	38.3	29.1	41.8
4	1	13.3	13.3	16.7	16.7
	2	13.9	16.7	17.3	20.5
	4	15.1	40.0	-	42.7
8	1	6.7	6.7	10.1	10.1
	2	7.1	8.3	10.5	11.2
	4	7.5	10.8	-	14.0
16	1	1.5	1.5	4.9	4.9
	2	1.7	2.1	5.1	5.4
	4	2.2	3.3	-	6.5

As in Section II-B, the P_i's ($i = 1, N$) that maximize $d^2_{\min}$ are difficult to calculate. Therefore, we study only the case of $P_i = 1$ ($i = 1, N$). Note also that for the optimum nonlinear processor, we calculated the error rate for the output vector rather than the output symbols as studied for the optimum linear processor. Thus, for given $\boldsymbol{C}$, the efficiency index is the same for all remotes, unlike with optimum linear processing.

Since results were obtained only for QPSK and BPSK, we study the ρ required for these two modulation techniques, and compare the required ρ to optimum nonlinear processing to that with optimum linear processing. For QPSK, the required ρ for optimum linear processing with $I_i = 2$ can be obtained as in Section II-B. For BPSK ($I_i = 1$), the required ρ with optimum linear processing can also be obtained as in Section II-B, but with the efficiency index given by

$$I_i = \tfrac{1}{2} \log_2 \left(1 + 3\,\mu_i/|\ln P_e|\right), \qquad i = 1, N \quad (19)$$

rather than (9).

Table I shows the ρ required for a 10^{-3} outage probability at a 10^{-4} error rate for QPSK and BPSK with given N and M. For fixed M, the required ρ increases as N increases with both optimum linear and nonlinear processing. However, with optimum nonlinear processing, up to M remotes can access the system with only a few dB increase in ρ, while with optimum linear processing, ρ increases dramatically with $N = M$. Thus, optimum nonlinear processing is significantly better than optimum linear processing only for $N > M/2$.

Finally, considering implementation of optimum nonlinear processing, we note that optimum nonlinear processing requires that $\boldsymbol{C}$ be known at the receiver. Thus, the channel must be probed prior to data transmission. However, performance close to that of optimum nonlinear processing can be achieved without a knowledge of $\boldsymbol{C}$ through the use of optimum linear processing followed by interference cancellation (such as with a bootstrap canceller [13]), but such a system has not been studied. Also, we note that we cannot use optimum nonlinear processing for base station to remote transmission and, therefore, the efficiency index for base station to remote transmission without coding can only be as high as that of optimum linear processing (Section II-B).

D. Capacity

Finally, we consider the distribution of an upper bound on the capacity normalized to the bandwidth for the channels between the remotes and the base station. If the remotes' signals were weighted and combined prior to transmission to the base station, the normalized capacity can be determined by analyzing the independent channels in $\boldsymbol{C}$ [4]. From [4], for given $\boldsymbol{C}$, the normalized capacity of the ith independent channel is given by

$$I_i = \log_2 (1 + \rho \lambda_i P_i), \qquad i = 1, N \quad (20)$$

where λ_i is the ith eigenvalue of $\boldsymbol{C}^{\dagger}\boldsymbol{C}$. Since the channels are independent (i.e., no cross-coupled interference), $P_i = 1$ ($i = 1, N$) maximizes the capacity in each channel. Thus, for given $\boldsymbol{C}$, the average capacity per remote is given by

$$I_s/N = \frac{1}{N}\sum_{i=1}^{N} I_i = \frac{1}{N}\sum_{i=1}^{N} \log_2 (1 + \rho\lambda_i). \quad (21)$$

Without the combining of the signals prior to transmission (i.e., with each remote transmitting only its signal), the average capacity per remote is less than or equal to (for $N = 1$) that given by (21). Thus, (21) upper bounds the average capacity per remote.

For $N = 1$, $\lambda_1 = \boldsymbol{C}^{\dagger}\boldsymbol{C}$, and, therefore, from (15), the distribution of the normalized capacity is given by

$$P(I_1) = 1 - e^{-(2^{I_1}-1)/\rho} \sum_{k=0}^{M-1} \frac{\left(\dfrac{2^{I_1}-1}{\rho}\right)^k}{k!}. \quad (22)$$

Thus, the normalized capacity at ρ' is the same as the efficiency index for a 10^{-4} error rate without coding at $\rho = \rho' |\ln (P_e/2)|/1.5$ (or $\rho = \rho' + 8.2$ dB).

