Principles of MIMO-OFDM Wireless Systems

Helmut Bölcskei

Communication Technology Laboratory
Swiss Federal Institute of Technology (ETH)
Sternwartstrasse 7, CH-8092 Zürich
Phone: +41-1-6323433, Fax: +41-1-6321209, Email: boelcskei@nari.ee.ethz.ch

Abstract

The use of multiple antennas at both ends of a wireless link (**MIMO** technology) holds the potential to drastically improve the *spectral efficiency* and **link reliability** in future wireless communications systems. A particularly promising candidate for next-generation fixed and mobile wireless systems is the combination of MIMO technology with **Orthogonal Frequency Division Multiplexing (OFDM)**. This chapter provides an overview of the basic principles of MIMO-OFDM.

1 Introduction

The major challenges in future wireless communications system design are *increased spectral efficiency* and *improved link reliability*. The wireless channel constitutes a hostile propagation medium, which suffers from **fading** (caused by destructive addition of multipath components) and *interference from other users*. Diversity provides the receiver with several (ideally independent) replicas of the transmitted signal and is therefore a powerful means to combat fading and interference and thereby improve link reliability. Common forms of diversity are time diversity (due to Doppler spread) and frequency diversity (due to delay spread). In recent years the use of spatial (or antenna) diversity has become very popular, which is mostly due to the fact that it can be provided without loss in spectral efficiency. *Receive diversity*, that is, the use of multiple antennas on the receive side of a wireless

link, is a well-studied subject [1]. Driven by mobile wireless applications, where it is difficult to deploy multiple antennas in the handset, the use of multiple antennas on the transmit side combined with signal processing and coding has become known under the name of **space-time coding** [2–4] and is currently an active area of research. The use of *multiple antennas* at *both ends of a wireless link* (multiple-input multiple-output (MIMO) technology) has recently been demonstrated to have the potential of achieving extraordinary data rates [5–9]. The corresponding technology is known as **spatial multiplexing** [5,9] or BLAST [6,10] and yields an impressive increase in spectral efficiency.

Most of the previous work in the area of MIMO wireless has been restricted to narrowband systems. Besides spatial diversity broadband MIMO channels, however, offer higher capacity and frequency diversity due to delay spread. *Orthogonal frequency division multiplexing* (OFDM) [11,12] *significantly reduces receiver complexity* in wireless broadband systems. The use of MIMO technology in combination with OFDM, i.e., MIMO-OFDM [8,9,13], therefore seems to be an attractive solution for future broadband wireless systems.

The purpose of this chapter is to provide a survey of the basic principles of MIMO-OFDM. Most of the material in this chapter appeared in (more detail) in [9, 13, 14].

Notation. $\mathcal{E}\{\cdot\}$ denotes the expectation operator, $\mathbf{I}_M$ is the $M \times M$ identity matrix, $\mathbf{0}_{M\times N}$ stands for the $M \times N$ all zeros matrix, $\|\mathbf{A}\|_F^2 = \sum_{i,j} |[\mathbf{A}]_{i,j}|^2$ is the squared Frobenius norm of the matrix $\mathbf{A}$, $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of the matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathrm{Tr}(\mathbf{A})$ and $\det(\mathbf{A})$ stand for the trace and the determinant of $\mathbf{A}$, respectively, $r(\mathbf{A})$ is the rank of $\mathbf{A}$, $\lambda_i(\mathbf{A})$ denotes the i-th eigenvalue of $\mathbf{A}$, $\sigma(\mathbf{A})$ stands for the eigenvalue spectrum of $\mathbf{A}$, $\mathbf{a}_i$ denotes the i-th column of $\mathbf{A}$, and $\mathbf{vec}(\mathbf{A}) = [\mathbf{a}_0^T \ \mathbf{a}_1^T \ \ldots \ \mathbf{a}_{N-1}^T]^T$. The superscripts ^{T,*,H} denote transpose, elementwise conjugation, and conjugate transpose, respectively. 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 independent, identically distributed (i.i.d.) $\mathcal{N}(0, \sigma^2/2)$.

2 The Broadband MIMO Fading Channel

In this section, we shall introduce a model for MIMO broadband fading channels taking into account real-world propagation conditions.

Basic assumptions. In the following, M_T and M_R denote the number of transmit and receive antennas, respectively. We assume that the discrete-time $M_R \times M_T$ matrix-valued channel has order $L-1$ with transfer function

$$\mathbf{H}(e^{j2\pi\theta}) = \sum_{l=0}^{L-1} \mathbf{H}_l e^{-j2\pi l\theta}, \quad 0 \leq \theta < 1, \tag{1}$$

where the $M_R \times M_T$ complex-valued random matrix $\mathbf{H}_l$ represents the l-th tap. One can think of each of the taps as corresponding to a significant scatterer cluster (see Fig. 1) with each of the paths emanating from within the same scatterer cluster experiencing the same delay. We write each of the taps as the sum of a fixed (possibly line-of-sight) component, $\bar{\mathbf{H}}_l = \mathcal{E}\{\mathbf{H}_l\}$, and a variable (or scattered) component $\widetilde{\mathbf{H}}_l$ as

$$\mathbf{H}_l = \bar{\mathbf{H}}_l + \widetilde{\mathbf{H}}_l, \quad l = 0, 1, \ldots, L-1.$$

The channel is said to be Rayleigh fading if $\bar{\mathbf{H}}_l = \mathbf{0}_{M_R \times M_T}$ for $l = 0, 1, \ldots, L-1$ and Ricean fading if $\bar{\mathbf{H}}_l \neq \mathbf{0}_{M_R \times M_T}$ for at least one $l \in [0, L-1]$. The elements of the matrices $\widetilde{\mathbf{H}}_l$ $(l = 0, 1, \ldots, L-1)$ are (possibly correlated) circularly symmetric complex Gaussian random variables. Different *scatterer clusters are assumed to induce uncorrelated fading* (or equivalently different taps fade independently), i.e.,

$$\mathcal{E}\left\{\mathbf{vec}(\widetilde{\mathbf{H}}_l)\left(\mathbf{vec}(\widetilde{\mathbf{H}}_{l'})\right)^H\right\} = \mathbf{0}_{M_R M_T \times M_R M_T} \quad \text{for } l \neq l'. \tag{2}$$

Each scatterer cluster has a mean angle of departure from the transmit array and a mean angle of arrival at the receive array denoted as $\bar{\theta}_{T,l}$ and $\bar{\theta}_{R,l}$ (see Fig. 1), respectively, a cluster angle spread as perceived by the transmitter $\sigma^2_{\theta_{T,l}}$ (proportional to the scattering radius of the cluster as observed by the transmitter), a cluster angle spread as perceived by the receiver $\sigma^2_{\theta_{R,l}}$ (proportional to the scattering radius of the cluster as observed by the receiver), and a path gain σ_l^2 (derived from the power delay profile of the channel). Finally, we define the total transmit and receive angle spread as the spread of the $\bar{\theta}_{T,l}$ and the $\bar{\theta}_{R,l}$, respectively.

Array geometry. We assume a uniform linear array at both the transmitter and the receiver with identical uni-polarized antenna elements. The relative transmit and receive antenna spacing is denoted as $\Delta_T = \frac{d_T}{\lambda}$ and $\Delta_R = \frac{d_R}{\lambda}$, respectively, where d_T and d_R stand for absolute antenna spacing and $\lambda = c/f_c$ is the wavelength of a narrowband signal with center frequency f_c.

Fading statistics. Spatial fading correlation can occur both at the transmitter and the receiver, the impact of which is modeled by decomposing the Rayleigh component of the l-th tap according to

$$\widetilde{\mathbf{H}}_l = \mathbf{R}_l^{1/2} \widetilde{\mathbf{H}}_{w,l} \left(\mathbf{S}_l^{1/2} \right)^T, \quad l = 0, 1, \ldots, L-1, \tag{3}$$

where $\mathbf{R}_l = \mathbf{R}_l^{1/2} \mathbf{R}_l^{1/2}$ and $\mathbf{S}_l = \mathbf{S}_l^{1/2} \mathbf{S}_l^{1/2}$ are the receive and transmit correlation matrices, respectively, and $\widetilde{\mathbf{H}}_{w,l}$ is an $M_R \times M_T$ matrix with i.i.d. $\mathcal{CN}(0, \sigma_l^2)$ entries. We note that the decomposition (3) does not incorporate the most general case of spatial fading correlation, but yields a reasonable compromise between analytical tractability and validity of the channel model.

Note that the power delay profile σ_l^2 has been incorporated into the matrices $\widetilde{\mathbf{H}}_{w,l}$. From (2) we have $\mathcal{E}\left\{ \mathbf{vec}(\widetilde{\mathbf{H}}_{w,l}) \left(\mathbf{vec}(\widetilde{\mathbf{H}}_{w,l'}) \right)^H \right\} = \mathbf{0}_{M_R M_T \times M_R M_T}$ for $l \neq l'$. In the following, we define $\rho(s\Delta, \bar{\theta}, \sigma_\theta)$ to be the fading correlation between two antenna elements spaced $s\Delta$ wavelengths apart. The correlation matrices $\mathbf{R}_l$ and $\mathbf{S}_l$ are consequently given by

$$[\mathbf{R}_l]_{m,n} = \rho((n-m)\Delta_R, \bar{\theta}_{R,l}, \sigma_{\theta_{R,l}}) \quad \text{and}$$

$$[\mathbf{S}_l]_{m,n} = \rho((n-m)\Delta_T, \bar{\theta}_{T,l}, \sigma_{\theta_{T,l}}).$$

Let us next assume that the actual angle of departure for the l-th path cluster is given by $\theta_{T,l} = \bar{\theta}_{T,l} + \hat{\theta}_{T,l}$ with $\hat{\theta}_{T,l} \sim \mathcal{N}(0, \sigma^2_{\theta_{T,l}})$ and the actual angle of arrival is $\theta_{R,l} = \bar{\theta}_{R,l} + \hat{\theta}_{R,l}$ with $\hat{\theta}_{R,l} \sim \mathcal{N}(0, \sigma^2_{\theta_{R,l}})$. With these assumptions, we obtain [15]

$$\rho(s\Delta, \bar{\theta}, \sigma_\theta) \approx e^{-j2\pi s\Delta \cos(\bar{\theta})} e^{-\frac{1}{2}(2\pi s\Delta \sin(\bar{\theta})\sigma_\theta)^2} \tag{4}$$

which implies that the correlation function is essentially a Gaussian with spread inversely proportional to the product of antenna spacing and cluster angle spread. Consequently large antenna spacing and/or large cluster angle spread lead to small spatial fading correlation and vice versa. It must be stressed, however, that (4) is an approximation which becomes inaccurate when the mean angle of arrival is close to 0 or close to π or when the cluster angle spread is large. We note that in the limiting case of zero receive angle spread, i.e., $\sigma_{\theta_{R,l}} = 0$, the receive correlation matrix has rank 1 with $\mathbf{R}_l = \mathbf{a}(\bar{\theta}_{R,l}) \mathbf{a}^H(\bar{\theta}_{R,l})$, where

$$\mathbf{a}(\theta) = [1 \;\; e^{j2\pi\Delta\cos(\theta)} \;\; \ldots \;\; e^{j2\pi(M_R-1)\Delta\cos(\theta)}]^T. \tag{5}$$

Likewise for $\sigma_{\theta_{T,l}} = 0$, we have $\mathbf{S}_l = \mathbf{a}(\bar{\theta}_{T,l}) \mathbf{a}^H(\bar{\theta}_{T,l})$.

