On the Capacity of OFDM-Based Spatial Multiplexing Systems

Helmut Bölcskei, *Member, IEEE*, David Gesbert, *Member, IEEE*, and Arogyaswami J. Paulraj, *Fellow, IEEE*

***Abstract*—This paper deals with the capacity behavior of wireless orthogonal frequency-division multiplexing (OFDM)-based *spatial multiplexing systems* in broad-band fading environments for the case where the channel is unknown at the transmitter and perfectly known at the receiver. Introducing a physically motivated multiple-input multiple-output (MIMO) broad-band fading channel model, we study the influence of physical parameters such as the amount of delay spread, cluster angle spread, and total angle spread, and system parameters such as the number of antennas and antenna spacing on ergodic capacity and outage capacity. We find that, in the MIMO case, unlike the single-input single-output (SISO) case, delay spread channels may provide advantages over flat fading channels not only in terms of outage capacity but also in terms of ergodic capacity. Therefore, MIMO delay spread channels will in general provide both higher diversity gain and higher multiplexing gain than MIMO flat-fading channels.**

***Index Terms*—Broad-band fading channels, diversity gain, ergodic capacity, MIMO, multiplexing gain, OFDM, outage capacity.**

I. Introduction and Outline

THE USE OF multiple antennas at both ends of a wireless link has recently been shown to have the potential of achieving extraordinary bit rates [1]–[4]. The corresponding technology is known as spatial multiplexing [1] or BLAST [2], [5] and allows an increase in bit rate in a wireless radio link without additional power or bandwidth consumption. So far, most of the research in this context has focused on the narrow-band flat-fading case. Extensive investigations on the capacity of narrow-band flat-fading (deterministic and stochastic) multiple-input multiple-output (MIMO) channels (assuming different levels of channel state information at the transmitter and the receiver) can be found in [2], [3], and [5]–[8].

1) Contributions: For a broad-band MIMO fading channel model, which is based on previous work reported in [9], [10], we provide expressions for the ergodic capacity and the outage capacity of orthogonal frequency-division multiplexing (OFDM)-based spatial multiplexing systems [4], [11] considering the case where the channel is unknown at the transmitter and perfectly known at the receiver. These expressions are then used to study (analytically and numerically) the influence of propagation parameters such as delay spread, cluster angle spread, and total angle spread, and system parameters such as the number of antennas and antenna spacing on capacity. We find that, in the MIMO case, unlike the single-input single-output (SISO) case, delay spread channels may provide an advantage over flat-fading channels not only in terms of outage capacity but also in terms of ergodic capacity. Consequently, MIMO delay spread channels provide not only higher diversity gain than MIMO flat-fading channels but also higher multiplexing gain.

2) Relation to Previous Work: Our channel model builds on research reported in [9] and [10]. In particular, it is an extension of the space–time channel model proposed in [9] to the case of multiple antennas at both the transmitter and the receiver. The capacity of deterministic MIMO channels with memory and full channel knowledge at the transmitter and the receiver was derived in [12]. In [13], the capacity of deterministic two-user multiaccess channels with memory is computed. Using a parametric MIMO channel model in which each path is described by an angle of departure, an angle of arrival, a (complex) path gain, and a path delay, the capacity of the corresponding deterministic MIMO delay spread channel (full channel knowledge at the transmitter and the receiver) has been provided in [4]. Using the same parametric channel model and defining the underlying parameters as random variables, a parametric MIMO fading channel model is established in [11], and an expression for the ergodic capacity is provided for the cases where the channel is either known at the receiver only or known at both the transmitter and the receiver. The channel model used in [4] and [11] does not capture the effects of spatial fading correlation, diffuse scattering, and scattering radius on capacity and is therefore fundamentally different from the MIMO fading channel model used in this paper. Furthermore, in the channel model used in [4] and [11], each path can only be a rank-1 contributor to capacity,[1] whereas in our model the rank depends on physically meaningful parameters such as cluster angle spread and antenna spacing. Our MIMO fading channel model is therefore more flexible than the one used in [4] and [11]. An interesting asymptotic (in the number of antenna elements) analysis for both the flat-fading and the frequency-selective fading cases appears in [14].

The use of OFDM in the context of spatial multiplexing has been proposed previously in [4], [11], [15], and [16]. However, it appears that no capacity studies of OFDM-based spatial multiplexing systems using the physically motivated MIMO fading

Paper approved by C. Tellambura, the Editor for Modulation and Signal Design of the IEEE Communications Society. Manuscript received October 12, 1999; revised September 18, 2000, and April 7, 2001. This work was supported in part by FWF under Grants J1629-TEC and J1868-TEC. This paper was presented in part at the IEEE ICASSP-00, Istanbul, Turkey, June 2000.

H. Bölcskei was with the Coordinated Science Laboratory and the Department of Electrical Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. He is now with the Communication Technology Laboratory, ETH Zurich, ETH Zentrum, ETF EIZZ, CH-8092 Zurich, Switzerland (e-mail: bolcskei@nari.ee.ethz.ch)).

D. Gesbert was with Iospan Wireless Inc., San Jose, CA 95134 USA. He is now with the Department of Informatics, University of Oslo, N-0316 Oslo Norway (e-mail: gesbert@ifi.uio.no).

A. J. Paulraj is with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305-9510 USA (e-mail: apaulraj@stanford.edu).

Publisher Item Identifier S 0090-6778(02)01368-5.

[1]This statement will be made more precise in Section II-B.

channel model provided in this paper have been performed so far. For the single-carrier narrow-band flat-fading case, the impact of spatial fading correlation and antenna array geometry on capacity has been studied in [7] and [8]. To the best of our knowledge, the impact of physical parameters (delay spread, cluster angle spread, and total angle spread) and system parameters (number of antennas and antenna spacing) on ergodic capacity and outage capacity in the broad-band OFDM case has not been studied in the literature so far.

3) Organization of the Paper: The rest of this paper is organized as follows. In Section II, we introduce our broad-band MIMO fading channel model. In Section III, we derive expressions for the ergodic capacity and the outage capacity of OFDM-based spatial multiplexing systems taking into account the new channel model. In Section IV, we study the influence of propagation parameters and system parameters on ergodic capacity and outage capacity. We furthermore demonstrate that, in the MIMO case, delay spread channels may provide advantage over flat-fading channels not only in terms of outage capacity but also in terms of ergodic capacity. In Section V, we provide numerical results complementing the analytical results in Section IV. Finally, Section VI provides our conclusions and some future research directions.

II. Broad-Band MIMO Fading Channel Model

In this section, we shall introduce a new model for broad-band MIMO fading channels based on a physical description of the propagation environment. Our channel model builds on previous work reported in [9] and [10].

A. General Assumptions

1) Propagation Scenario: We assume that the subscriber unit (SU) is surrounded by local scatterers so that fading at the SU antennas is spatially uncorrelated. The base transceiver station (BTS), however, is sufficiently high so that it is unobstructed and no local scattering occurs. Therefore, spatial fading at the BTS will be correlated with the exact correlation depending on the BTS antenna spacing and the angle spread observed at the BTS array [17]. Our model incorporates the power delay profile of the channel, but neglects shadowing. These assumptions on the propagation scenario are typical for cellular suburban deployments [17], where the BTS is on a tower or on the roof of a building and the terminal is on the street level and experiences local scattering. For the sake of simplicity, throughout the paper, we restrict our attention to the uplink case. The results for the downlink case are similar. In the following, M_T and M_R denote the number of transmit (i.e., SU) and receive (i.e., BTS) antennas, respectively.