For $N \geq 2$, Monte Carlo simulation was again used to determine the distribution of the average normalized capacity. For given N, M, and ρ, the I_i's and I_s/N were determined for 10^4 randomly generated $\boldsymbol{C}$ matrices, and the distribution of I_s/N was calculated.

Fig. 5 shows the upper bound on the average normalized capacity per remote at a 10^{-3} outage probability versus ρ for $N = 1, 2, 4, 8$ and $M = 1, 2, 4, 8, 16$. For fixed M, in most cases the capacity increases with N, since as N increases, I_s/N is averaged over more channels. Note that

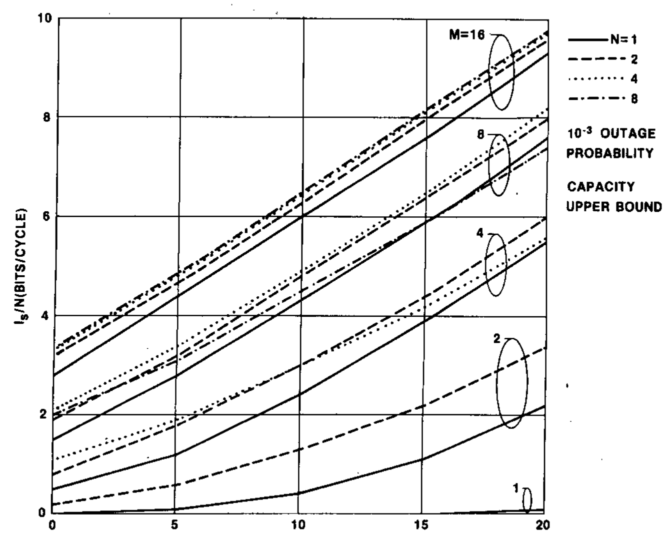

Fig. 5. Upper bound on the average normalized capacity per remote versus ρ for a 10^{-3} outage probability with N remotes and a base station with M antennas.

this increase in capacity with N is greatest for small M where the fading effects are the strongest. However, for fixed M, as N increases, the interference also increases, which reduces the capacity. Whether the combined effect of the averaging and interference increases or decreases the capacity depends on N, M, and ρ. Fig. 5 shows that the interference dominates (with the capacity reduced) for $N = M$, $M \geq 4$, and that the reduction in capacity increases with ρ.

III. Two Users with Multiple Antennas

A. System Description

Fig. 6 shows a radio system consisting of two users, one with M antennas and the other with N ($N \leq M$) antennas. We assume 1) a transmit power constraint on both users, 2) independent Rayleigh (nondispersive) fading between the transmit and receive antennas, and 3) independent additive Gaussian noise at each receiver antenna. With multiple antennas for both users, the same maximum data rate can be obtained in both transmission directions. We study only the case of linear processing at both the transmitter and receiver.

The system can be represented in matrix form as in Section II. The differences with the system of Section II are, first, with multiple transmit antennas, the input signals can be combined prior to transmission, i.e., $\boldsymbol{P}$ need not be diagonal. Second, with the power constraint

$$\sum_{i=1}^{N} [\boldsymbol{P}^{\dagger}\boldsymbol{P}]_{ii} = 1, \tag{23}$$

the total transmit power is $1/N$ times the total transmit power of the system of Section II. Third, we are interested in the maximum total data rate for given $\boldsymbol{C}$ and the distribution of this data rate, rather than the distribution of the maximum data rate in each channel.

Fig. 6. Radio system consisting of two users, one with M antennas and the other with N antennas.

B. Optimum Linear Processing at the Receiver

We first consider optimum linear processing at the receiver. The analysis is similar to that of Section II-B, except that we calculate

$$I_s = \sum_{i=1}^{N} I_i \tag{24}$$

where I_i is given in (9), and we have the power constraint of (23). For $N = 1$, the results are, of course, identical to those of Section II-B.