Ricean component. The Ricean component of the l-th tap is modeled as

$$\bar{\mathbf{H}}_l = \sum_{i=0}^{P_l-1} \beta_{l,i}\, \mathbf{a}\big(\bar{\theta}_{R,l,i}\big)\, \mathbf{a}^T\big(\bar{\theta}_{T,l,i}\big)\,, \tag{6}$$

where $\bar{\theta}_{R,l,i}$ and $\bar{\theta}_{T,l,i}$ denote the angle of arrival and the angle of departure, respectively, of the i-th component of $\bar{\mathbf{H}}_l$ and $\beta_{l,i}$ is the corresponding complex-valued path amplitude. We can furthermore associate a Ricean K-factor with each of the taps by defining

$$K_l = \frac{\|\bar{\mathbf{H}}_l\|_F^2}{\mathcal{E}\{\|\widetilde{\mathbf{H}}_l\|_F^2\}}, \quad l = 0, 1, \ldots, L-1.$$

We note that large cluster angle spread will in general result in a high-rank Ricean component.

Comments on the channel model. For the sake of simplicity of exposition, we assumed that different scatterer clusters can be resolved in time and hence correspond to different delays. In practice, this is not necessarily the case. We emphasize, however, that allowing different scatterer clusters to have the same delay does in general not yield significant new insights into the impact of the propagation conditions on the performance of MIMO-OFDM systems.

3 Capacity of Broadband MIMO-OFDM Systems

This section is devoted to the capacity of broadband MIMO-OFDM systems operating in spatial multiplexing mode. Spatial multiplexing [5,9], also referred to as BLAST [6,10] increases the capacity of wireless radio links drastically with no additional power or bandwidth consumption. The technology requires multiple antennas at both ends of the wireless link and realizes capacity gains (denoted as **spatial multiplexing gain**) by sending independent data streams from different antennas.

MIMO-OFDM

The main motivation for using OFDM in a MIMO channel is the fact that OFDM modulation turns a frequency-selective MIMO channel into a set of parallel frequency-flat MIMO channels. This renders multi-channel equalization particularly simple, since for each OFDM-tone only a constant matrix has to be inverted [8,9].

In a MIMO-OFDM system with N subcarriers (or tones) the individual data streams are first passed through OFDM modulators which perform an **IFFT** on blocks of length N followed by a

parallel-to-serial conversion. A cyclic prefix (CP) of length $L_{cp} \geq L$ containing a copy of the last L_{cp} samples of the parallel-to-serial converted output of the N-point IFFT is then prepended. The resulting OFDM symbols of length $N + L_{cp}$ are launched simultaneously from the individual transmit antennas. The CP is essentially a guard interval which serves to eliminate interference between OFDM symbols and turns linear convolution into circular convolution such that the channel is diagonalized by the **FFT**. In the receiver the individual signals are passed through OFDM demodulators which first discard the CP and then perform an N-point FFT. The outputs of the OFDM demodulators are finally separated and decoded. Fig. 2 shows a schematic of a MIMO-OFDM system. For a more detailed discussion of the basic principles of OFDM the interested reader is referred to [16]. The assumption of the length of the CP being greater or equal than the length of the discrete-time baseband channel impulse response (i.e. $L_{cp} \geq L$) guarantees that the frequency-selective MIMO fading channel indeed decouples into a set of parallel frequency-flat MIMO fading channels [11]. Organizing the transmitted data symbols into frequency vectors $\mathbf{c}_k = [c_k^{(0)} \;\; c_k^{(1)} \;\; \ldots \;\; c_k^{(M_T-1)}]^T$ $(k = 0, 1, \ldots, N-1)$ with $c_k^{(i)}$ denoting the data symbol transmitted from the i-th antenna on the k-th tone, the reconstructed data vector for the k-th tone is given by [8, 9]

$$\mathbf{r}_k = \mathbf{H}(e^{j\frac{2\pi}{N}k})\,\mathbf{c}_k + \mathbf{n}_k, \quad k = 0, 1, \ldots, N-1, \tag{7}$$

where $\mathbf{n}_k$ is complex-valued circularly symmetric additive white Gaussian noise satisfying $\mathcal{E}\{\mathbf{n}_k \mathbf{n}_l^H\} = \sigma_n^2 \mathbf{I}_{M_R} \delta[k-l]$.

Capacity of MIMO-OFDM Spatial Multiplexing Systems

In the following, we ignore the loss in spectral efficiency due to the presence of the CP (recall that the CP contains redundant information). We assume that the channel is purely Rayleigh fading, ergodic, remains constant over a block spanning at least one OFDM symbol and changes in an independent fashion from block to block. For the sake of simplicity of exposition, we restrict our attention to the case of receive correlation only (i.e. the transmit antennas fade in an uncorrelated fashion).

In OFDM-based spatial multiplexing systems statistically independent data streams are transmitted from different antennas and different tones and the total available power is allocated uniformly across all space-frequency subchannels [9]. Assuming that coding and **interleaving** are performed

across OFDM-symbols and that the number of fading blocks spanned by a codeword goes to infinity whereas the blocksize (which equals the number of tones in the OFDM system multiplied by the number of OFDM symbols spanning one channel use or equivalently one block) remains constant (and finite), an ergodic or Shannon capacity exists and is given by[1] [9]

$$C = \mathcal{E}\left\{\log\det\left(\mathbf{I}_{M_R} + \rho\,\mathbf{\Lambda}\widetilde{\mathbf{H}}_w\widetilde{\mathbf{H}}_w^H\right)\right\} \quad \text{bps/Hz} \tag{8}$$

where $\mathbf{\Lambda} = \text{diag}\{\lambda_i(\mathbf{R})\}_{i=0}^{M_R-1}$ with $\mathbf{R} = \sum_{l=0}^{L-1}\sigma_l^2\,\mathbf{R}_l$, $\rho = \frac{P}{M_T N\sigma_n^2}$ with P denoting the total available power, $\widetilde{\mathbf{H}}_w$ is an $M_R \times M_T$ i.i.d. random matrix with $\mathcal{CN}(0,1)$ entries, and the expectation is taken with respect to $\widetilde{\mathbf{H}}_w$. The operational meaning of C is as follows. At rates lower than C, the error probability (for a good code) decays exponentially with the transmission length. Capacity can be achieved in principle by transmitting a codeword over a very large number of independently fading blocks.

It is instructive to study the case of fixed M_R with M_T large, where $\frac{1}{M_T}\widetilde{\mathbf{H}}_w\widetilde{\mathbf{H}}_w^H \rightarrow \mathbf{I}_{M_R}$ and consequently

$$C \;=\; \log\det\left(\mathbf{I}_{M_R} + \bar{\rho}\,\mathbf{\Lambda}\right) \tag{9}$$

with $\bar{\rho} = M_T\rho = \frac{P}{N\sigma_n^2}$. For small $\bar{\rho}$, it follows from (9) that in the large M_T limit [9]

$$C \;\approx\; \log\left(1 + \bar{\rho}\,\text{Tr}(\mathbf{R})\right),$$

where all the higher-order terms in $\bar{\rho}$ have been neglected. Thus, in the low signal-to-noise ratio (SNR) regime the ergodic capacity is governed by $\text{Tr}(\mathbf{R})$. In the high-SNR regime, we obtain a fundamentally different conclusion. Starting from (9) we have

$$C \;=\; \sum_{i=0}^{r(\mathbf{R})-1}\log(1 + \bar{\rho}\,\lambda_i(\mathbf{R})) \tag{10}$$

with $\lambda_i(\mathbf{R})$ denoting the nonzero eigenvalues of $\mathbf{R}$. The rank and eigenvalue spread of the sum correlation matrix $\mathbf{R} = \sum_{l=0}^{L-1}\sigma_l^2\,\mathbf{R}_l$ therefore critically determine ergodic capacity in the high-SNR regime. In fact, it follows directly from (10) that the multiplexing gain in the large M_T limit is given by $r(\mathbf{R})$. Moreover, for a given $\text{Tr}(\mathbf{R})$, the right-hand-side (RHS) in (10) is maximized if $r(\mathbf{R}) = M_R$ and all the $\lambda_i(\mathbf{R})\,(i = 0, 1, ..., M_R - 1)$ are equal [9]. 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.

[1]Throughout the chapter all logarithms are to the base 2.

Impact of Propagation Parameters on Capacity

We shall next show how the propagation parameters impact the eigenvalues of the sum correlation matrix $\mathbf{R}$ and hence ergodic capacity. Since the individual correlation matrices $\mathbf{R}_l$ $(l = 0, 1, ..., L-1)$ are Toeplitz the sum correlation matrix $\mathbf{R}$ is Toeplitz as well. Using (4) and applying Szegö's theorem [17] to $\mathbf{R}$, we obtain the limiting ($M_R \rightarrow \infty$) distribution[2] of the eigenvalues of $\mathbf{R} = \sum_{l=0}^{L-1} \sigma_l^2 \, \mathbf{R}_l$ as

$$\lambda(\nu) = \sum_{l=0}^{L-1} \sigma_l^2 \underbrace{\vartheta_3\Big(\pi(\nu - \Delta_R \cos(\bar{\theta}_{R,l})), e^{-\frac{1}{2}(2\,\pi\,\Delta_R\,\sin(\bar{\theta}_{R,l})\sigma_{\theta_{R,l}})^2}\Big)}_{\lambda_l(\nu)}, \quad 0 \leq \nu < 1 \tag{11}$$

with the third-order theta function given by $\vartheta_3(\nu, q) = \sum_{n=-\infty}^{\infty} q^{n^2} \, e^{2jn\nu}$. Although this expression yields the exact eigenvalue distribution only in the limiting case $M_R \rightarrow \infty$, good approximations of the eigenvalues for finite M_R can be obtained by sampling $\lambda(\nu)$ uniformly, which allows us to assume that the eigenvalue distribution in the finite M_R case follows the shape of $\lambda(\nu)$. This observation combined with (11) shall next be used to relate propagation and system parameters to the eigenvalues of $\mathbf{R}$ and hence ergodic capacity.

Impact of cluster angle spread and antenna spacing. Let us start by investigating the influence of receive cluster angle spreads and receive antenna spacing on ergodic capacity. For the sake of simplicity consider a single-tap channel (i.e. $L = 1$) with associated receive correlation matrix $\mathbf{R}_0$. The limiting eigenvalue distribution of $\mathbf{R} = \sigma_0^2 \, \mathbf{R}_0$ is given by

$$\lambda(\nu) = \sigma_0^2 \, \vartheta_3 \left(\pi(\nu - \Delta_R \cos(\bar{\theta}_{R,0})), e^{-\frac{1}{2}(2\,\pi\,\Delta_R\,\sin(\bar{\theta}_{R,0})\sigma_{\theta_{R,0}})^2}\right).$$

Now, noting that the correlation function $\rho(s\Delta_R, \bar{\theta}_{R,0}, \sigma_{\theta_{R,0}})$ as a function of s is essentially a modulated Gaussian function with its spread decreasing for increasing antenna spacing and/or increasing cluster angle spread and vice versa, it follows that $\lambda(\nu)$ 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 $\lambda(\nu)$ will be peaky. Figs. 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 $\lambda(\nu)$ (or equivalently high spatial fading correlation) and vice versa.