2) Channel: Following [9] and [10], we model the delay spread by assuming that there are L significant scatterer clusters (see Fig. 1) and that each of the paths emanating from within the same scatterer cluster experiences the same delay. In practice, local scatterers in the cluster introduce micro delay variations, which will be neglected in our model. With $\mathbf{x}[n]$ denoting the discrete-time $M_T \times 1$ transmitted signal vector and $\mathbf{y}[n]$ the discrete-time $M_R \times 1$ received signal vector, respectively, we can write

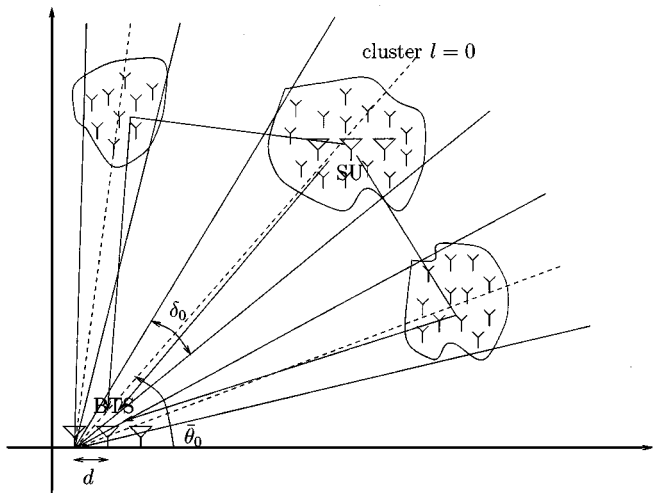


Fig. 1. Schematic representation of the MIMO delay spread channel composed of multiple clustered paths. Each path cluster has a mean angle of arrival $\bar{\theta}_l$ and an angle spread δ_l. The absolute antenna spacing is denoted by d.

$$\mathbf{y}[n] = \sum_{l=0}^{L-1} \mathbf{H}_l \mathbf{x}[n-l] \tag{1}$$

where the $M_R \times M_T$ complex-valued random matrix $\mathbf{H}_l$ represents the lth tap of the discrete-time MIMO fading channel impulse response. Note that in general there will be a continuum of delays. The channel model (1) is derived from the assumption of having L resolvable paths, where $L = \lfloor B\tau \rfloor$ with B and τ denoting the signal bandwidth and delay spread, respectively. The elements of the individual $\mathbf{H}_l$ are (possibly correlated) circularly symmetric complex Gaussian random variables.[2] Different scatterer clusters are uncorrelated, i.e.,[3]

$$\mathcal{E}\left[\mathrm{vec}\{\mathbf{H}_l\}\mathrm{vec}^H\{\mathbf{H}_{l'}\}\right] = \mathbf{0}_{M_R M_T} \quad \text{for } l \neq l' \tag{2}$$

where

$$\mathrm{vec}\{\mathbf{H}_l\} = \left[\mathbf{h}_{l,0}^T \quad \mathbf{h}_{l,1}^T \quad \ldots \quad \mathbf{h}_{l,M_T-1}^T\right]^T$$

with $\mathbf{h}_{l,k} = \left[h_{l,k}^{(0)} \quad h_{l,k}^{(1)} \quad \ldots \quad h_{l,k}^{(M_R-1)}\right]^T$ denoting the kth column of the matrix $\mathbf{H}_l$, and $\mathbf{0}_{M_R M_T}$ denoting the all-zero matrix of size $M_R M_T \times M_R M_T$. Each scatterer cluster has a mean angle of arrival at the BTS denoted as $\bar{\theta}_l$, a cluster angle spread δ_l (proportional to the scattering radius of the cluster), and a path gain σ_l^2 (derived from the power delay profile of the channel).

3) Array Geometry: For the sake of simplicity, we assume a uniform linear array (ULA) at both the BTS and the SU with identical antenna elements. Most of our results, can however, be extended to nonuniform arrays. The relative antenna spacing is denoted as $\Delta = d/\lambda$, where d is the absolute antenna spacing and $\lambda = c/f_c$ is the wavelength of a narrow-band signal with center frequency f_c.

4) Fading Statistics: We assume that the $\mathbf{h}_{l,k}$ ($l = 0, 1, \ldots, L-1$; $k = 0, 1, \ldots, M_T - 1$) have zero mean (i.e.,

[2] A circularly symmetric complex Gaussian random variable is a random variable $z = (x + jy) \sim \mathcal{CN}(0, \sigma^2)$, where x and y are i.i.d. $\mathcal{N}(0, \sigma^2/2)$.

[3] $\mathcal{E}$ denotes the expectation operator and the superscript ^{H} stands for conjugate transposition.

pure Rayleigh fading) and that the $M_R \times M_R$ correlation matrix $\mathbf{R}_l = \mathcal{E}\left\{\mathbf{h}_{l,k}\mathbf{h}_{l,k}^H\right\}$ is independent of k, or, equivalently, the fading statistics are the same for all transmit antennas. Defining $\rho_l(s\Delta, \bar{\theta}_l, \delta_l) = \mathcal{E}\left\{h_{l,k}^{(r)}\left(h_{l,k}^{(r+s)}\right)^*\right\}$ for $l = 0, 1, \ldots, L-1$, $k = 0, 1, \ldots, M_T - 1$ to be the fading correlation between two BTS antenna elements spaced $s\Delta$ wavelengths apart, the correlation matrix $\mathbf{R}_l$ can be written as

$$[\mathbf{R}_l]_{m,n} = \sigma_l^2 \rho_l((n-m)\Delta, \bar{\theta}_l, \delta_l). \tag{3}$$

Note that we have absorbed the power delay profile of the channel into the correlation matrices.

Factoring the $M_R \times M_R$ correlation matrix $\mathbf{R}_l$ according to $\mathbf{R}_l = \mathbf{R}_l^{1/2}\mathbf{R}_l^{1/2}$, where $\mathbf{R}_l^{1/2}$ is of size $M_R \times M_R$, the $M_R \times M_T$ matrices $\mathbf{H}_l$ can be written as

$$\mathbf{H}_l = \mathbf{R}_l^{1/2}\mathbf{H}_{w,l}, \quad l = 0, 1, \ldots, L-1 \tag{4}$$

with the $\mathbf{H}_{w,l}$ being uncorrelated $M_R \times M_T$ matrices with i.i.d. $\mathcal{CN}(0,1)$ entries. We have therefore decomposed the lth tap of the stochastic MIMO channel impulse response into the product of a deterministic matrix $\mathbf{R}_l^{1/2}$ taking into account the spatial fading correlation at the BTS and a stochastic matrix of i.i.d. complex Gaussian random variables.

Let us next assume that the angle of arrival for the lth ($l = 0, 1, \ldots, L-1$) path cluster at the BTS is Gaussian distributed around the mean angle of arrival $\bar{\theta}_l$, i.e., the actual angle of arrival is given by $\theta_l = \bar{\theta}_l + \hat{\theta}_l$ with $\hat{\theta}_l \sim \mathcal{N}(0, \sigma_{\theta_l}^2)$. The variance $\sigma_{\theta_l}^2$ is proportional to the angular spread δ_l and hence the scattering radius of the lth path cluster. It is shown in [9] that for small angular spread the correlation function can be approximated as

$$\rho_l(s\Delta, \bar{\theta}_l, \delta_l) \approx e^{-j2\pi s\Delta\cos(\bar{\theta}_l)} e^{-(1/2)(2\pi s\Delta \sin(\bar{\theta}_l)\sigma_{\theta_l})^2}. \tag{5}$$

Although this approximation is accurate only for small angular spread, it does provide the correct trend for large angular spread, namely uncorrelated spatial fading. Note that in the case $\sigma_{\theta_l} = 0$, the correlation matrix $\mathbf{R}_l$ collapses to a rank-1 matrix and can be written as $\mathbf{R}_l = \sigma_l^2 \mathbf{a}(\bar{\theta}_l)\mathbf{a}^H(\bar{\theta}_l)$ with the array response vector of the ULA given by

$$\mathbf{a}(\theta) = \begin{bmatrix} 1 & e^{j2\pi\Delta\cos(\theta)} & \ldots & e^{j2\pi(M_R-1)\Delta\cos(\theta)} \end{bmatrix}^T. \tag{6}$$

B. Differences to the Parametric Fading Channel Model

In the parametric fading channel model proposed in [11], each tap $\mathbf{H}_l$ can be written as

$$\mathbf{H}_l = \beta_l \mathbf{a}_R(\bar{\theta}_{R,l})\mathbf{a}_D^T(\bar{\theta}_{D,l}), \qquad l = 0, 1, \ldots, L-1$$

where β_l denotes the complex Gaussian distributed path gain, $\bar{\theta}_{R,l}$ and $\bar{\theta}_{D,l}$ are the random angle-of-arrival and angle-of-departure, respectively, of the lth path, and $\mathbf{a}_R(\theta)$ and $\mathbf{a}_D(\theta)$ the $M_R \times 1$ and $M_T \times 1$ receive and transmit array response vectors (cf. (6)), respectively. Note that in this model every realization of $\mathbf{H}_l$ has rank 1. In our MIMO fading channel model, the rank of the matrices $\mathbf{H}_l$ is controled by the fading correlation at the BTS. If the angular spread of the lth path cluster is large, $\mathbf{H}_l$ will have high rank; for decreasing angular spread the rank of $\mathbf{H}_l$ will decrease. This follows from (4) and the fact that the correlation matrix $\mathbf{R}_l$ loses rank if the angular spread decreases. The MIMO fading channel model proposed in this paper is therefore more flexible than the parametric fading channel model [11] and seems to be a more adequate description of a real-world scattering environment.

III. Mutual Information and Capacity of OFDM-Based Spatial Multiplexing Systems

In this section, we derive an expression for the mutual information of OFDM-based spatial multiplexing systems. This expression is then used to compute the ergodic capacity and study the outage properties of the system.