For $N \geq 2$, we again need to consider the P_i's. As in Section II-B, the P_i's for maximum total data rate could not be easily determined. Thus, implementation of a system that optimizes the P_i's does not appear practical. One practical method, since we are interested in maximizing the total data rate, is to adjust the transmit power of each antenna proportional to its received signal power (similar to the processing method used for optimum linear processing without interference). Thus, the transmit powers are given by

$$P_i = \frac{[\boldsymbol{C}^{\dagger}\boldsymbol{C}]_{ii}}{\sum_{j=1}^{N} [\boldsymbol{C}^{\dagger}\boldsymbol{C}]_{jj}}. \tag{25}$$

Results, however, show that the maximum total data rate is slightly higher when all the transmit powers are equal, i.e., $P_i = 1/N$ ($i = 1, N$). Therefore, in the analysis below, we present results only for equal transmit powers.

For $N \geq 2$, the distribution of I_s was determined by Monte Carlo simulation. For given N, M, ρ, and P_e, the I_s's were determined for 10^4 randomly generated $\boldsymbol{C}$ matrices and the distribution of I_s was calculated to determine the outage probability.

Now consider the average maximum data rate per channel I_s/N at a given outage probability. For fixed M, the signal-to-noise ratio per channel is decreased by $1/N$ with N transmit antennas. Thus, it might be expected that the ρ required for a given I_s/N would increase linearly with N. However, there are two other effects. First, as N increases, the I_s/N is averaged over more channels, which decreases the ρ required for a given I_s/N at a given outage probability. Second, as N increases, the interference increases, which increases the ρ required for a given I_s/N. Which of these two effects dominates depends on N and M as shown below.

Fig. 7 shows the average efficiency index per channel (I_s/N) versus ρ for a 10^{-3} outage probability at a 10^{-4} error rate, with $N = 1, 2, 4, 8$ and $M = 1, 2, 4, 8, 16$.

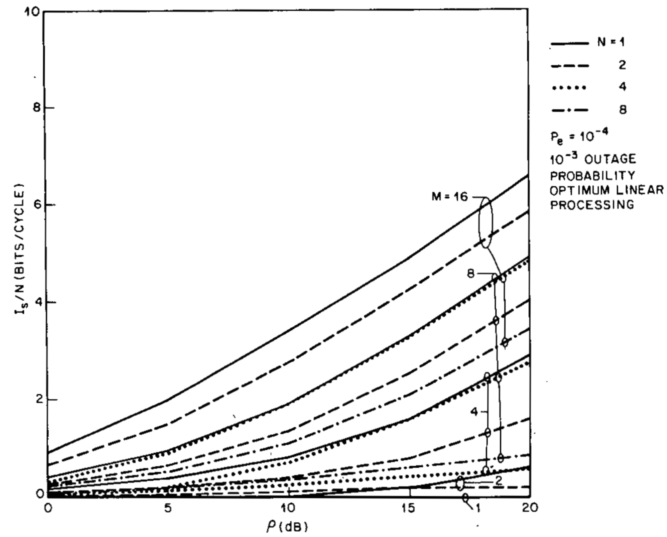

Fig. 7. The average efficiency index per channel versus ρ for a 10^{-3} outage probability at a 10^{-4} error rate for optimum linear processing at one receiver; N transmitting antennas and M receiving antennas.

With large M/N, the averaging effect is seen to dominate. For example, with $M = 16$, the ρ required for a given I_s/N increases by only 2 dB when N is increased from 1 to 2. However, as N increases further, the interference dominates. Thus, with fixed M, the ρ required for a given I_s/N increases linearly with N for $N \leq M/2$, but for $N = M$, the required ρ is much higher. This is because of the greatly increased probability of high cross correlation between N random signal vectors as N approaches M, i.e., for the same reason that in a multiple remote system I_i is greatly reduced as N approaches M (as shown in Fig. 4 and discussed in Section II-B).