[2]Note that for $M_R \rightarrow \infty$ the eigenvalues of $\mathbf{R}$ are characterized by a periodic continuous function [17]. Thus, in the following whenever we use the term *eigenvalue distribution*, we actually refer to this function.

Impact of total angle spread. We shall next study the impact of total receive angle spread on ergodic capacity. Assume that either the individual scatterer cluster angle spreads are small or that antenna spacing at the receiver is small or both. Hence, the individual $\lambda_l(\nu)$ will be peaky. Now, from (11) we can see that the limiting distribution $\lambda(\nu)$ is obtained by adding the individual limiting distributions $\lambda_l(\nu)$ weighted by the σ_l^2. Note furthermore that $\lambda_l(\nu)$ is essentially a Gaussian centered around $\Delta_R \cos(\bar{\theta}_{R,l})$. Now, if the total angle spread, i.e., the spread of the $\bar{\theta}_{R,l}$ is large the sum limiting distribution $\lambda(\nu)$ can still be flat even though the individual $\lambda_l(\nu)$ are peaky. For given small cluster angle spreads, Figs. 4 (a) and (b) show example limiting eigenvalue distributions for a 3-tap channel (assuming a uniform power delay profile) with a total angle spread of 22.5 degrees and 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 leads to reduced ergodic capacity.

Ergodic capacity in the SISO and in the MIMO case. It is well known that in the single-input single-output (SISO) case delay spread channels do not offer advantage over flat-fading channels in terms of ergodic capacity [18] (provided the receive SNR is kept constant). This can easily be seen from (8) by noting that in the SISO case $\mathbf{R} = \sum_{l=0}^{L-1} \sigma_l^2$ which implies that ergodic capacity is only a function of the total energy in the channel and does not depend on how this energy is distributed across taps. In the MIMO case the situation can be fundamentally different. Fix $\mathrm{Tr}(\mathbf{R})$, and take a flat-fading scenario (i.e., $L = 1$) with small antenna spacing such that $\mathbf{R} = \sigma_0^2 \mathbf{R}_0$ has rank 1. In this case the matrix $\mathbf{\Lambda} \widetilde{\mathbf{H}}_w \widetilde{\mathbf{H}}_w^H$ has rank 1 with probability one and hence only one spatial data pipe can be opened up, or equivalently there is no spatial multiplexing gain. Now, compare the flat-fading scenario to a frequency-selective fading scenario where $L \geq M_R$ and each of the $\mathbf{R}_l$ $(l = 0, 1, ..., L-1)$ has rank 1 but the sum correlation matrix $\mathbf{R} = \sum_{l=0}^{L-1} \sigma_l^2 \mathbf{R}_l$ has full rank. For this to happen a sufficiently large total angle spread is needed. Clearly, in this case $\min(M_T, M_R)$ 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 $\mathbf{R}_l = \mathbf{I}_{M_R}$ $(l = 0, 1, ..., L-1)$ this effect does not occur. However, this scenario corresponds to fully uncorrelated spatial fading on all taps and is therefore unlikely. *We can conclude that in practice*

MIMO delay spread channels offer advantage over MIMO flat-fading channels in terms of ergodic capacity. However, 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 broadband channels in the 2.5GHz band [19].

Numerical Results

We conclude this section with a numerical result studying the impact of delay spread on ergodic capacity. In this example the power delay profile was taken to be exponential, tap spacing was uniform, the relative receive antenna spacing was set to $\Delta_R = \frac{1}{2}$, SNR was defined as $\text{SNR} = \bar{\rho} = \frac{P}{N\sigma_n^2}$, and the number of antennas was $M_T = M_R = 4$. In order to make the following comparisons fair we normalize the energy in the channel by setting $\text{Tr}(\mathbf{R}) = 1$ for all cases. The cluster angle spreads were assumed to satisfy $\sigma_{\theta_{R,l}} = 0\,(l = 0, 1, ..., L-1)$. In the flat-fading case the mean angle of arrival was set to $\bar{\theta}_{R,0} = \pi/2$. In the frequency-selective case we assumed a total receive angle spread of 90 degrees. Fig. 5(a) shows the ergodic capacity (in bps/Hz) (obtained through evaluation of (8) using 1,000 independent Monte Carlo runs) as a function of SNR for different values of L. We can see that ergodic capacity indeed increases for increasing L and hence increasing rank of $\mathbf{R}$. Moreover, we observe that increasing the number of resolvable (i.e. independently fading) taps beyond 4 does not further increase ergodic capacity. The reason for this is that the multiplexing gain $\min(r(\mathbf{R}), M_T)$ cannot exceed 4. Fig. 5(b) shows the ergodic capacity for the same parameters as above except for the cluster angle spreads increased to $\sigma_{\theta_{R,l}} = 0.25\,(l = 0, 1, ..., L-1)$. In this case the rank of the individual correlation matrices $\mathbf{R}_l$ 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 the conclusions drawn in this simulation result are a consequence of the assumption that delayed paths increase the rank of the sum correlation matrix $\mathbf{R}$.

4 Space-Frequency Coded MIMO-OFDM

While spatial multiplexing realizes increased spectral efficiency, **space-frequency coding** [13,14,20] is used to improve link reliability through (spatial and frequency) **diversity gain**. In this section,

we describe the basics of space-frequency coded MIMO-OFDM.

Space-Frequency Coding

Using the notation introduced in Section 3, we start from the input-output relation

$$\mathbf{r}_k = \sqrt{E_s}\,\mathbf{H}(e^{j\frac{2\pi}{N}k})\,\mathbf{c}_k + \mathbf{n}_k, \quad k = 0, 1, \ldots, N-1, \tag{12}$$

where the data symbols $c_k^{(i)}$ are taken from a finite complex alphabet and have average energy 1. The constant E_s is an energy normalization factor. Throughout the remainder of this chapter, we assume that the channel is constant over the duration of at least one OFDM symbol, the transmitter has no channel knowledge, and the receiver knows the channel perfectly. We furthermore assume that coding is performed only within one OFDM symbol such that one data burst consists of N vectors of size $M_T \times 1$ or equivalently one spatial OFDM symbol.

The maximum likelihood (ML) decoder computes the vector sequence $\hat{\mathbf{c}}_k\,(k = 0, 1, \ldots, N-1)$ according to

$$\widehat{\mathbf{C}} = \arg\min_{\mathbf{C}} \sum_{k=0}^{N-1} \|\mathbf{r}_k - \sqrt{E_s}\,\mathbf{H}(e^{j\frac{2\pi}{N}k})\,\mathbf{c}_k\|^2,$$

where $\mathbf{C} = [\mathbf{c}_0\ \ \mathbf{c}_1\ \ \ldots\ \ \mathbf{c}_{N-1}]$ and $\widehat{\mathbf{C}} = [\hat{\mathbf{c}}_0\ \ \hat{\mathbf{c}}_1\ \ \ldots\ \ \hat{\mathbf{c}}_{N-1}]$. In the remainder of this section, we employ the channel model introduced in Section 2 in its full generality and we assume that $N > M_T L$.

Error Rate Performance

Assume that $\mathbf{C} = [\mathbf{c}_0\ \ \mathbf{c}_1\ \ \ldots\ \ \mathbf{c}_{N-1}]$ and $\mathbf{E} = [\mathbf{e}_0\ \ \mathbf{e}_1\ \ \ldots\ \ \mathbf{e}_{N-1}]$ are two different space-frequency codewords of size $M_T \times N$. The average (with respect to the random channel) probability that the receiver decides erroneously in favor of the signal $\mathbf{E}$ assuming that $\mathbf{C}$ was transmitted can be upper-bounded by the pairwise error probability (PEP) as [14]

$$P(\mathbf{C} \to \mathbf{E}) \le \exp\left(-\frac{E_s}{4\sigma_n^2}\|\bar{\mathbf{Y}}\|^2\right) \prod_{i=0}^{r(\mathbf{C}_Y)-1} \exp\left(\left(\frac{E_s}{4\sigma_n^2}\right)^2 \frac{|b_i|^2\lambda_i^2(\mathbf{C}_Y)}{1+\frac{E_s}{4\sigma_n^2}\lambda_i(\mathbf{C}_Y)}\right) \frac{1}{1+\frac{E_s}{4\sigma_n^2}\lambda_i(\mathbf{C}_Y)}, \tag{13}$$

where

$$\mathbf{C}_Y = \sum_{l=0}^{L-1} \sigma_l^2 \left[\mathbf{D}^l(\mathbf{C}-\mathbf{E})^T\mathbf{S}_l(\mathbf{C}-\mathbf{E})^*\mathbf{D}^{-l}\right] \otimes \mathbf{R}_l \tag{14}$$

with $\mathbf{D} = \text{diag}\{e^{-j\frac{2\pi}{N}k}\}_{k=0}^{N-1}$, $\bar{\mathbf{Y}} = [\bar{\mathbf{y}}_0^T \ \bar{\mathbf{y}}_1^T \ \ldots \ \bar{\mathbf{y}}_{N-1}^T]^T$ with $\bar{\mathbf{y}}_k = \bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})(\mathbf{c}_k - \mathbf{e}_k)$ and $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k}) = \sum_{l=0}^{L-1} \bar{\mathbf{H}}_l \, e^{-j\frac{2\pi}{N}kl}$, $\mathbf{C}_Y = \mathbf{U}\mathbf{\Phi}\mathbf{U}^H$ with $\mathbf{\Phi} = \text{diag}\{\lambda_0(\mathbf{C}_Y), \ldots, \lambda_{r(\mathbf{C}_Y)-1}(\mathbf{C}_Y), 0, \ldots, 0\}$, and $\mathbf{b} = [b_0 \ b_1 \ \ldots \ b_{M_R N-1}]^T = \text{diag}\left\{\frac{1}{\sqrt{\lambda_0(\mathbf{C}_Y)}}, \ldots, \frac{1}{\sqrt{\lambda_{r(\mathbf{C}_Y)-1}(\mathbf{C}_Y)}}, 0, \ldots, 0\right\} \mathbf{U}^H \bar{\mathbf{Y}}$.

Maximum Diversity Order and Coding Gain

Based on (13), we are now ready to establish the maximum achievable diversity order and coding gain for various propagation scenarios. We focus on the high-SNR regime, i.e., $\frac{E_s}{4\sigma_n^2} \gg 1$, and discuss the cases of Rayleigh fading and Ricean fading separately.