A. OFDM-Based Spatial Multiplexing

Spatial multiplexing [1], also refered to as BLAST [2], [5], has the potential to drastically increase the capacity of wireless radio links with no additional power or bandwidth consumption. The technology requires multiple antennas at both ends of the wireless link. The gain in terms of ergodic capacity over SISO systems resulting from the use of multiple antennas is termed *multiplexing gain*. The main reason for using OFDM in this context is the fact that OFDM modulation turns a frequency-selective MIMO fading channel into a set of parallel frequency-flat MIMO fading channels. This renders multichannel equalization particularly simple, since for each OFDM tone a narrow-band receiver can be employed [4], [11]. In OFDM-based spatial multiplexing, the (possibly coded) data streams are first passed through OFDM modulators and then launched from the individual antennas. Note that this transmission takes place simultaneously from all M_T transmit antennas. In the receiver, the individual signals are passed through OFDM demodulators, separated, and then decoded. Fig. 2 shows a schematic of an OFDM-based spatial multiplexing system. Throughout the paper, we assume that the length of the cyclic prefix (CP) in the OFDM system is greater than the length of the discrete-time baseband channel impulse response. This assumption guarantees that the frequency-selective fading channel indeed decouples into a set of parallel frequency-flat fading channels [18].

Organizing the transmitted data symbols into frequency vectors $\mathbf{c}_k = \begin{bmatrix} c_k^{(0)} & c_k^{(1)} & \ldots & c_k^{(M_T-1)} \end{bmatrix}^T$ with $c_k^{(i)}$ denoting the data symbol transmitted from the ith antenna on the kth ($k = 0, 1, \ldots, N-1$) tone and defining $\mathbf{H}\left(e^{j2\pi\theta}\right) = \sum_{l=0}^{L-1} \mathbf{H}_l e^{-j2\pi l\theta}$ ($0 \le \theta < 1$), it can be shown that

$$\hat{\mathbf{c}}_k = \mathbf{H}\left(e^{j2\pi(k/N)}\right)\mathbf{c}_k + \mathbf{n}_k, \quad k = 0, 1, \ldots, N-1 \tag{7}$$

where $\hat{\mathbf{c}}_k$ denotes the reconstructed data vector for the kth tone, and $\mathbf{n}_k$ is additive white Gaussian noise (AWGN) satisfying

$$\mathcal{E}\left\{\mathbf{n}_k\mathbf{n}_l^H\right\} = \sigma_n^2 \mathbf{I}_{M_R}\delta[k-l] \tag{8}$$

where $\mathbf{I}_{M_R}$ is the identity matrix of size M_R. From (7), it can be seen that equalization requires application of a narrow-band receiver for each tone $k = 0, 1, \ldots, N-1$.

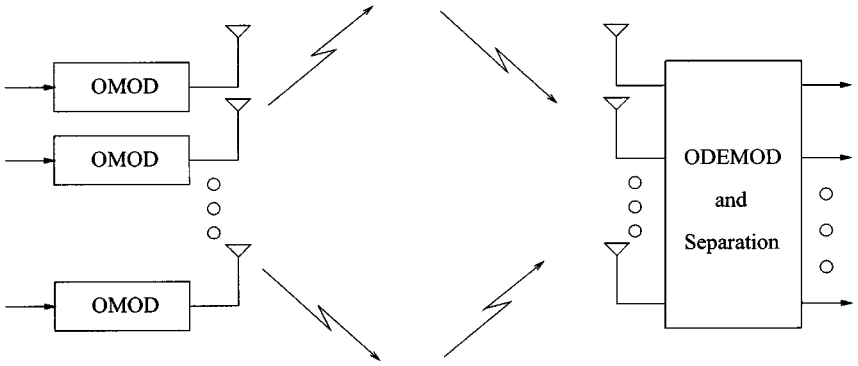


Fig. 2. OFDM-based spatial multiplexing system (OMOD and ODEMOD denote an OFDM-modulator and demodulator, respectively).

B. Mutual Information

We start by stacking the vectors $\hat{\mathbf{c}}_k$, $\mathbf{c}_k$, and $\mathbf{n}_k$ according to

$$\begin{aligned}\hat{\mathbf{c}} &= [\,\hat{\mathbf{c}}_0^T \quad \hat{\mathbf{c}}_1^T \quad \dots \quad \hat{\mathbf{c}}_{N-1}^T\,]^T \\ \mathbf{c} &= [\,\mathbf{c}_0^T \quad \mathbf{c}_1^T \quad \dots \quad \mathbf{c}_{N-1}^T\,]^T \\ \mathbf{n} &= [\,\mathbf{n}_0^T \quad \mathbf{n}_1^T \quad \dots \quad \mathbf{n}_{N-1}^T\,]^T\end{aligned}$$

where $\hat{\mathbf{c}}$ and $\mathbf{n}$ are $M_R N \times 1$ vectors and $\mathbf{c}$ is an $M_T N \times 1$ vector. Note that (8) implies that the noise vector $\mathbf{n}$ is white, i.e.,

$$\mathcal{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2 \mathbf{I}_{M_R N}.$$

We furthermore define the $NM_R \times NM_T$ block-diagonal matrix

$$\mathbf{H} = \operatorname{diag}\left\{\mathbf{H}\left(e^{j2\pi(k/N)}\right)\right\}_{k=0}^{N-1}.$$

With these definitions, (7) can be rewritten as

$$\hat{\mathbf{c}} = \mathbf{H}\mathbf{c} + \mathbf{n}. \tag{9}$$

In the following we assume that for each channel use (corresponding to at least one OFDM symbol) an independent realization of the random channel impulse response matrices $\mathbf{H}_l$ is drawn and that the channel remains constant within one channel use. Using (9), the mutual information (in b/s/Hz) of the OFDM-based spatial multiplexing system under an average transmitter power constraint is given by[4] [19], [20]

$$I = \frac{1}{N}\log\left[\det\left(\mathbf{I}_{M_R N} + \frac{1}{\sigma_n^2}\mathbf{H}\boldsymbol{\Sigma}\mathbf{H}^H\right)\right] \tag{10}$$

where $\boldsymbol{\Sigma}$ with[5] $\operatorname{Tr}(\boldsymbol{\Sigma}) \leq P$ is the covariance matrix of the Gaussian input vector $\mathbf{c}$ and P is the maximum overall transmit power. Note that mutual information is normalized by N, since N data symbols are transmitted in one OFDM symbol and that we ignored the loss in spectral efficiency due to the presence of the CP. The $NM_T \times NM_T$ matrix $\boldsymbol{\Sigma}$ is a block-diagonal matrix given by

$$\boldsymbol{\Sigma} = \operatorname{diag}\{\boldsymbol{\Sigma}_k\}_{k=0}^{N-1}$$

where the $M_T \times M_T$ matrices $\boldsymbol{\Sigma}_k$ are the covariance matrices of the Gaussian vectors $\mathbf{c}_k$, and as such determine the power allocation across the transmit antennas and across the OFDM tones. If the channel is perfectly known at the transmitter, the optimum power allocation is obtained by distributing the total available power P according to the water-filling solution [4]. In OFDM-based spatial multiplexing systems, statistically independent data symbols are transmitted from different antennas and different tones and the total available power is allocated uniformly across all space–frequency subchannels [4], [11]. In the following, we set $\boldsymbol{\Sigma}_k = (P/(M_T N))\mathbf{I}_{M_T}$ $(k = 0, 1, \dots, N-1)$, which is easily verified to result in $\operatorname{Tr}(\boldsymbol{\Sigma}) = P$. Using (10), we therefore obtain

$$\begin{aligned}I &= \frac{1}{N}\sum_{k=0}^{N-1} I_k \\ &= \frac{1}{N}\sum_{k=0}^{N-1}\log\left[\det\left(\mathbf{I}_{M_R} + \rho\mathbf{H}(e^{j2\pi(k/N)})\mathbf{H}^H(e^{j2\pi(k/N)})\right)\right]\end{aligned} \tag{11}$$

where $\rho = P/(M_T N\sigma_n^2)$. The quantity I_k is the mutual information of the kth MIMO OFDM subchannel. Note that, since $\mathbf{H}\left(e^{j2\pi(k/N)}\right)$ is random, I_k is a random entity as well. We shall next show that the distribution of I_k is independent of k and hence all the I_k $(k = 0, 1, \dots, N-1)$ have the same distribution. In the following, the notation $x \sim y$ means that the distribution of the random variable x is equal to the distribution of the random variable y.