C. Optimum Transmitter/Receiver Processing

We now consider the combining of the input signals prior to transmission. In this case, we are interested in the optimum linear transmitter/receiver processor pair where the processor pair is optimum in the sense of minimizing the total MSE at the receiver output. For given $\boldsymbol{C}$, with the jointly optimum $\boldsymbol{P}$ and $\boldsymbol{W}$, the total MSE is of the form [6]

$$\text{MSE} = \sum_{i=1}^{N} \text{MSE}_i = \sum_{i=1}^{N} \frac{1}{1 + \rho\lambda_i P_i} \tag{26}$$

where the λ_i's are the eigenvalues of $\boldsymbol{C}^\dagger \boldsymbol{C}$ and

$$\sum_{i=1}^{N} P_i = 1. \tag{27}$$

At this point, we could, of course, determine the P_i's that minimize the total MSE. However, we wish to maximize the total data rate, which is given by [from (9), (10), and (26)]

$$I_s = \sum_{i=1}^{N} \log_2 \left(1 + 1.5\rho P_i \lambda_i \Big/ \left|\ln \frac{P_e}{2}\right|\right). \tag{28}$$

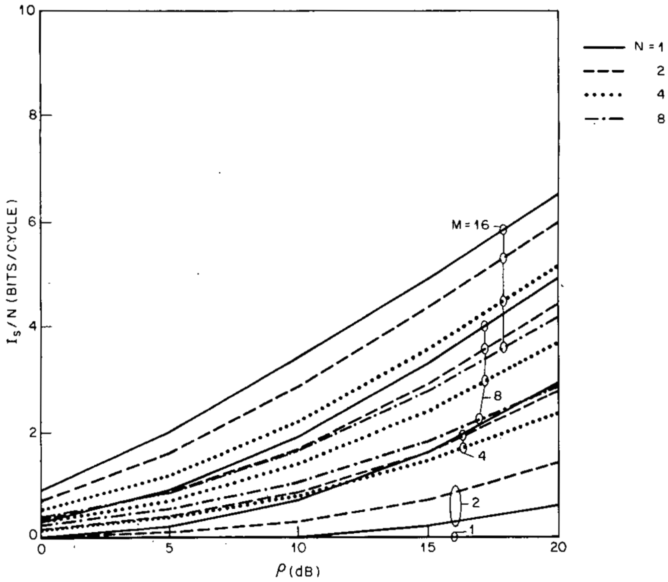

Fig. 8. The average efficiency index per channel versus ρ for a 10^{-3} outage probability at a 10^{-4} error rate for the optimum transmitter/receiver matrix pair; N transmitting antennas and M receiving antennas.

The P_i's that maximize I_s can be found by using the water fill analogy [14], i.e.,

$$P_i = \begin{cases} J - \left[1.5\rho\lambda_i \Big/ \left|\ln \frac{P_e}{2}\right|\right]^{-1} \\ \qquad \text{if } \left[1.5\rho\lambda_i \Big/ \left|\ln \frac{P_e}{2}\right|\right]^{-1} < J \\ 0 \qquad \text{otherwise} \end{cases} \tag{29}$$

where

$$J = \frac{1}{m}\left(1 + \sum_i \left[1.5\rho\lambda_i \Big/ \left|\ln \frac{P_e}{2}\right|\right]^{-1}\right) \tag{30}$$

and the sum is over the m terms where $[1.5\rho\lambda_i/|\ln(P_e/2)|]^{-1} < J$.

Fig. 8 shows the average efficiency index per channel versus ρ for a 10^{-3} outage probability at a 10^{-4} bit error rate with $N = 1, 2, 4, 8$ and $M = 1, 2, 4, 8, 16$. With fixed M, the ρ required for fixed I_s/N increases by less than 3 dB as N is doubled. Thus, the averaging effect described in Section III-B dominates for $N \leq M$. Note that because of the averaging effect, for $M = 2$, the I_s/N even increases as N increases from 1 to 2. For $N < M/2$, the average efficiency index per channel is similar to that with optimum linear processing at the receiver only. However, for $N > M/2$, the I_s/N is much higher.

Finally, considering implementation of the optimum transmitter/receiver processor pair, we note that to determine the jointly optimum $\boldsymbol{P}$ and $\boldsymbol{W}$, $\boldsymbol{C}$ must be known. However, with these matrices, the efficiency index is the same in both transmission directions.