Rayleigh Fading

Let us start by assuming that the channel is purely Rayleigh fading, i.e., $\bar{\mathbf{H}}_l = \mathbf{0}_{M_R \times M_T}$ for $l = 0, 1, \ldots, L-1$. In this case (13) simplifies to

$$\mathbf{P}(\mathbf{C} \to \mathbf{E}) \leq \left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_Y)} \prod_{i=0}^{r(\mathbf{C}_Y)-1} \lambda_i^{-1}(\mathbf{C}_Y). \tag{15}$$

We recall the following definitions [2]. The *diversity order* d achieved by a space-frequency code is given by the minimum rank of $\mathbf{C}_Y$ over all codeword pairs $\{\mathbf{C}, \mathbf{E}\}$. The *coding gain* is defined as the minimum of $\left(\prod_{i=0}^{d-1} \lambda_i(\mathbf{C}_Y)\right)^{\frac{1}{d}}$ over all matrices $\mathbf{C}_Y$ with $r(\mathbf{C}_Y) = d$. In the case of i.i.d. channels, where $\mathbf{R}_l = \mathbf{I}_{M_R}$ and $\mathbf{S}_l = \mathbf{I}_{M_T}$ for $l = 0, 1, \ldots, L-1$, the maximum achievable diversity order is $d = M_T M_R |\mathcal{L}|$, where $\mathcal{L} = \{i : \sigma_i^2 > 0\}$ ($i \in [0, L-1]$), and $|\mathcal{L}|$ denotes the size of the set $\mathcal{L}$.

The propagation parameters determine the maximum achievable diversity order and coding gain through the rank and the eigenvalues of the matrix $\mathbf{C}_Y$. In the following, we shall therefore study the dependence of $\mathbf{C}_Y$ on the correlation matrices $\mathbf{R}_l$ and $\mathbf{S}_l$. According to the conclusions obtained, the discussion is organized into three different cases, namely receive correlation only, transmit correlation only, and joint transmit and receive correlation.

Receive correlation only. In this case $\mathbf{S}_l = \mathbf{I}_{M_T}$ for $l = 0, 1, \ldots, L-1$, and (14) specializes to

$$\mathbf{C}_Y = \sum_{l=0}^{L-1} \sigma_l^2 \left[\mathbf{D}^l (\mathbf{C} - \mathbf{E})^T (\mathbf{C} - \mathbf{E})^* \mathbf{D}^{-l}\right] \otimes \mathbf{R}_l.$$

We start with the assumption that all the correlation matrices $\mathbf{R}_l$ with $l \in \mathcal{L}$ have full rank. Defining

$$\mathbf{C}_u = \sum_{l=0}^{L-1} \sigma_l^2 \left[\mathbf{D}^l (\mathbf{C} - \mathbf{E})^T (\mathbf{C} - \mathbf{E})^* \mathbf{D}^{-l}\right] \tag{16}$$

we obtain [14]

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_u)M_R} \prod_{i=0}^{r(\mathbf{C}_u)M_R-1} \theta_i^{-1} \prod_{i=0}^{r(\mathbf{C}_u)-1} \lambda_i^{-M_R}(\mathbf{C}_u), \tag{17}$$

where $0 < \lambda_{min}(\widehat{\mathbf{R}}) \leq \theta_i \leq \lambda_{max}(\widehat{\mathbf{R}})$ with $\widehat{\mathbf{R}} = \text{diag}\{\mathbf{R}_l\}_{l \in \mathcal{L}}$. Noting that the diversity order achieved by a space-frequency code in the i.i.d. case is given by $M_R \min_{\mathbf{C}_u} r(\mathbf{C}_u)$ with the minimum taken over all codeword pairs $\{\mathbf{C}, \mathbf{E}\}$, it follows that the diversity order achieved by a space-frequency code in the presence of receive correlation only with all correlation matrices $\mathbf{R}_l\,(l = 0, 1, ..., L-1)$ full rank is equal to the diversity order achieved by the code in the absence of spatial fading correlation. The corresponding coding gain is given by the coding gain achieved in the uncorrelated case $\left(\prod_{i=0}^{r(\mathbf{C}_u)-1} \lambda_i(\mathbf{C}_u)\right)^{\frac{1}{r(\mathbf{C}_u)}}$ multiplied by $\left(\prod_{i=0}^{r(\mathbf{C}_u)M_R-1} \theta_i\right)^{\frac{1}{r(\mathbf{C}_u)M_R}}$. In the special instance where the space-frequency code achieves full diversity gain in the i.i.d. case, i.e., the minimum rank of $\mathbf{C}_u$ over all pairs of codeword matrices $\{\mathbf{C}, \mathbf{E}\}$ is $M_T|\mathcal{L}|$, we get

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left(\frac{E_s}{4\sigma_n^2}\right)^{-M_T M_R|\mathcal{L}|} \left(\det(\widehat{\mathbf{R}})\right)^{-M_T} \prod_{i=0}^{M_T|\mathcal{L}|-1} \lambda_i^{-M_R}(\mathbf{C}_u). \tag{18}$$

From the discussion above and (18) it follows that the diversity order achieved in this case is $d = M_T M_R|\mathcal{L}|$, which is the maximum possible diversity order achievable in the i.i.d. case [13]. The coding gain is given by the coding gain achieved in the i.i.d. case multiplied by $\left(\det(\widehat{\mathbf{R}})\right)^{\frac{1}{M_R|\mathcal{L}|}}$. Using the normalization $\text{Tr}(\mathbf{R}) = M_R|\mathcal{L}|$, it is easily seen that $\det(\widehat{\mathbf{R}}) \leq 1$ where the upper bound is achieved if all the $\mathbf{R}_l$ with $l \in \mathcal{L}$ are unitary. Thus, we conclude that spatial fading correlation always leads to a loss in performance with the degradation being determined by the eigenvalue spread of $\widehat{\mathbf{R}}$.

Finally, if $\widehat{\mathbf{R}}$ is rank-deficient (which is the case if at least one of the $\mathbf{R}_l$ is rank-deficient) and the space-frequency code achieves full diversity gain in the i.i.d. case, the achievable diversity order is upper-bounded as [14]

$$d \leq M_T \sum_{l \in \mathcal{L}} r(\mathbf{R}_l). \tag{19}$$

Transmit correlation only. In the case of transmit correlation only $\mathbf{R}_l = \mathbf{I}_{M_R}$ for $l = 0, 1, \ldots, L-1$. For nonsingular $\widehat{\mathbf{S}} = \text{diag}\{\mathbf{S}_l\}_{l \in \mathcal{L}}$ we have

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left(\frac{E_s}{4\sigma_n^2}\right)^{-M_R r(\mathbf{C}_u)} \prod_{i=0}^{r(\mathbf{C}_u)-1} \theta_i^{-M_R} \prod_{i=0}^{r(\mathbf{C}_u)-1} \lambda_i^{-M_R}(\mathbf{C}_u), \tag{20}$$

where $0 < \lambda_{min}(\widehat{\mathbf{S}}) \leq \theta_i \leq \lambda_{max}(\widehat{\mathbf{S}})$, which implies that the diversity order is given by $d = M_R \min_{\mathbf{C}_u} r(\mathbf{C}_u)$ with the minimum taken over all codeword pairs $\{\mathbf{C}, \mathbf{E}\}$. Hence the achievable

diversity order in the case of transmit correlation only with full-rank $\widehat{\mathbf{S}}$ is equal to the diversity order achieved by the space-frequency code in the i.i.d. case. For general $\widehat{\mathbf{S}}$, focusing on a minimum-rank error event $\mathbf{C}-\mathbf{E}$ (i.e. an error event that leads to minimum $r(\mathbf{C}_u)$), we get

$$M_R(r(\mathbf{C}_u)+r(\widehat{\mathbf{S}})-M_T|\mathcal{L}|) \;\leq\; d \;\leq\; M_R \min\{r(\mathbf{C}_u), r(\widehat{\mathbf{S}})\}. \tag{21}$$

We note that (21) shows that if $\widehat{\mathbf{S}}$ is singular and the space-frequency code does not achieve full diversity gain in the i.i.d. case, the diversity order achieved in the correlated case can only be lower bounded by $M_R(r(\mathbf{C}_u)+r(\widehat{\mathbf{S}})-M_T|\mathcal{L}|)$. In this case, it is difficult to make statements on the exact diversity order achieved since d will be a function of the eigenspaces of the $\mathbf{S}_l$ and the eigenspace of $(\mathbf{C}-\mathbf{E})^*(\mathbf{C}-\mathbf{E})^T$. In general, as evidenced by (21) in the presence of transmit correlation, it is crucial that the space-frequency code excites the range spaces of the $\mathbf{S}_l$ in order to obtain good performance in terms of error probability.

Joint transmit-receive correlation. In the case of joint transmit-receive correlation, it follows from (14) that the maximum achievable diversity order is given by $d=\sum_{l\in\mathcal{L}} r(\mathbf{S}_l)r(\mathbf{R}_l)$. Noting that the diversity order offered by the l-th tap is $r(\mathbf{R}_l)r(\mathbf{S}_l)$, this result says that the maximum achievable diversity order is simply the number of total degrees of freedom available in the channel.

Ricean Fading

So far, we have restricted our attention to Rayleigh fading. Let us next consider the Ricean case. Again assuming high SNR ($\frac{E_s}{4\sigma_n^2} \gg 1$), the PEP upper bound in (13) simplifies to

$$\mathbf{P}(\mathbf{C}\rightarrow\mathbf{E}) \leq \exp\left\{\frac{E_s}{4\sigma_n^2}\left(\sum_{i=0}^{r(\mathbf{C}_Y)-1}|b_i|^2\lambda_i(\mathbf{C}_Y)-\|\bar{\mathbf{Y}}\|^2\right)\right\}\left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_Y)}\prod_{i=0}^{r(\mathbf{C}_Y)-1}\lambda_i^{-1}(\mathbf{C}_Y). \tag{22}$$

We can see that the PEP upper bound consists of an exponential term multiplied by the RHS of (15), which is the PEP upper bound in the Rayleigh fading case. The behavior of the second and third terms of the RHS in (22) was studied above in the section on Rayleigh fading. We shall therefore focus on the first term of the RHS of (22), which represents the contribution due to the Ricean component, and start by noting that

$$\sum_{i=0}^{r(\mathbf{C}_Y)-1}|b_i|^2\lambda_i(\mathbf{C}_Y)-\|\bar{\mathbf{Y}}\|^2 = \bar{\mathbf{Y}}^H\mathbf{U}\,\mathrm{diag}\{\underbrace{0,0,\ldots,0}_{r(\mathbf{C}_Y)},-1,-1,\ldots,-1\}\,\mathbf{U}^H\bar{\mathbf{Y}},$$

which after application of the Rayleigh-Ritz Theorem [21] yields

$$e^{-\frac{E_s}{4\sigma_n^2}\|\bar{\mathbf{Y}}\|^2} \leq \exp\left\{\frac{E_s}{4\sigma_n^2}\left(\sum_{i=0}^{r(\mathbf{C}_Y)-1}|b_i|^2\lambda_i(\mathbf{C}_Y) - \|\bar{\mathbf{Y}}\|^2\right)\right\} \leq 1. \tag{23}$$

We can, therefore, conclude that the performance in the Ricean case depends strongly on the angle between $\bar{\mathbf{Y}}$ and the eigenspace of $\mathbf{C}_Y$. Assuming that $\bar{\mathbf{Y}}$ is such that the lower bound in (23) is achieved, we get

$$\mathbf{P}(\mathbf{C} \to \mathbf{E}) \leq e^{-\frac{E_s}{4\sigma_n^2}\|\bar{\mathbf{Y}}\|^2}\left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_Y)}\prod_{i=0}^{r(\mathbf{C}_Y)-1}\lambda_i^{-1}(\mathbf{C}_Y).$$

In this case, the PEP upper bound is minimized if $\|\bar{\mathbf{Y}}\|^2$ is maximized, which implies that the error probability performance is determined by the properties of the matrices $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ $(k = 0, 1, \ldots, N-1)$ and, hence, the matrices $\bar{\mathbf{H}}_l$. In particular, we note that error events with code difference vectors lying in the null-spaces of the $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ will yield small $\|\bar{\mathbf{Y}}\|^2$ and hence result in high error probability. For high-rate codes such as spatial multiplexing it is therefore crucial for the matrices $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ to have high rank in order to ensure good performance in terms of error probability. Let us first investigate the flat-fading case where $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k}) = \bar{\mathbf{H}}_0$ for $k = 0, 1, \ldots, N-1$. Now, using (6) it follows that $\bar{\mathbf{H}}_0$ has high rank and is well-conditioned if the transmit and receive angle spreads and the number of paths contributing to $\bar{\mathbf{H}}_0$ are sufficiently large. We can therefore expect that large P_0 and large transmit and receive angle spreads will lead to small error probability. In the frequency-selective case the matrices $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k}) = \sum_{l=0}^{L-1}\bar{\mathbf{H}}_l\, e^{-j\frac{2\pi}{N}kl}$ can have full rank and be well-conditioned even if the individual matrices $\bar{\mathbf{H}}_l$ are rank-deficient (or equivalently the individual cluster angle spreads are small), and the performance will depend on the total number of paths $\sum_{l=0}^{L-1}P_l$ and the angle spread measured over all the taps (i.e. the total angle spread). We can therefore conclude that in the MIMO Ricean case presence of frequency-selectivity can be beneficial as it can help to restore the rank of the $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ matrices. The situation is fundamentally different in the SISO case, where frequency-selectivity causes dips in the frequency response, and the performance degrades compared to the flat-fading case.