Proposition 1: The distribution of I_k $(k = 0, 1, \dots, N-1)$ is independent of k and given by

$$I_k \sim \log\left[\det\left(\mathbf{I}_{M_R} + \rho\boldsymbol{\Lambda}\mathbf{H}_w\mathbf{H}_w^H\right)\right], \qquad \text{for } k = 0, 1, \dots, N-1 \tag{12}$$

where $\boldsymbol{\Lambda} = \operatorname{diag}\{\lambda_i(\mathbf{R})\}_{i=0}^{M_R-1}$, $\mathbf{H}_w$ is an $M_R \times M_T$ i.i.d. random matrix with $\mathcal{CN}(0,1)$ entries, and $\mathbf{R} = \sum_{l=0}^{L-1}\mathbf{R}_l$. Finally, $\lambda_i(\mathbf{R})$ denotes the ith eigenvalue of $\mathbf{R}$.

Proof: Gaussianity of the $\mathbf{H}_l$ implies Gaussianity of $\mathbf{H}\left(e^{j2\pi(k/N)}\right)$ for $k = 0, 1, \dots, N-1$. Now, using

$$\mathbf{H}\left(e^{j2\pi(k/N)}\right) = \sum_{l=0}^{L-1}\mathbf{R}_l^{1/2}\mathbf{H}_{w,l}e^{-j2\pi(k/N)l} \tag{13}$$

[4]Throughout the paper, all logarithms are to the base 2.

[5]$\operatorname{Tr}(\mathbf{A})$ stands for the trace of the matrix $\mathbf{A}$.

it follows that the columns of $\mathbf{H}\left(e^{j2\pi(k/N)}\right)$ $(k = 0,1,\ldots,N-1)$ are uncorrelated and have the same statistics. Denoting the first column of $\mathbf{H}\left(e^{j2\pi(k/N)}\right)$ as $\mathbf{h}\left(e^{j2\pi(k/N)}\right)$ and the first column of $\mathbf{H}_{w,l}$ as $\mathbf{h}_{w,l}$ and using (2) it follows that

$$\begin{aligned}&\mathcal{E}\left\{\mathbf{h}\left(e^{j2\pi(k/N)}\right)\mathbf{h}^H\left(e^{j2\pi(k/N)}\right)\right\}\\&\quad=\mathcal{E}\left\{\sum_{l=0}^{L-1}\mathbf{R}_l^{1/2}\mathbf{h}_{w,l}e^{-j2\pi(k/N)l}\sum_{l'=0}^{L-1}\mathbf{h}_{w,l'}^H\mathbf{R}_{l'}^{1/2}e^{j2\pi(k/N)l'}\right\}\\&\quad=\sum_{l=0}^{L-1}\mathbf{R}_l^{1/2}\mathbf{R}_l^{1/2}=\mathbf{R}.\end{aligned}$$

Note that the correlation matrix is independent of k. We have thus shown that

$$\mathbf{H}\left(e^{j2\pi(k/N)}\right)\sim\mathbf{R}^{1/2}\mathbf{H}_w$$

with $\mathbf{H}_w$ denoting an i.i.d. matrix with $\mathcal{CN}(0,1)$ entries. Hence, it follows that

$$I_k\sim\log\left[\det\left(\mathbf{I}_{M_R}+\rho\mathbf{R}\mathbf{H}_w\mathbf{H}_w^H\right)\right],\quad k=0,1,\ldots,N-1$$

which can be rewritten as

$$I_k\sim\log\left[\det\left(\mathbf{I}_{M_R}+\rho\mathbf{U}\mathbf{\Lambda}\mathbf{U}\mathbf{H}_w\mathbf{H}_w^H\right)\right].$$

Using the fact that $\mathbf{U}\mathbf{H}_w\sim\mathbf{H}_w$[21], we obtain

$$I_k\sim\log\left[\det\left(\mathbf{I}_{M_R}+\rho\mathbf{U}\mathbf{\Lambda}\mathbf{H}_w\mathbf{H}_w^H\right)\right].$$

Now, without changing the distribution, we can right-multiply $\mathbf{H}_w^H$ by $\mathbf{U}^H$ to obtain

$$I_k\sim\log\left[\det\left(\mathbf{I}_{M_R}+\rho\mathbf{U}\mathbf{\Lambda}\mathbf{H}_w\mathbf{H}_w^H\mathbf{U}^H\right)\right].$$

Using $\det(\mathbf{I}+\mathbf{X}\mathbf{Y})=\det(\mathbf{I}+\mathbf{Y}\mathbf{X})$ and exploiting the unitarity of $\mathbf{U}$, we finally get

$$I_k\sim\log\left[\det\left(\mathbf{I}_{M_R}+\rho\mathbf{\Lambda}\mathbf{H}_w\mathbf{H}_w^H\right)\right]$$

which concludes the proof. □

C. Ergodic Capacity and Outage Capacity

We shall next establish the information-theoretic value of the results derived in Section III-B. Two scenarios are considered, the *ergodic* and the *nonergodic* case. In both cases, we assume that the channel remains fixed within one channel use (at least one OFDM symbol) and then changes in an independent fashion to a new realization.

1) Ergodic Case: The basic assumption here is that the transmission time is long enough to reveal the long-term ergodic properties of the fading channel. In this case, a Shannon capacity exists and is given by $C=\mathcal{E}\{I\}$ with I defined in (11). At rates lower than C, the error probability (for a good code) decays exponentially with the transmission length. The assumption here is that the fading process is ergodic, coding and interleaving are performed across OFDM symbols, and that the number of fading blocks spanned by a codeword goes to infinity whereas the block size (which equals the number of tones in the OFDM system multiplied by the number of OFDM symbols spanning one channel use) remains constant (and finite). Capacity can be achieved in principle by transmitting a codeword over a very large number of independently fading blocks. We furthermore note that the capacity obtained for an OFDM-based spatial multiplexing system is a lower bound for the capacity of the underlying broad-band MIMO fading channel.

2) Nonergodic Case: In this case, we assume that a codeword spans an arbitrary but fixed number of blocks while the block size goes to infinity. This situation typically occurs when stringent delay constraints are imposed, as is the case, for example, in speech transmission over wireless channels. These assumptions give rise to error probabilities which do not decay with an increase of block length. A capacity in the Shannon sense does not exist since, with nonzero probability, which is independent of the code length, the mutual information I in (11) falls below any positive rate, as small as it may be. Thus, the concept of capacity versus outage [22], [23] has to be invoked. Assuming that codewords extend over a single block, the outage (or failure) probability for a given rate is the probability that I falls below that rate. In this case, capacity is viewed as a random entity [22], [23] since it depends on the instantaneous random channel parameters.

IV. Influence of Channel and System Parameters on Capacity

In this section, we study the influence of channel and system parameters on ergodic capacity and outage capacity. In particular, we demonstrate that in the MIMO case, unlike the SISO case, delay spread channels may provide advantage over flat-fading channels in terms of ergodic capacity. While the ergodic case is to some extent amenable to analytic studies, the nonergodic case will mainly be discussed by means of simulation results in Section V. Analytic results seem hard to obtain in the nonergodic case. Some statements of qualitative nature on the nonergodic case will be made in this section.

A. The Ergodic Case

The ergodic capacity is obtained from (11) as

$$C=\mathcal{E}\left\{\frac{1}{N}\sum_{k=0}^{N-1}I_k\right\}.$$

Now, using Proposition 1, which says that the I_k $(k = 0,1,\ldots,N-1)$ all have the same distribution given by (12), the ergodic capacity is obtained as

$$C=\mathcal{E}\left\{\log\left[\det\left(\mathbf{I}_{M_R}+\rho\mathbf{\Lambda}\mathbf{H}_w\mathbf{H}_w^H\right)\right]\right\}\tag{14}$$

where expectation is taken with respect to $\mathbf{H}_w$. A semi-analytic result for this expectation has been provided by Telatar in [3] for the case where $\mathbf{\Lambda}=\mathbf{I}$. In the general case, the evaluation of the expectation in (14) requires the concept of zonal polynomials [21] and is significantly more complicated. We shall therefore resort to a simple asymptotic analysis by assuming that M_T is large. It follows from the law of large numbers that, for fixed

M_R as M_T gets large, $(1/M_T)\mathbf{H}_w\mathbf{H}_w^H \to \mathbf{I}_{M_R}$. Hence, in the large M_T limit, we get

$$C = \log\left[\det\left(\mathbf{I}_{M_R} + \bar{\rho}\mathbf{\Lambda}\right)\right] \tag{15}$$

where $\bar{\rho} = M_T\rho = (P/(N\sigma_n^2))$. In the low SNR regime, i.e., for small $\bar{\rho}$, it follows from (15) that in the large M_T limit

$$C = \log\left(\prod_{i=0}^{M_R-1}\left(1+\bar{\rho}\lambda_i(\mathbf{R})\right)\right) \approx \log(1+\bar{\rho}M_T\mathrm{Tr}(\mathbf{R}))$$

where all the higher order terms in $\bar{\rho}$ have been neglected. Thus, in the low SNR regime, the ergodic capacity is driven by the Trace of the sum correlation matrix $\mathbf{R}$. Here, comparing channels on the basis of fixed energy, i.e., fixed $\mathrm{Tr}(\mathbf{R})$ leads to the conclusion that delay spread has no impact on ergodic capacity.