D. Capacity

Finally, we consider the distribution of the total normalized capacity (Shannon limit) for the channels between the two users. The capacity can be determined by analyzing the independent channels in C [4], as in Section II-D. Thus, the total normalized capacity is given by

$$I_s = \sum_{i=1}^{N} \log_2 (1 + \rho \lambda_i P_i) \tag{31}$$

where λ_i is the ith eigenvalue of $C^\dagger C$ and $\Sigma_{i=1}^N P_i = 1$. This is the same formula as (28) with ρ replaced by $1.5\rho / |\ln (P_e/2)|$. Thus, Fig. 8 also shows the average normalized capacity power channel at $\rho' = 1.5\rho/|\ln (P_e/2)|$. That is, the normalized capacity at ρ is the same as the maximum data rate for a 10^{-4} error rate without coding at ρ + 8.2 dB.

IV. Summary and Conclusions

In this paper, we studied, for given outage probability, the maximum data rate (at a given error rate) and the capacity of multiple antenna systems in a Rayleigh fading environment. In such an environment, up to M independent channels can be established between M transmit and M receive antennas. Results show that for a base station with M antennas, up to $M/2$ remotes (each with one antenna) can access the base station that uses optimum linear processing, with about the same maximum data rate as a single remote. However, the maximum data rate per remote is much lower with M remotes. With optimum nonlinear processing at the base station, up to M remotes can access the base station with about the same maximum data rate as a single remote.

Results for two users, each with M antennas, show that with optimum linear processing at one receiver, up to $M/2$ independent channels can be established between the users, with each channel having about the same maximum data rate as a single channel. With the optimum transmitter/receiver processor pair, up to M channels can be established between the users, with each channel having about the same maximum data rate as a single channel. The capacity (the maximum data rate with an essentially zero error rate) at a given average received signal-to-noise ratio ρ was shown to be the same as the maximum data rate without coding at $1.5\rho/|\ln (P_e/2)|$ (ρ + 8 dB for $P_e = 10^{-4}$).

In summary, we have described and studied the fundamental limits on systems that exploit multipath to allow multiple simultaneous users (or channels) in the same bandwidth. Results show the potential for large capacity in systems with limited bandwidth.

References

[1] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 528-539, July 1984, (also *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 144-155, Aug. 1984).

[2] —, "Optimum combining for indoor radio systems with multiple users," submitted to *IEEE Trans. Commun.*, also to be presented at ICC '87.

[3] G. J. Foschini and J. Salz, "Digital communications over fading radio channels," *Bell. Syst. Tech. J.*, vol. 62, pp. 429-456, Feb. 1983.

[4] N. Amitay and J. Salz, "Linear equalization theory in digital data transmission over dually polarized fading radio channels," *Bell Syst. Tech. J.*, vol. 63, pp. 2215-2259, Dec. 1984.

[5] J. Salz, "Digital transmission over cross-coupled linear channels," *Bell Syst. Tech. J.*, vol. 64, pp. 1147-1159, July-Aug. 1985.

[6] J. Salz and A. D. Wyner, "On data transmission over cross coupled linear channels," submitted to *IEEE Trans. Commun.*

[7] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays.* New York: Wiley, 1980, p. 104.

[8] W. C. Jakes, Jr., *Microwave Mobile Communications.* New York: Wiley, 1974, p. 319.

[9] K. H. Awadalla, "Direction diversity in mobile communications," *IEEE Trans. Veh. Technol.*, vol. VT-30, pp. 121-123, Aug. 1981.

[10] R. T. Compton, Jr., "On the performance of a polarization sensitive adaptive array," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 718-725, Sept. 1981.

[11] D. C. Cox, "Antenna diversity performance in mitigating the effects of portable radiotelephone orientation and multipath propagation," *IEEE Trans. Commun.*, vol. COM-31, pp. 620-628, May 1983.

[12] W. Van Etten, "Maximum likelihood receiver for multiple channel transmission systems," *IEEE Trans. Commun.*, vol. COM-24, pp. 276-283, Feb. 1976.

[13] M. L. Steinberger, "Design of a terrestrial cross pol canceller," in *Proc. Int. Conf. Commun.*, Philadelphia, PA, June 1982, pp. 2B.6.1-2B.6.5.

[14] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968, p. 344.

Jack H. Winters (S'77-M'82) for a biography and photograph, see this issue, p. 805.