Low rate codes, such as space-frequency block codes, will be less sensitive to the rank and the condition number of the matrices $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$, as there will be fewer code difference vectors that tend to lie in the null-spaces of the $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ and hence cause high error probability. More discussions and conclusions along these lines can be found in Simulation Example 4 in Section 5.

5 Impact of Propagation Parameters on Space-Frequency Codes

We have seen in the previous section that the rank and the eigenvalues of the individual correlation matrices $\mathbf{S}_l$ and $\mathbf{R}_l$ determine the diversity gain and the coding gain achieved by a space-frequency code. In this section, we shall first relate angle spread and antenna spacing to the eigenvalues of the correlation matrices, and then use these results to study the impact of the propagation environment on the performance of space-frequency codes.

Impact of Propagation Parameters

Impact of cluster angle spread. Let us restrict our attention to the transmit correlation matrix $\mathbf{S}_l$ (the same analysis applies to the $\mathbf{R}_l$). Since $\mathbf{S}_l$ is Toeplitz its limiting ($M_T \to \infty$) eigenvalue distribution can be obtained as $\lambda_l(\nu) = \vartheta_3\left(\pi(\nu - \Delta_T \cos(\bar{\theta}_{T,l})), e^{-\frac{1}{2}(2\,\pi\,\Delta_T\,\sin(\bar{\theta}_{T,l})\sigma_{\theta_{T,l}})^2}\right)$ with the third-order theta function defined in (11). Similar to the discussion in Section 3, it follows that $\lambda_l(\nu)$ 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 (i.e. high spatial fading correlation) $\lambda_l(\nu)$ will be peaky. For any M_T, using the fact that $\mathbf{S}_l = \mathbf{S}_l^H$, the eigenvalues of $\mathbf{S}_l$ can be lower and upper-bounded by the infimum and the supremum of $\lambda_l(\nu)$. In particular, defining

$$m = \operatorname*{ess\,inf}_{\nu\,\in\,[0,1)} \lambda_l(\nu), \qquad\qquad M = \operatorname*{ess\,sup}_{\nu\,\in\,[0,1)} \lambda_l(\nu)$$

we have that [22]

$$m \;\leq\; \lambda_i(\mathbf{S}_l) \;\leq\; M \qquad i = 0, 1, \ldots, M_T - 1.$$

This result allows to provide an upper bound on the PEP in terms of $\lambda_l(\nu)$, and, hence, makes the dependence of the PEP on cluster angle spread and antenna spacing more explicit. For example, assuming flat-fading (i.e., $L = 1$) and a nonsingular $\mathbf{S}_0$, we get from (20) that

$$P(\mathbf{C} \to \mathbf{E}) \;\leq\; \left(\frac{mE_s}{4\sigma_n^2}\right)^{-M_R\,r(\mathbf{C}_u)} \prod_{i=0}^{r(\mathbf{C}_u)-1} \lambda_i^{-M_R}(\mathbf{C}_u),$$

where $\mathbf{C}_u = \sigma_0^2(\mathbf{C}-\mathbf{E})^T(\mathbf{C}-\mathbf{E})^*$. Now, using the normalization $\mathrm{Tr}(\mathbf{S}_0) = M_T$, it follows that $m \leq 1$ with $m = 1$ if and only if $\mathbf{S}_0$ is unitary or equivalently spatial fading is uncorrelated. The presence of spatial fading correlation can therefore be interpreted as a reduction in the effective SNR by a factor of m.

Impact of total angle spread. We shall next investigate the impact of total angle spread on the performance of space-frequency codes. Let us start with the case of receive correlation only and assume that the receive antenna spacing and/or the receive cluster angle spreads are small so that the individual receive correlation matrices $\mathbf{R}_l$ are rank-deficient. Consider the extreme scenario $\sigma_{\theta_{R,l}} = 0$ for $l = 0, 1, \ldots, L-1$, where $\mathbf{R}_l = \mathbf{a}(\bar{\theta}_{R,l})\mathbf{a}^H(\bar{\theta}_{R,l})$ $(l = 0, 1, \ldots, L-1)$. Next, assume that the $\mathbf{R}_l$ span mutually orthogonal subspaces, i.e., the $\mathbf{a}(\bar{\theta}_{R,l})$ are orthogonal to each other and hence $\mathbf{R}_l^{1/2}\mathbf{R}_{l'}^{1/2} = \mathbf{0}$ for $l \neq l'$. Of course, for this to hold we need $L \leq M_R$. Roughly speaking mutual orthogonality of the $\mathbf{a}(\bar{\theta}_{R,l})$ requires that either the total receive angle spread be sufficiently large or M_R is large so that the receive array provides high spatial resolution. Now, if the correlation matrices $\mathbf{R}_l$ indeed span mutually orthogonal subspaces, it follows that $\sigma(\mathbf{C}_Y) = \left\{\sigma(\widetilde{\mathbf{C}}_Y), 0, ..., 0\right\}$ with [14]

$$\widetilde{\mathbf{C}}_Y = \mathrm{diag}\{\sigma_l^2[(\mathbf{C}-\mathbf{E})^*(\mathbf{C}-\mathbf{E})^T] \otimes \mathbf{R}_l\}_{l \in \mathcal{L}}. \tag{24}$$

The block-diagonality of $\widetilde{\mathbf{C}}_Y$ implies several interesting properties. First, we note that (24) immediately yields $d = d_f \sum_{l \in \mathcal{L}} r(\mathbf{R}_l)$, where d_f is the diversity order achieved by the space-frequency code in the flat-fading spatially uncorrelated case. Moreover, we can infer from (24) that the space-frequency code design criteria reduce to the well-known space-time code design criteria first reported in [2]. Hence, in this case space-frequency code design for the broadband case reduces to classical space-time code design for the narrowband case. This result has a physically intuitive explanation which is as follows. Start from the i.i.d. case (i.e. $\mathbf{R}_l = \mathbf{I}_{M_R}$ for $l = 0, 1, \ldots, L-1$), where the design criteria are the classical rank and determinant criteria [2] applied to the stacked matrix $\left[\sqrt{\sigma_0^2}\,(\mathbf{C}-\mathbf{E})^T \;\; \ldots \;\; \sqrt{\sigma_{L-1}^2}\,\mathbf{D}^{L-1}(\mathbf{C}-\mathbf{E})^T\right]$. In order to achieve good performance in terms of error probability we need the columns of $(\mathbf{C}-\mathbf{E})^T$ to be "as orthogonal as possible" to each other and the columns of $\mathbf{D}^l(\mathbf{C}-\mathbf{E})^T$ to be "as orthogonal as possible" to the columns of $\mathbf{D}^{l'}(\mathbf{C}-\mathbf{E})^T$ for $l \neq l'$. Next, write $(\mathbf{C}-\mathbf{E})^T = \mathbf{F}(\mathbf{C}_t-\mathbf{E}_t)^T$ where $\mathbf{F}$ with $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}mn}$ is the $N \times N$ FFT matrix and exploit $\mathbf{F}^{-1}\mathbf{D}^l\mathbf{F}(\mathbf{C}_t-\mathbf{E}_t)^T = (\mathbf{C}_{t-l}-\mathbf{E}_{t-l})^T$, where $(\mathbf{C}_{t-l}-\mathbf{E}_{t-l})$ denotes the matrix obtained

by cyclically shifting the rows of $(\mathbf{C}_t - \mathbf{E}_t)$ by l positions to the right. The design criteria can now be rephrased (exploiting the unitarity of the FFT) as the columns of $(\mathbf{C}_t - \mathbf{E}_t)^T$ being "as orthogonal as possible" to each other and the columns of $(\mathbf{C}_{t-l} - \mathbf{E}_{t-l})^T$ being "as orthogonal as possible" to the columns of $(\mathbf{C}_{t-l'} - \mathbf{E}_{t-l'})^T$ for $l \neq l'$. This means that the cyclically shifted versions of $(\mathbf{C}_t - \mathbf{E}_t)^T$ should be "as orthogonal as possible" to each other, so that the receiver can separate them and exploit frequency diversity gain. This can be achieved by judicious space-frequency code design [23]. Now, in the case of receive correlation only where the $\mathbf{R}_l$ span mutually orthogonal subspaces, the separation of the delayed replicas is provided by the channel as the different delayed versions excite mutually orthogonal subspaces. Roughly speaking, if the total receive angle spread is large or M_R is large, the spatial component of the channel orthogonalizes the delayed versions of the transmitted signal, and frequency-diversity gain is provided for free (assuming ML decoding).

We shall finally discuss the impact of total transmit angle spread on the performance of space-frequency codes in the case of transmit correlation only. As already mentioned, in the presence of transmit correlation, the performance of space-frequency codes depends very much on the geometry of the code relative to the geometry of the transmit correlation matrices. Let us start by considering the flat-fading case (i.e., $L = 1$), where $\mathbf{C}_Y = \sigma_0^2 (\mathbf{C} - \mathbf{E})^T \mathbf{S}_0 (\mathbf{C} - \mathbf{E})^*$. Now, assume that either the transmit cluster angle spread and/or the transmit antenna spacing is small so that $\mathbf{S}_0$ is rank-deficient. In this case, the error events $(\mathbf{C} - \mathbf{E})^*$ lying in the null space of $\mathbf{S}_0$ will cause large PEP. Next, consider the case of multi-path propagation where $L > 1$ and retain the assumption of small cluster angle spreads and/or small transmit antenna spacing. If the total transmit angle spread is large, the $\mathbf{S}_l$ will span different subspaces and hence have different null spaces. Consequently, if $(\mathbf{C} - \mathbf{E})^*$ happens to lie in the null space of one of the $\mathbf{S}_l$, it can still lie in the range of one of the other $\mathbf{S}_l$ and hence using (14) it follows that the PEP performance may still be good. This effect will be studied further and quantified in Simulation Example 1 below.