In the high- SNR case, we obtain

$$C = \sum_{i=0}^{M_R-1}\log(1+\bar{\rho}\lambda_i(\mathbf{R})). \tag{16}$$

The eigenvalue spread of the sum correlation matrix $\mathbf{R} = \sum_{l=0}^{L-1}\mathbf{R}_l$ therefore critically determines ergodic capacity. In fact, we have

Lemma 2: For $\mathrm{Tr}(\mathbf{R}) = 1$, the right-hand side (RHS) in (16) is maximized for $\lambda_i(\mathbf{R}) = (1/M_R)$ $(i = 0, 1, \ldots, M_R - 1)$.

Proof: The proof follows easily by applying Jensen's inequality to the RHS of (16). □

Using the developments in Section IV-A and [3, Theorem 1], it can even be shown that, for $\mathrm{Tr}(\mathbf{R}) = 1$, $\lambda_i(\mathbf{R}) = (1/M_R)$ $(i = 0, 1, \ldots, M_R - 1)$ maximizes the exact (finite M_T) expression (14). A deviation of $\lambda_i(\mathbf{R})$ as a function of i from a constant function will therefore result in a loss in terms of ergodic capacity or equivalently reduced multiplexing gain. In the following, we restrict our attention to the high SNR case. We shall next show how the propagation and system parameters impact the eigenvalues of $\mathbf{R}$ and hence ergodic capacity. Since the individual correlation matrices $\mathbf{R}_l$ are Toeplitz, the sum correlation matrix $\mathbf{R}$ is Toeplitz as well. We can thus invoke Szegö's theorem [24] to obtain the limiting $(M_R \to \infty)$ distribution[6] of the eigenvalues of $\mathbf{R}$ as

$$\begin{aligned}\lambda(\nu) &= \sum_{l=0}^{L-1}\lambda_l(\nu)\\ &= \sum_{l=0}^{L-1}\sum_{s=-\infty}^{\infty}\rho_l(s\Delta,\bar{\theta}_l,\delta_l)e^{-j2\pi s\nu}, \qquad 0 \le \nu < 1.\end{aligned}$$

Using (5), we obtain

$$\lambda(\nu) = \sum_{l=0}^{L-1}\vartheta_3\left(\pi(\nu-\Delta\cos(\bar{\theta}_l)), e^{-(1/2)(2\pi\Delta\sin(\bar{\theta}_l)\sigma_{\theta,l})^2}\right) \tag{17}$$

with the third-order theta function given by [25] $\vartheta_3(\nu, q) = \sum_{n=-\infty}^{\infty} q^{n^2}e^{2jn\nu}$. Although this expression yields the exact eigenvalue distribution only in the limiting case $M_R \to \infty$, in the case of finite M_R good approximations of the eigenvalues can be obtained by sampling $\lambda(\nu)$ uniformly on the unit circle [24], which allows us to assume that the eigenvalue distribution in the finite case follows the distribution given by $\lambda(\nu)$. We are now able to study the impact of various propagation and system parameters on the eigenvalue distribution of $\mathbf{R}$ and hence the ergodic capacity.

[6]Note that for $M_R \to \infty$ the eigenvalues of $\mathbf{R}$ are characterized by a periodic continuous function [24]. Thus, whenever we use the term eigenvalue distribution, we actually refer to this function.

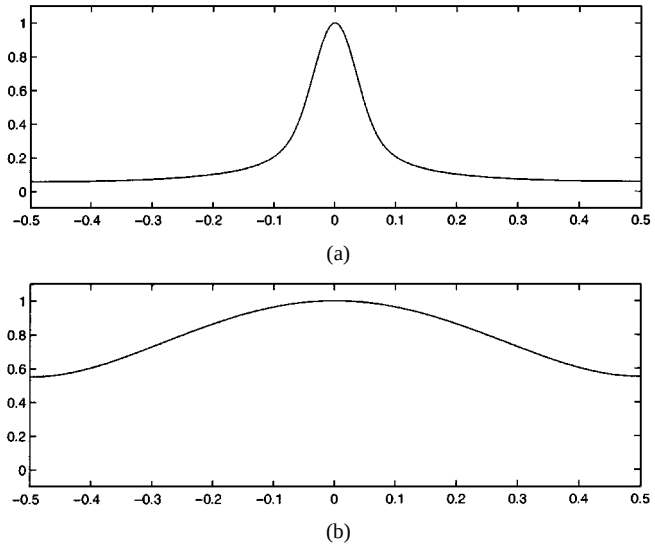


Fig. 3. Limiting eigenvalue distribution of the correlation matrix $\mathbf{R}_0$ for the cases of (a) high spatial fading correlation and (b) low spatial fading correlation.

1) Impact of Cluster Angle Spread and Antenna Spacing: Let us first investigate the influence of cluster angle spread and antenna spacing on ergodic capacity. For the sake of simplicity, take one path only and its associated correlation matrix $\mathbf{R}_0$. The limiting eigenvalue distribution of $\mathbf{R}_0$ is given by $\lambda_0(\nu) = \vartheta_3\left(\pi(\nu-\Delta\cos(\bar{\theta}_0)), e^{-(1/2)(2\pi\Delta\sin(\bar{\theta}_0)\sigma_{\theta,0})^2}\right)$. Now, noting that the correlation function $\rho_0(s\Delta,\bar{\theta}_0,\delta_0)$ as a function of s is essentially a modulated Gaussian function with its spread increasing for increasing antenna spacing and/or increasing cluster angle spread and vice versa, it follows that $\lambda_0(\nu)$ will be more flat in the case of large antenna spacing and/or large cluster angle spread (i.e., low spatial fading correlation). For small antenna spacing and/or small cluster angle spread, $\lambda_0(\nu)$ will be peaky. Fig. 3(a) and (b) show the limiting eigenvalue distribution of $\mathbf{R}_0$ for high and low spatial fading correlation, respectively. From our previous discussion, it thus follows that the ergodic capacity will decrease for increasing concentration of $\lambda_0(\nu)$ and vice versa.

2) Impact of Total Angle Spread: We shall next study the impact of total angle spread on ergodic capacity. Assume that either the individual scatterer cluster angle spreads are small or that antenna spacing at the BTS is small or both. Hence, the individual $\lambda_l(\nu)$ are peaky. Now, from (17), we can see that the limiting distribution $\lambda(\nu)$ is obtained by adding the individual limiting distributions $\lambda_l(\nu)$. Note furthermore that $\lambda_l(\nu)$ is essentially a Gaussian centered around $\Delta\cos(\bar{\theta}_l)$. Now, if the total angle spread, i.e., the spread of the $\bar{\theta}_l$, is large, the sum-limiting distribution $\lambda(\nu)$ can still be flat even though the individual $\lambda_l(\nu)$ are peaky. For a given small cluster angle spread, Fig. 4(a) and (b) show example limiting distributions

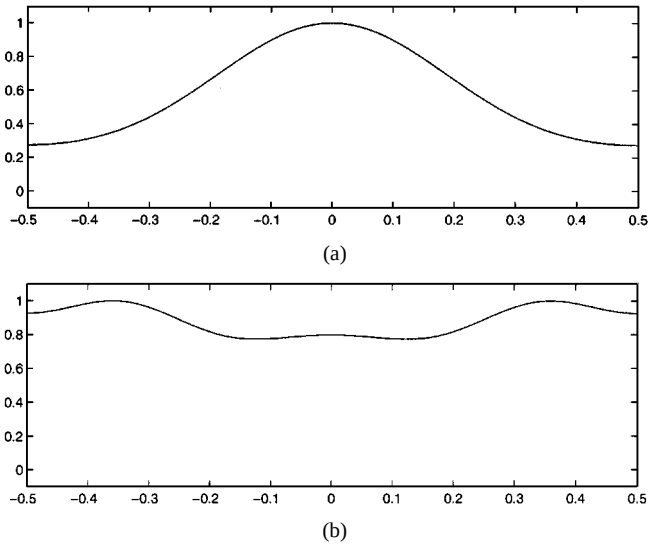


Fig. 4. Limiting eigenvalue distribution of the sum correlation matrix $\mathbf{R} = \sum_{l=0}^{L-1} \mathbf{R}_l$ for fixed cluster angle spread and for the cases of (a) small total angle spread and (b) large total total angle spread.