Summarizing, we can conclude that increased total angle spread not only leads to an increased ergodic capacity as observed in Section 3, but generally also to improved performance from an error probability point of view.

Simulation Results

We conclude this section with simulation results that corroborate and quantify some of the analytical results provided in this chapter. Unless specified otherwise, we simulated a space-frequency coded MIMO-OFDM system with $M_T = M_R = 2$, $L = 2$, $\Delta_T = \Delta_R = 1, \sigma_0^2 = \sigma_1^2 = \frac{1}{2}$, $N = 32$ tones, no **forward error correction**, and ML decoding. The SNR is defined as SNR $= 10 \log_{10}\left(\frac{M_T E_s}{\sigma_n^2}\right)$. In all simulations the channel was normalized to satisfy $\sum_{l \in \mathcal{L}} \mathcal{E}\{\|\mathbf{H}_l\|_F^2\} = M_T M_R$. We employed a space-frequency coding scheme where an arbitrary (inner) space-time code is used to transmit data in the first $\frac{N}{2}$ tones followed by a repetition of this block. This construction ensures that the additionally available frequency diversity is fully exploited [23].

Simulation Example 1. In this simulation example, we study the impact of total transmit angle spread $\Delta\bar{\theta}_T = \bar{\theta}_{T,1} - \bar{\theta}_{T,0}$ in the presence of high transmit correlation (caused by small cluster angle spreads) on the performance of space-frequency codes. As inner codes we used QPSK-based spatial multiplexing [5, 6, 9], which amounts to transmitting independent data symbols on each tone and each antenna, and the Alamouti scheme [3] based on QPSK. The transmit cluster angle spreads were chosen as $\sigma_{\theta_{T,0}} = \sigma_{\theta_{T,1}} = 0$ so that $\mathbf{S}_l = \mathbf{a}(\bar{\theta}_{T,l})\mathbf{a}^H(\bar{\theta}_{T,l})$ $(l = 0, 1)$. The receive antennas were assumed to fade in an uncorrelated fashion (i.e., $\mathbf{R}_0 = \mathbf{R}_1 = \mathbf{I}_2$). The top diagram in Fig. 6 shows the block error rate as a function of $\Delta\bar{\theta}_T$ with $\bar{\theta}_{T,0} = 0$ for spatial multiplexing at an SNR of 14 dB and for the Alamouti scheme at an SNR of 7 dB, respectively. We can clearly see that spatial multiplexing is very sensitive to total transmit angle spread and that the performance improves significantly for the case where the array response vectors $\mathbf{a}(\bar{\theta}_{T,0})$ and $\mathbf{a}(\bar{\theta}_{T,1})$ are close to orthogonal to each other. This is reflected by displaying the condition number of the 2×2 matrix $\mathbf{B} = [\mathbf{a}(\bar{\theta}_{T,0})\ \ \mathbf{a}(\bar{\theta}_{T,1})]$ in the bottom diagram of Fig. 6. The lower rate Alamouti scheme is virtually unaffected by the total transmit angle spread, which is due to the fact that it orthogonalizes the channel irrespectively of the channel realization and hence its performance is independent of the rank properties of the channel. For spatial multiplexing the situation is different since the performance depends critically on the rank of the channel realizations. To be more specific, we note that each transmit correlation matrix $\mathbf{S}_l$ spans a one-dimensional subspace with the angle between these subspaces being a function of total angle spread (as evidenced by the bottom plot of Fig. 6). If the total transmit angle spread

is small, the two subspaces tend to be aligned so that transmit signal vectors which happen to lie in the orthogonal complement of this subspace will cause high error probability. If the total angle spread is large, the $\mathbf{S}_l$ tend to span different subspaces. Hence if a transmit vector excites the null space of $\mathbf{S}_0$, it will lie in the range space of $\mathbf{S}_1$ and vice versa. This leads to good performance in terms of error rate as none of the transmit vectors will be attenuated by the channel (on average, where the average is with respect to the random channel).

Simulation Example 2. This example investigates the impact of transmit and receive antenna correlation on the performance of space-frequency codes. Fig. 7 displays the block error rate for a system employing the two-antenna 16-state QPSK Trellis code proposed in [2] as inner code. For $M_T = 2$ and $M_R = 3$, we show the cases of no correlation (i.i.d. channel), transmit correlation only ($\Delta_T = 0.1, \sigma_{\theta_{T,0}} = \sigma_{\theta_{T,1}} = 0.25, \bar{\theta}_{T,0} = \bar{\theta}_{T,1} = \pi/4$), and receive correlation only ($\Delta_R = 0.1, \sigma_{\theta_{R,0}} = \sigma_{\theta_{R,1}} = 0.25, \bar{\theta}_{R,0} = \bar{\theta}_{R,1} = \pi/4$), respectively. We note that this choice of channel parameters results in $|\rho(\Delta, \bar{\theta}_l, \sigma_{\theta_l})| = 0.994$ and $|\rho(2\Delta, \bar{\theta}_l, \sigma_{\theta_l})| = 0.976$ for $l = 0, 1$, which basically amounts to fully correlated spatial fading. Fig. 7 shows that best performance is achieved in the i.i.d. case (as expected), performance degrades in the presence of transmit correlation only, and degrades even more in the presence of receive correlation only. This asymmetry between the transmit and receive correlation cases can be explained as follows. In the i.i.d. case the channel offers diversity order $d = M_T M_R L = 12$. For the correlated cases we make the simplifying assumption $|\rho(\Delta, \bar{\theta}_l, \sigma_{\theta_l})| = |\rho(2\Delta, \bar{\theta}_l, \sigma_{\theta_l})| = 1$ for $l = 0, 1$ which implies $r(\mathbf{S}_0) = r(\mathbf{S}_1) = r(\mathbf{R}_0) = r(\mathbf{R}_1) = 1$. In the case of transmit correlation only, the diversity order follows from (21) as $d = M_R(r(\mathbf{S}_0)+r(\mathbf{S}_1)) = 6$, whereas in the case of receive correlation only (19) yields $d \leq M_T(r(\mathbf{R}_0) + r(\mathbf{R}_1)) = 4$. These differences in the achievable diversity order are reflected in the error rate performance depicted in Fig. 7. While not shown in the diagram, we emphasize that the performance gap between transmit correlation only and receive correlation only is even more pronounced for higher SNR.

Simulation Example 3. This simulation example compares the performance of orthogonal and nonorthogonal space-frequency codes in the presence of transmit correlation[3]. We have seen above that the impact of transmit antenna correlation is highly dependent on how the code excites

[3]A space-frequency code is referred to as orthogonal or nonorthogonal if the inner space-time code is orthogonal or nonorthogonal, respectively.

the channel (i.e., the code geometry) relative to the channel geometry. Given that the transmitter does not know the channel, an orthogonal scheme such as the Alamouti scheme, which excites all spatial directions uniformly, should exhibit maximum robustness with respect to this kind of channel impairment. Nonorthogonal schemes such as spatial multiplexing do not excite all spatial directions uniformly and hence suffer from widely varying performance losses in the presence of transmit correlation. In order to demonstrate this effect, we compared a 16-QAM based Alamouti scheme as a simple orthogonal inner code to spatial multiplexing based on QPSK (nonorthogonal inner code). Note that this choice of symbol constellations ensures that the transmission rate is the same in both cases. The SNR was kept constant at 12 dB while the absolute value of the transmit antenna correlation coefficient was varied from 0 to 1. The correlation matrix was chosen to be identical for both taps. Fig. 8 shows the resulting bit error rates assuming Gray encoding. It is clearly seen that whilst with uncorrelated transmit antennas the Alamouti scheme performs inferior compared to spatial multiplexing, performance degradation with increasing correlation is much worse for spatial multiplexing such that at full correlation the Alamouti scheme performs significantly better. These performance differences along with the results of Simulation Example 1 lead us to the conclusion that space-frequency block codes exhibit superior robustness with respect to varying propagation conditions when compared to (spatially) high-rate schemes such as spatial multiplexing.

Simulation Example 4. The last simulation example studies the impact of K-factor and geometry of the Ricean component on the performance of space-frequency codes. We compared a 2 bits/tone spatial multiplexing BPSK code to a simple BPSK-based delay diversity scheme with delay 4 (achieved by premultiplying the signals with a linear phase factor in the frequency domain). In both cases, we used a rate 1/2, constraint length-4 convolutional code with generator polynomials 15 and 17, and random interleaving. Hence, the overall transmission rate for delay diversity was half the transmission rate for spatial multiplexing. We emphasize that the difference in overall transmission rate results from the fact that the spatial rate of delay diversity (1 symbol per vector transmission) is half the spatial rate of spatial multiplexing (2 symbols per vector transmission). This ensures that the performance differences between the two schemes are due to the different spatial signaling structures and rates. The receiver for spatial multiplexing first separates the spatial signals using an

MMSE front-end and then performs soft Viterbi decoding, whereas in the case of delay diversity a soft Viterbi decoder only was used (spatial separation is not required as the two transmit antennas are effectively collapsed into one antenna). The matrices $\bar{\mathbf{H}}_l$ were generated using the parametric model described in (6). The four cases depicted in Figs. 9 and 10, respectively, are:

- (Scenario 1) Pure Rayleigh fading on both taps, i.e., $K_0 = K_1 = 0$.
- (Scenario 2) In this and the remaining cases $K_0 = K_1 = 16$dB. Only one significant path is present in each of the two taps, i.e., $P_0 = P_1 = 1$, therefore $r(\bar{\mathbf{H}}_0) = r(\bar{\mathbf{H}}_1) = 1$. The angle spread across the two taps is zero so that $\bar{\mathbf{H}}_0 = \bar{\mathbf{H}}_1$. Due to the finite constellation used here, the specific realization of the Ricean channel component has a large impact on performance. In order to make our conclusions representative, we average over a uniform distribution of angles of arrival and departure (which determine the Ricean channel component) between 0 and 2π.
- (Scenario 3) Again both taps have only one significant path, but there is an angle spread of $\frac{\pi}{2}$ across the taps, such that $\bar{\mathbf{H}}_0$ and $\bar{\mathbf{H}}_1$ are still rank 1, but not identical any more. Also, we again average over a uniform angle of arrival and departure distribution, but the angle *difference* between the two taps is kept constant at $\frac{\pi}{2}$ for both the angle of arrival and departure.
- (Scenario 4) Flat-fading scenario (i.e., $L = 1$), but two paths adding up at this single tap, i.e., $P_0 = 2$ resulting in a rank-2 matrix $\bar{\mathbf{H}}_0$. The angles of arrival and departure were chosen such that $\bar{\mathbf{H}}_0$ is orthogonal.

Spatial multiplexing. From Fig. 9 it is clearly seen that depending on the propagation environment a high K-factor can either be beneficial or detrimental. In Scenario 2, the matrices $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ have rank 1 for $k = 0, 1, \ldots, N-1$, thus severely impeding spatial multiplexing performance. Scenario 4 leads to good performance since $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k}) = \bar{\mathbf{H}}_0$ is orthogonal for $k = 0, 1, ..., N-1$. Finally, we can see that in Scenario 3 we obtain significantly better performance than in Scenario 2, which is due to the fact that the high angle spread across the two taps results in full-rank $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$, even though $\bar{\mathbf{H}}_0$ and $\bar{\mathbf{H}}_1$ each have rank 1. In Scenario 3 if only one tap were present the $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ $(k = 0, 1, ..., N-1)$ would be rank-deficient and hence the performance would be comparable to Scenario 2. It is therefore interesting to observe that the presence of frequency-selectivity (provided

that the delayed paths increase the total angle spread) can improve performance in the Ricean MIMO case. This is in contrast to the SISO Ricean case where frequency-selectivity leads to a performance degradation when compared to the frequency-flat case.

Delay diversity combined with convolutional code. The same trends are true for the lower rate delay diversity scheme, but the differences are less pronounced. In particular, Scenario 2 exhibits a much smaller performance loss compared to the Rayleigh fading case. This is due to the fact that since we are dealing with a scheme with half the spatial signaling rate, the rank properties and the condition numbers of the $\bar{\mathbf{H}}(e^{j\frac{2\pi}{N}k})$ are less important than in the case of spatial multiplexing.

Defining Terms

Diversity gain: Improvement in link reliability obtained by transmitting the same data on independently fading branches.

Fading: Fluctuation in the signal level due to shadowing and multipath effects.

FFT: Fast Fourier Transform.

Forward error correction (FEC): A technique that inserts redundant bits during transmission to help detect and correct bit errors during reception.

IFFT: Inverse Fast Fourier Transform.

Interleaving: A form of data scrambling that spreads bursts of bit errors evenly over the received data allowing efficient forward error correction.

Link reliability: Measure for the quality of a wireless communication link.

MIMO: Multiple-input multiple-output (MIMO) wireless communication system employing an antenna array at both ends of the link.

(Spatial) multiplexing gain: Capacity gain at no additional power or bandwidth consumption obtained through the use of multiple antennas at both ends of a wireless radio link.

Orthogonal Frequency Division Multiplexing (OFDM): Modulation scheme which divides the available frequency band into subcarriers (or tones) of smaller bandwidth and thereby drastically simplifies equalization.

Space-frequency coding: Coding technique that realizes spatial and frequency diversity gains without knowing the channel in the transmitter by spreading information across antennas (space) and frequency.

Space-time coding: Coding technique that realizes spatial diversity gain without knowing the channel in the transmitter by spreading information across antennas (space) and time.

Spatial multiplexing: Signaling technique to realize spatial multiplexing gain.

Further Information

The books [24–26] contain excellent introductions to and overviews of MIMO wireless and space-time coding. Four special issues have appeared on MIMO wireless in the past 2 years in the IEEE Trans. Signal Processing Oct. 2002 and Oct. 2003, IEEE Journal on Selected Areas in Communications April and June 2003, and IEEE Trans. Information Theory Oct. 2003.

References

[1] W. C. Jakes, *Microwave mobile communications*. New York: Wiley, 1974.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[4] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 543–564, 2000.

[5] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," *U. S. Patent*, no. 5,345,599, 1994.

[6] 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.

[7] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[8] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, 1998.

[9] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.

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

[11] 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.

[12] B. LeFloch, M. Alard, and C. Berrou, "Coded orthogonal frequency division multiplex," *Proc. of IEEE*, vol. 83, no. 6, pp. 982–996, June 1995.

[13] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE WCNC-2000*, Chicago, IL, Sept. 2000, pp. 1–6.

[14] H. Bölcskei, M. Borgmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 427–439, Apr. 2003.

[15] D. Asztély, "On antenna arrays in mobile communication systems: Fast fading and GSM base station receiver algorithms," Royal Institute of Technology, Stockholm, Sweden, Tech. Rep. IR-S3-SB-9611, Mar. 1996.

[16] M. Sandell, "Design and analysis of estimators for multicarrier modulation and ultrasonic imaging," Ph.D. dissertation, Lulea University of Technology, Lulea, Sweden, 1996.

[17] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications.* New York: Chelsea Publishing Company, 1984.

[18] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[19] V. Erceg, Private communication, Apr. 2001.

[20] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE GLOBECOM*, vol. 2, San Francisco, CA, Nov. 2000, pp. 1000–1004.

[21] R. A. Horn and C. R. Johnson, *Topics in matrix analysis.* New York: Cambridge Press, 1991.

[22] R. M. Gray, "Toeplitz and circulant matrices," Stanford University, ISL, Tech. Rep., Aug. 2002. [Online]. Available: http://ee-www.stanford.edu/~gray/toeplitz.html

[23] H. Bölcskei, M. Borgmann, and A. J. Paulraj, "Space-frequency coded MIMO-OFDM with variable multiplexing-diversity tradeoff," in *Proc. IEEE Int. Conf. Communications (ICC)*, Anchorage, AK, May 2003, pp. 2837–2841.

[24] A. J. Paulraj, R. U. Nabar, and D. A. Gore, *Introduction to space-time wireless communications.* Cambridge, UK: Cambridge Univ. Press, 2003.

[25] E. G. Larsson and P. Stoica, *Space-time block coding for wireless communications.* Cambridge, UK: Cambridge Univ. Press, 2003.

[26] G. B. Giannakis, Z. Liu, X. Ma, and S. Zhou, *Space-time coding for broadband wireless communications.* Wiley, 2004.

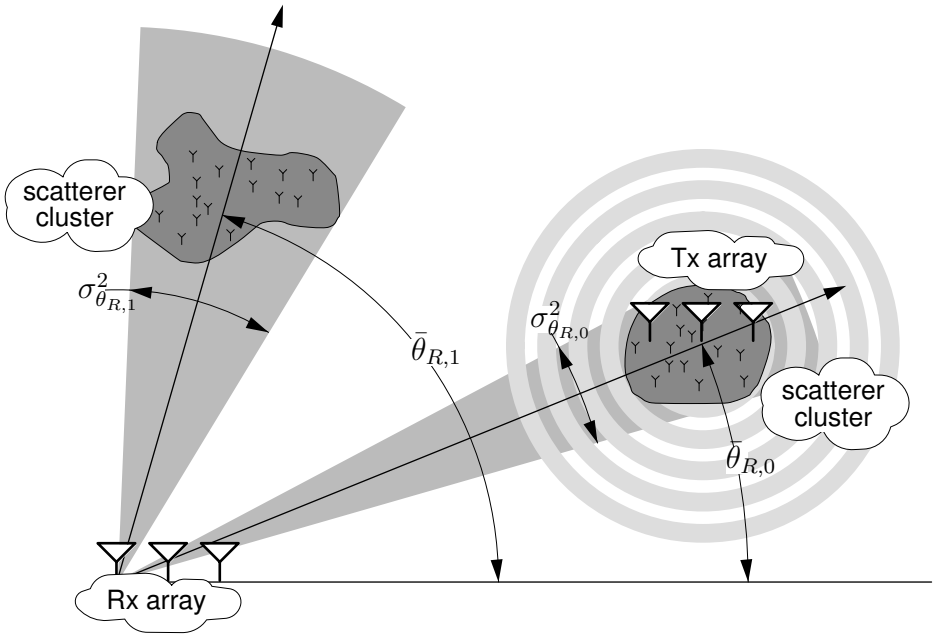


Figure 1: Schematic representation of MIMO broadband channel composed of multiple clustered paths. For simplicity, only the relevant angles for the receive array are shown — the transmit array situation is reciprocal. The two clusters correspond to different delay taps.

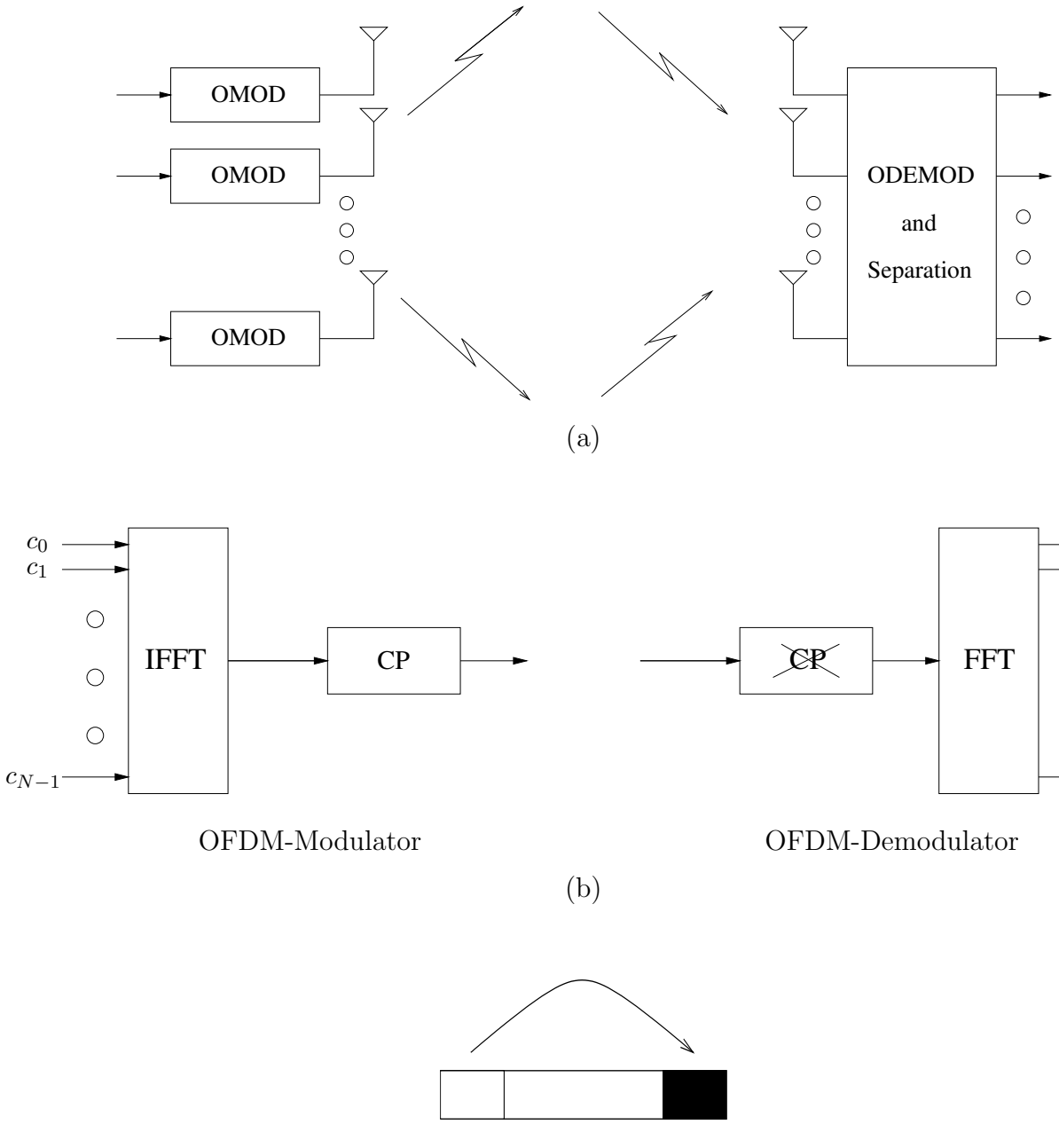


Figure 2: (a) Schematic of a MIMO-OFDM system. (OMOD and ODEMOD denote an OFDM-modulator and demodulator, respectively), (b) Single-antenna OFDM modulator and demodulator, (c) Adding the cyclic prefix.