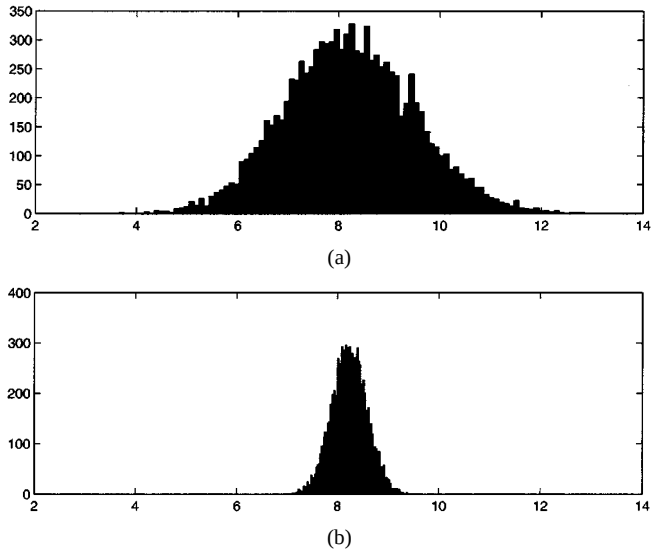


Fig. 5. Example histograms of the mutual information I in b/s/Hz in the (a) flat-fading case and (b) the high delay-spread case.

for a three-path channel with a total angle spread of 22.5 degrees and a total angle spread of 90 degrees, respectively. We can clearly see the impact of total angle spread on the limiting eigenvalue distribution $\lambda(\nu)$ and hence on ergodic capacity. Large total angle spread renders $\lambda(\nu)$ flat and therefore increases ergodic capacity, whereas small total angle spread makes $\lambda(\nu)$ peaky and hence reduces ergodic capacity. This impact can further be illustrated by studying the extreme case $\sigma_{\theta,l} \approx 0$, i.e., small cluster angle spread (or equivalently large distance between BTS and SU). In this case, the sum correlation matrix is given by

$$\mathbf{R} = \sum_{l=0}^{L-1} \sigma_l^2 \mathbf{a}(\bar{\theta}_l)\mathbf{a}^H(\bar{\theta}_l)$$

with $\mathbf{a}(\theta)$ defined in (6). Take the simple example $L = 2$, $\Delta = 1/2$, and $M_R = 2$. In this case, for $\bar{\theta}_0 = 0$ and $\bar{\theta}_1 = \pi/4$, we get $\sigma(\mathbf{R}) = \{0.21, 3.79\}$ and for $\bar{\theta}_0 = 0$ and $\bar{\theta}_1 = \pi/2$ we have $\sigma(\mathbf{R}) = \{2, 2\}$. For $\rho = 10$, the ergodic capacities obtained by Monte Carlo evaluation of (14) are $C = 7.3$ b/s/Hz in the case of small total angle spread and $C = 8.87$ b/s/Hz in the case of large total angle spread.

3) Ergodic Capacity in the SISO and in the MIMO Case: It is well known that in the SISO case delay spread channels do not offer advantage over flat-fading channels in terms of ergodic capacity [22], [23]. This can easily be seen from (14) by noting that in the SISO case $\sum_{l=0}^{L-1} \mathbf{R}_l = \sum_{l=0}^{L-1} \sigma_l^2$ and hence ergodic capacity is only a function of the total energy in the channel. In the MIMO case, the situation is in general different. Fix $\mathrm{Tr}(\mathbf{R})$, and take a flat-fading scenario with small cluster angle spread where $\mathbf{R} = \mathbf{R}_0$ has rank 1. In this case, the matrix $\mathbf{\Lambda H}_w \mathbf{H}_w^H$ has rank 1 with probability one and hence only one spatial data pipe can be opened up, or equivalently there is no multiplexing gain. Now, compare this scenario to a delay-spread scenario where $L \geq M_R$ and each of the $\mathbf{R}_l$ $(l = 0, 1, \ldots, L-1)$ has rank 1 but the sum-correlation matrix $\mathbf{R}$ has full rank. For this to happen, a sufficiently large total angle spread is necessary. Clearly, in this case, M_R spatial data pipes can be opened up and we will get a higher ergodic capacity because the rank of $\mathbf{R}$ is higher than in the flat-fading case. We note that in the case where all the correlation matrices satisfy $\mathbf{R}_l = \sigma_l^2 \mathbf{I}_{M_R}$ $(l = 0, 1, \ldots, L-1)$ this effect does not occur. However, since this scenario corresponds to fully uncorrelated spatial fading it is very unlikely. We can therefore conclude that in practice MIMO delay spread channels offer advantage over MIMO flat-fading channels in terms of ergodic capacity. We caution the reader that this conclusion is a result of the assumption that delayed paths increase the total angle spread. This assumption has been verified by measurement for outdoor MIMO broad-band channels in the 2.5-GHz band [26].

B. The Nonergodic Case

In [22], it has been demonstrated that SISO delay-spread channels offer significant advantage over flat-fading channels in terms of outage probabilities or outage capacity. The outage properties are determined by the number of diversity degrees of freedom in the channel. In our case, we have both spatial diversity and frequency diversity available. We can therefore expect that both diversity sources will contribute to the outage characteristics of the system. Assuming that a codeword spans one block, we recall that the outage probability for a given rate R is the probability that $I = (1/N)\sum_{k=0}^{N-1} I_k$ falls below that rate. The distribution of I is hard to compute analytically[7]. We

[7]In this context, we would like to point out an error in [27] and simplifications of some of the results reported in [27], in which the D_k defined in [27, eq. (9)] should read $C \sim \frac{1}{N}\sum_{k=0}^{N-1} \log\left[\det(\mathbf{I}_{M_B} + \rho \mathbf{D} V_k \mathbf{H}\mathbf{H}^H V_k^H \mathbf{D}^H)\right]$ and [27, eq. (11)] should be replaced by $\mathcal{E}\{C\} = \mathcal{E}\left\{(1/N)\sum_{k=0}^{N-1} \log\left[\det\left(\mathbf{I}_{M_B} + \rho \mathbf{\Lambda}_k \mathbf{H}_W \mathbf{H}_W^H \mathbf{\Lambda}_k^H\right)\right]\right\}$. Furthermore, it follows from Proposition 1 in this paper that [27eq. (11)] can be simplified to yield $\mathcal{E}\{C\} = \mathcal{E}\left\{\log\left[\det\left(\mathbf{I}_{M_B} + \rho \mathbf{\Lambda} \mathbf{H}_w \mathbf{H}_w^H\right)\right]\right\}$ with $\mathbf{\Lambda}$ defined in Proposition 1 in this paper. It can also be shown that the $|\lambda_{k,l}|^2$ in [27] are independent of k and equal the $\lambda_l(\mathbf{R})$ defined in Proposition 1 in this paper, and that the r_k in [27] are independent of k and equal $\mathrm{rank}(\mathbf{R})$ with $\mathbf{R} = \sum_{l=0}^{L-1} \mathbf{R}_l$. With this, [27, eq. (12)] can be simplified to yield $\mathcal{E}\{C\} \leq \sum_{l=0}^{r-1} \log\left(1 + \rho M_S \lambda_l(\mathbf{R})\right)$. Furthermore, [27, eq. (13)] should be replaced by $\mathcal{E}\{C\} \geq \mathcal{E}\left\{\sum_{l=0}^{r-1} \log(1 + \rho \lambda_l(\mathbf{R})|r_{l,l}|^2)\right\}$ and the last equation in [27] has to be replaced by $\mathcal{E}\{C\} \leq \mathcal{E}\left\{\sum_{l=0}^{r-1} \log\left[1 + \rho\left(\lambda_l(\mathbf{R})|r_{l,l}|^2 + \sum_{m=l+1}^{r-1} \lambda_m(\mathbf{R})|r_{l,m}|^2\right)\right]\right\}$.

therefore resort to numerical studies presented in Section V and make a few qualitative statements below.

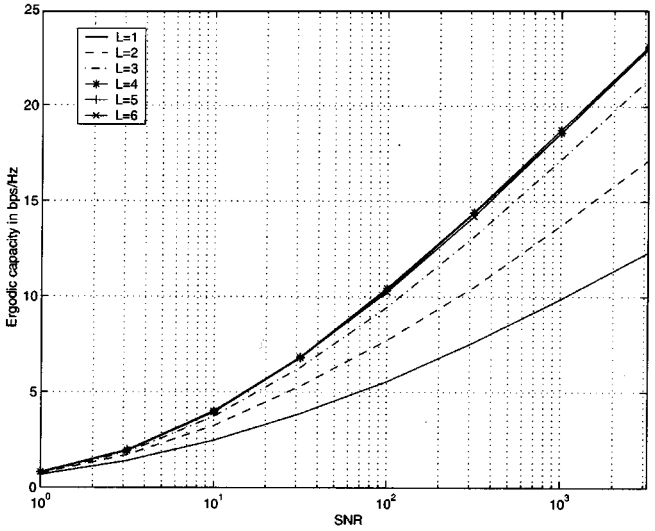
The individual I_k all have the same distribution. The correlation between the I_k, however, strongly depends on the amount of delay spread in the system. In order to establish the value of space–frequency diversity in terms of outage properties, let us consider the two limiting cases of flat-fading (i.e., no frequency diversity) and high delay spread. In the high delay spread case, we assume that the correlation between the I_k is small, which corresponds to the assumption of a high number of independently fading taps in the channel. The mean of I is independent of the correlation between the I_k and is given by (14). The variance of I, however, and hence the outage properties depend significantly on the amount of space–frequency diversity. Denote the variance of I_k as σ_I^2 (recall that the distribution of I_k is independent of k). In the flat-fading case, we have $\mathrm{var}\{I\} = \sigma_I^2$, whereas in the high delay-spread case (under the idealistic assumption of full decorrelation of the I_k) we obtain $\mathrm{var}\{I\} \approx (1/N)\sigma_I^2$. Fig. 5(a) and (b) illustrate example histograms of I for a 64-tone OFDM system in the flat-fading case and in the high delay-spread case, respectively. It can be seen that in the high delay-spread case the distribution is significantly more concentrated around the mean. Take a rate of say 7.5 b/s/Hz. For this rate, clearly from Fig. 5 the outage probability will be much lower for the high delay-spread case than for the flat-fading case.

Since the rank of the individual correlation matrices $\mathbf{R}_l$ $(l = 0, 1, \ldots, L-1)$ determines the number of spatial degrees of freedom in each path of the MIMO channel, it is to be expected that the rank of the individual correlation matrices and not the rank of the sum correlation matrix $\mathbf{R}$ determines the outage properties. This can be illustrated by assuming a simple example where all the $\mathbf{R}_l$ have rank 1 but are such that the sum correlation matrix has full rank. In this case, it readily follows from (4) that the number of degrees of freedom in each path is M_T and hence the total number of degrees of freedom in the channel is LM_T, irrespectively of the rank of the sum correlation matrix $\mathbf{R}$. In the case where the individual correlation matrices $\mathbf{R}_l$ are full rank, the sum correlation matrix $\mathbf{R}$ is also full rank,[8] but the number of degrees of freedom in the channel will be LM_TM_R, and hence significantly better outage properties than in the fully correlated case can be expected. These statements will be corroborated by means of simulation results in Section V. We conclude by noting that, while the multiplexing gain is determined by the rank of the sum correlation matrix $\mathbf{R}$, the diversity gain will be governed by the rank of the individual correlation matrices $\mathbf{R}_l$.

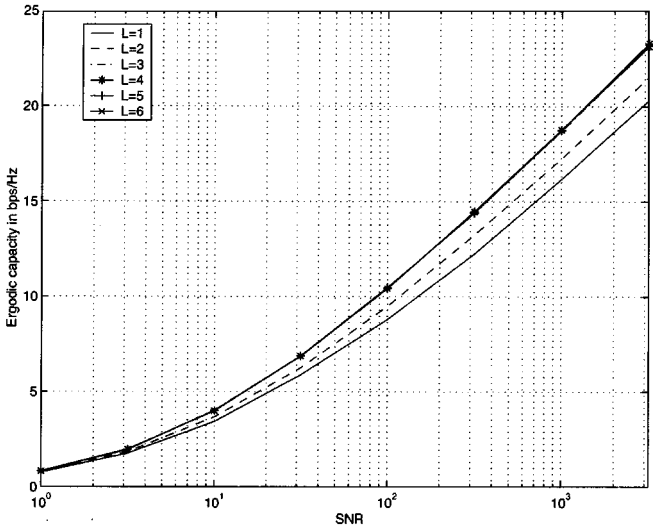
V. Simulation Results

In every simulation example, 1000 independent Monte Carlo runs were performed. Unless specified otherwise, the power delay profile was taken to be exponential, the number of tones in the OFDM system was $N = 512$, the CP length was 64, and the relative antenna spacing was set to $\Delta = 0.5$. For the sake of simplicity, we assume uniform tap spacing in all simulation examples. Finally, the SNR was defined as $\mathrm{SNR} = M_T\rho = \bar{\rho} = (P/(N\sigma_n^2))$.

[8]This follows from application of [28, Lemma 4.1] to $\mathbf{R} = \sum_{l=0}^{L-1} \mathbf{R}_l$ and noting that $\vartheta_3(\nu, q)$ is nonnegative.



(a)



(b)

Fig. 6. Ergodic capacity (in b/s/Hz) as a function of SNR for various values of L and (a) small cluster angle spread and (b) large cluster angle spread.

A. Simulation Results

1) Simulation Example 1: In the first simulation example, we study the impact of delay spread on ergodic capacity corroborating the statement that in the MIMO case delay spread may provide advantage over the flat-fading case in terms of ergodic capacity (provided that the total angle spread is large). The number of antennas was $M_R = M_T = 4$. In order to make the comparison fair, we normalize the energy in the channel by setting $\mathrm{Tr}(\mathbf{R}) = 1$ for all cases. The cluster angle spread was assumed to be $\sigma_{\theta,l} = 0$ $(l = 0, 1, \ldots, L-1)$. In the flat-fading case, the mean angle of arrival was set to $\bar{\theta}_0 = \pi/2$. In the delay-spread case, we assumed a total angle spread of 90 degrees. Fig. 6(a) shows the ergodic capacity (in b/s/Hz) as a function of SNR for different values of L. We can see that ergodic capacity indeed increases for increasing delay spread. We can furthermore observe that increasing the number of resolvable

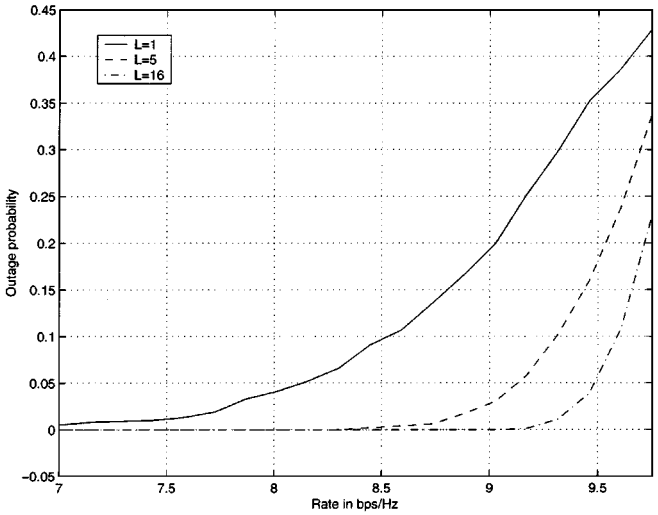


Fig. 7. Outage probability for $L = 1$, 5, and 16 at an SNR of 10 dB.

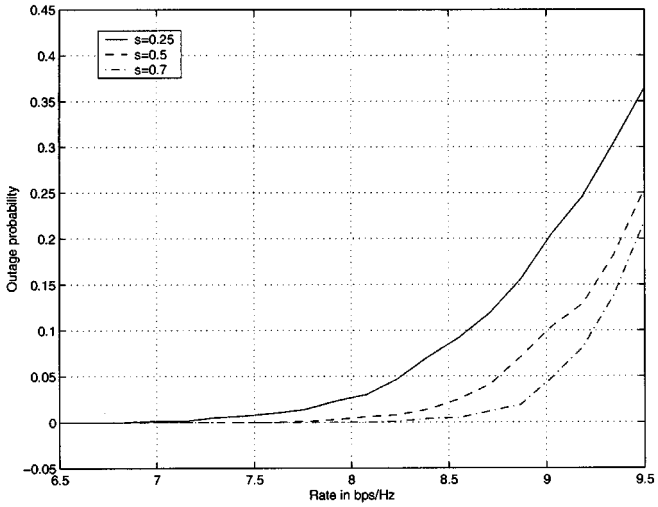


Fig. 8. Outage probability for various values of $s = \sigma_{\theta,l}$ at an SNR of 10 dB.

taps beyond 4 does not further increase ergodic capacity. The reason for this is that the number of transmit and receive antennas was set to 4 and hence the maximum rank of the sum correlation matrix $\mathbf{R}$ is 4. Fig. 6(b) shows the ergodic capacity for the same parameters as above except for the cluster angle spread which was increased to $\sigma_{\theta,l} = 0.25\,(l = 0, 1, \ldots, L-1)$. In this case, the rank of the individual correlation matrices is higher than 1 and the improvement in terms of ergodic capacity resulting from the presence of multiple taps is less pronounced. We emphasize that, as already stated in Section IV-A, this result is a consequence of the assumption that delayed paths tend to increase the total angle spread.

2) Simulation Example 2: In the second simulation example, we investigate the impact of delay spread on the outage properties of the system. Again, for fixed $\mathrm{Tr}(\mathbf{R}) = 1$, Fig. 7 shows the outage probability for $L = 1$, 5 and 16 and an SNR of 10 dB. Here, we assumed that there is no spatial fading correlation. It is clearly seen that the outage probability decreases significantly with increasing delay spread.

3) Simulation Example 3: In the last simulation example, we investigate the impact of spatial fading correlation on outage probability. For $M_T = M_R = 4$, $L = 10$ and $s = \sigma_{\theta,l} = 0.25$, 0.5, and 0.7, Fig. 8 shows the outage probability as a function of rate for an SNR of 10 dB. In all three simulations, the mean angles of arrival were chosen such that the sum correlation matrix $\mathbf{R}$ was full rank. Again, in all cases $\mathrm{Tr}(\mathbf{R}) = 1$. It can be seen that, even though the sum correlation matrix $\mathbf{R}$ has full rank, the outage probability depends critically on the individual cluster angle spreads and hence the rank of the individual correlation matrices $\mathbf{R}_l$.

VI. Conclusion

Based on a physically motivated model for broad-band MIMO fading channels, we derived expressions for the ergodic capacity and for outage capacity of OFDM-based spatial multiplexing systems for the case where the channel is unknown at the transmitter and perfectly known at the receiver. We studied the influence of propagation parameters and system parameters on ergodic capacity and outage probability and demonstrated the beneficial impact of delay spread and angle spread on capacity. Specifically, we showed that, in the MIMO case as opposed to the SISO case, delay-spread channels may provide advantage over flat-fading channels not only in terms of outage capacity but also in terms of ergodic capacity (provided the assumption that delayed paths tend to increase the total angle spread is true). We furthermore found that, while the multiplexing gain is governed by the rank of the sum correlation matrix $\mathbf{R}$, the diversity gain seems to be governed by the rank of the individual correlation matrices $\mathbf{R}_l$.

Directions for further work include the analysis of the case where there is scattering at both the transmitter and the receiver. A question of particular importance seems to be the analysis of the influence of scattering radii and distance between BTS and SU on capacity. Furthermore, a detailed study of the influence of different antenna geometries on the capacity of OFDM-based spatial multiplexing systems appears to be of interest. This problem has been studied to some extent in [7] for the narrow-band frequency-flat fading case.

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive criticism which helped to significantly improve the quality of the paper and the exposition. They would furthermore like to thank G. Wunder and R. W. Heath, Jr., for their detailed comments on the paper.

References

[1] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," U. S. Patent 5 345 599, 1994.

[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, Autumn 1996.

[3] I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," AT & T Bell Laboratories, BL0 112 170-950 615-07TM, 1995.

[4] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[5] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.

[6] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.
[7] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
[8] A. Gorokhov, "Capacity of multiple-antenna Rayleigh channel with a limited transmit diversity," in *Proc. IEEE Int. Symp. on Information Theory, 2000*, Sorrento, Italy, June 2000, p. 411.
[9] D. Asztély, "On Antenna Arrays in Mobile Communication Systems: Fast Fading and GSM Base Station Receiver Algorithms," Royal Institute of Technology, Stockholm, Sweden, IR-S3-SB-9611, 1996.
[10] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas," in *Proc. Inst. Elect. Eng.*, vol. 145, Feb. 1998, pp. 32–41.
[11] G. G. Raleigh and V. K. Jones, "Multivariate modulation and coding for wireless communication," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 851–866, 1999.
[12] L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case," *Bell Syst. Tech. J.*, vol. 53, pp. 745–778, May–June 1974.
[13] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, vol. 39, pp. 773–785, Mar. 1993.
[14] R. Müller, "A random matrix theory of communication via antenna arrays," *IEEE Trans. Inform. Theory*, submitted for publication.
[15] D. Agarwal, V. Tarokh, A. F. Naguib, and N. Seshadri, "Space-time coded OFDM for high data rate wireless communication over wideband channels," in *Proc. VTC 98*, May 1998, pp. 2232–2236.
[16] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 461–471, Mar. 1999.
[17] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Commun.*, pp. 10–22, Feb. 1998.
[18] A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," in *Proc. IEEE ICASSP-80*, Denver, CO, 1980, pp. 964–967.
[19] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[21] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.
[22] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.
[23] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.
[24] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. New York: Chelsea, 1984.
[25] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. San Diego, CA: Academic, 1994.
[26] V. Erceg, private communication, Apr. 2001.
[27] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based multi-antenna systems," in *Proc. IEEE ICASSP-00*, vol. 5, Istanbul, Turkey, June 2000, pp. 2569–2572.
[28] R. M. Gray, "Toeplitz and circulant matrices," Stanford Univ. ISL, 2000.

Helmut Bölcskei (M'98) was born in Mödling, Austria, on May 29, 1970. He received the Dipl.-Ing. and Dr. techn. degrees in electrical engineering/communications from Vienna University of Technology, Vienna, Austria, in 1994 and 1997, respectively.

From 1994 to 2001, he was with the Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology. From March 2001 to January 2002, he was an Assistant Professor of Electrical Engineering at the University of Illinois at Urbana-Champaign. Since February 2002, he has been an Assistant Professor of Communication Theory at ETH Zürich, Switzerland. From February to May 1996, he was a Visiting Researcher at Philips Research Laboratories Eindhoven, The Netherlands. From February to March 1998, he visited the Signal and Image Processing Department at ENST Paris, France. From February 1999 to February 2001, he was an Erwin Schrödinger Fellow of the Austrian National Science Foundation (FWF) performing research in the Smart Antennas Research Group in the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA. From 1999 to 2001, he was a consultant for Iospan Wireless, Inc. (formerly Gigabit Wireless Inc.), San Jose, CA. His research interests include communication and information theory and statistical signal processing with special emphasis on wireless communications, multi-input multi-output (MIMO) antenna systems, space-time coding, orthogonal frequency division multiplexing (OFDM), and wireless networks.

Dr. Bölcskei received a 2001 IEEE Signal Processing Society Young Author paper awards and serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

David Gesbert (S'96–M'99) was born in France in 1969. He received the M.Sc. degree in electrical engineering from National Institute for Telecommunications (INT), Evry, France, in 1993 and the Ph.D. degree from Ecole Nationale Superieure des Telecommunications, Paris, France, in 1997.

From 1993 to 1997, he was with France Telecom Research, where he was involved in the development and study of receiver algorithms for digital radio communications systems, with emphasis on blind signal detection. From April 1997 to October 1998, he was a Post-Doctoral Fellow in the Smart Antennas Research Group, Information Systems Laboratory, Stanford University, Stanford, CA. In October 1998, he took part in the founding engineering team of Iospan Wireless, Inc., formerly Gigabit Wireless Inc., San Jose, CA, a startup company promoting high-speed wireless data networks using smart antennas. In January 2001, he joined the Signal Processing Group, Department of Informatics at the University of Oslo, Norway, as an Adjunct Associate Professor in parallel to his activities at Iospan. His research interests are in the area of signal processing for digital communications, blind array processing, multi-input multi-output (MIMO) systems, multi-user communications, and adaptive wireless networks.

Arogyaswami J. Paulraj (SM'85–F'91) received the Ph.D. degree from the Naval Engineering College and Indian Institute of Technology, Bangalore, in 1973.

He has been a Professor at the Department of Electrical Engineering, Stanford University, Stanford, CA, since 1993, where he supervises the Smart Antennas Research Group. This group consists of approximately 12 researchers working on applications of space–time signal processing for wireless communications networks. His research group has developed many key fundamentals of this new field and helped shape a worldwide research and development focus onto this technology. His nonacademic positions included Head, Sonar Division, Naval Oceanographic Laboratory, Cochin, India; Director, Center for Artificial Intelligence and Robotics, Bangalore; Director, Center for Development of Advanced Computing; Chief Scientist, Bharat Electronics, Bangalore, and Chief Technical Officer and Founder, Iospan Wireless, Inc., San Jose, CA. He has also held visiting appointments at Indian Institute of Technology, Delhi, Loughborough University of Technology, and Stanford University. He sits on several boards of directors and advisory boards for U.S. and Indian companies/venture partnerships. His research has spanned several disciplines, emphasizing estimation theory, sensor signal processing, parallel computer architectures/algorithms and space–time wireless communications. His engineering experience included development of sonar systems, massively parallel computers, and more recently broad-band wireless systems. He is the author of over 250 research papers and holds eight patents.

Dr. Paulraj has won several awards for his engineering and research contributions. These include two President of India Medals, the CNS Medal, the Jain Medal, the Distinguished Service Medal, the Most Distinguished Service Medal, the VASVIK Medal, and the IEEE Best Paper Award (Joint), amongst others. He is a member of the Indian National Academy of Engineering.