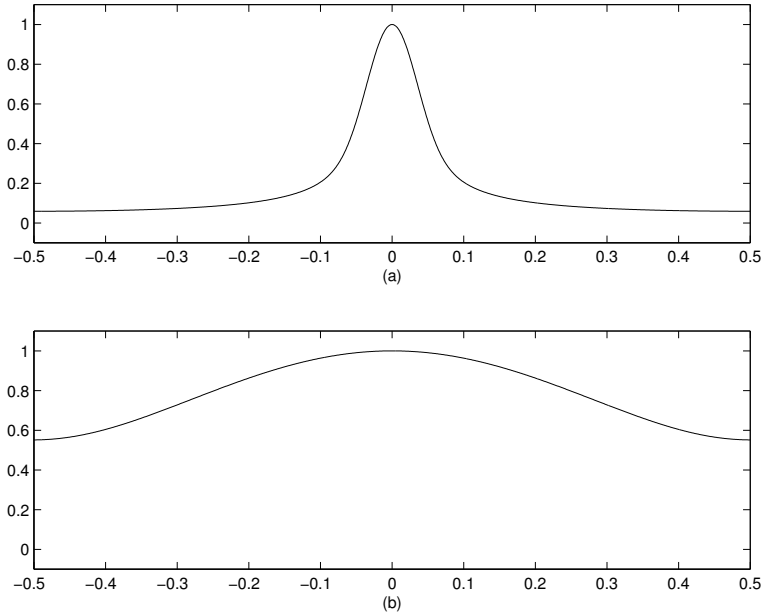


Figure 3: Limiting eigenvalue distribution of the correlation matrix $\mathbf{R}_0$ for the cases of (a) high spatial fading correlation, and (b) low spatial fading correlation.

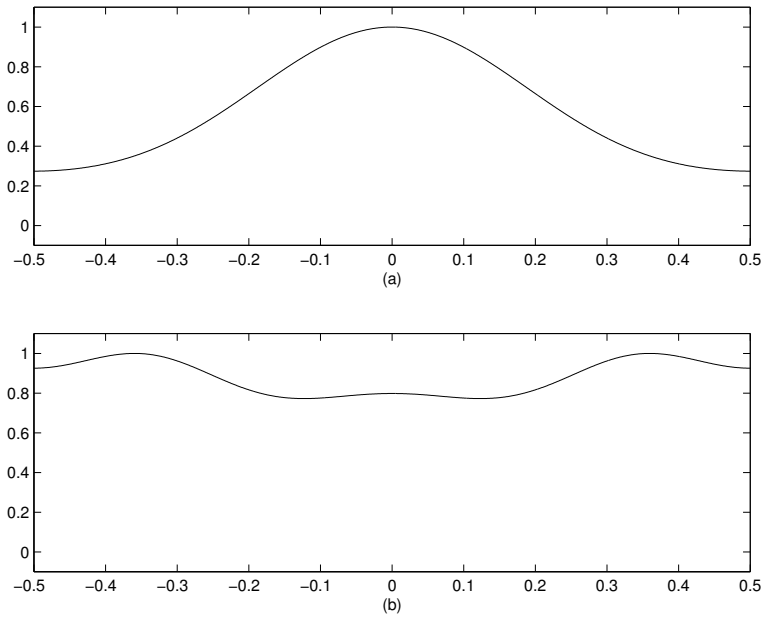
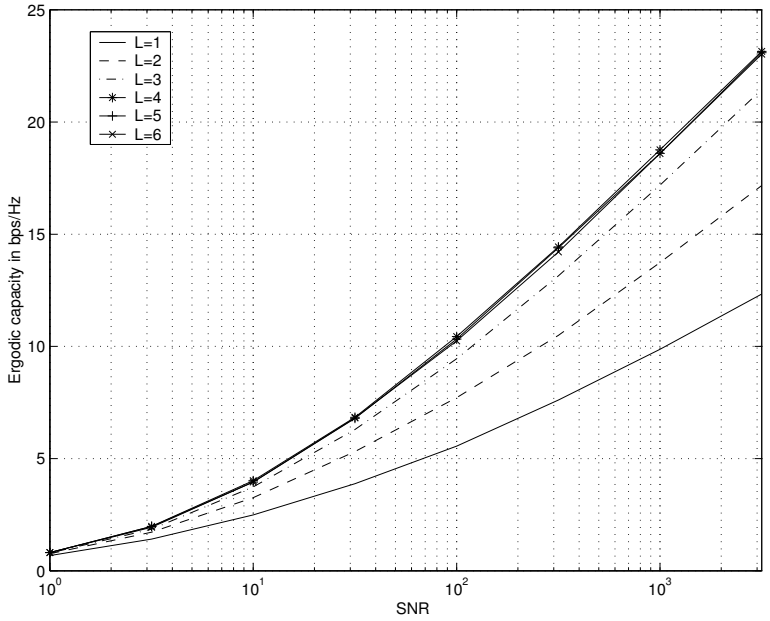
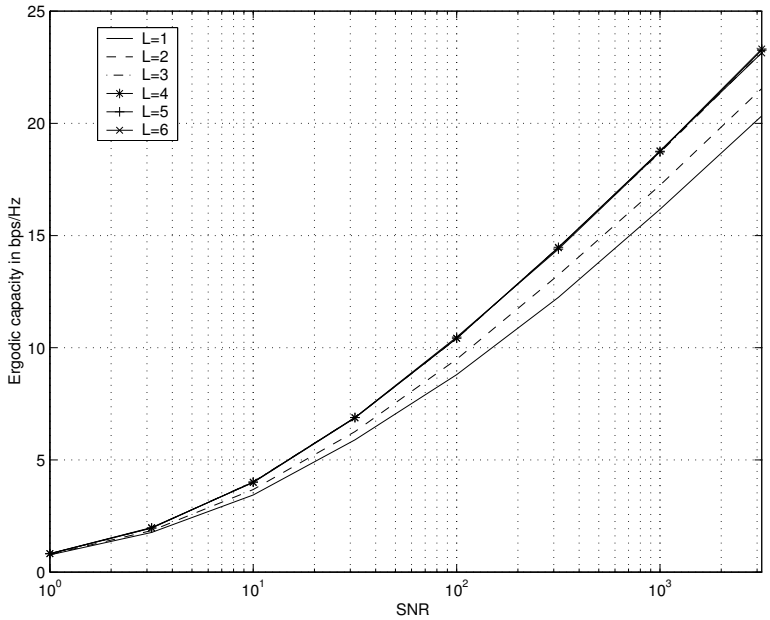


Figure 4: Limiting eigenvalue distribution of the sum correlation matrix $\mathbf{R} = \sum_{l=0}^{2} \sigma_l^2 \, \mathbf{R}_l$ for fixed cluster angle spread and for the cases of (a) small total angle spread, and (b) large total angle spread.



(a)



(b)

Figure 5: Ergodic capacity (in bps/Hz) as a function of SNR for various values of L and (a) small cluster angle spreads, and (b) large cluster angle spreads.

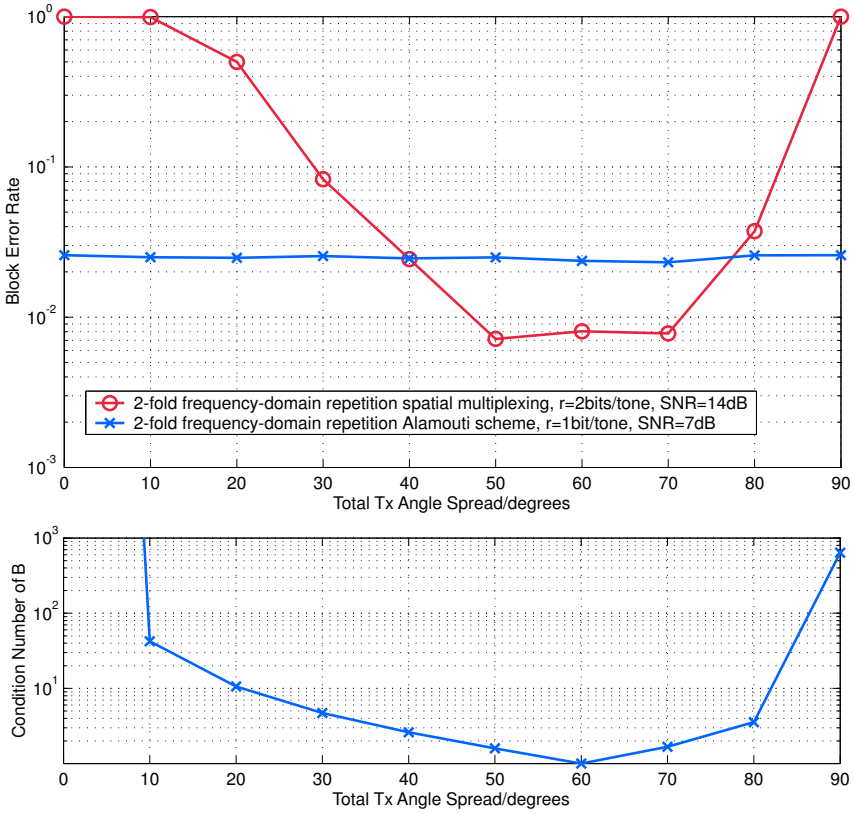


Figure 6: Impact of total transmit angle spread on performance of space-frequency codes.

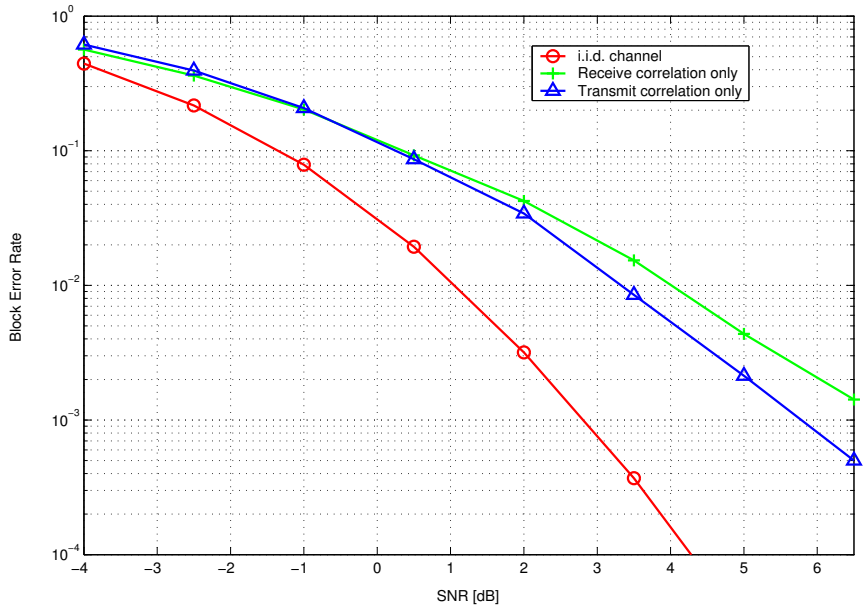


Figure 7: Impact of transmit and receive antenna correlation on the performance of space-frequency Trellis codes.

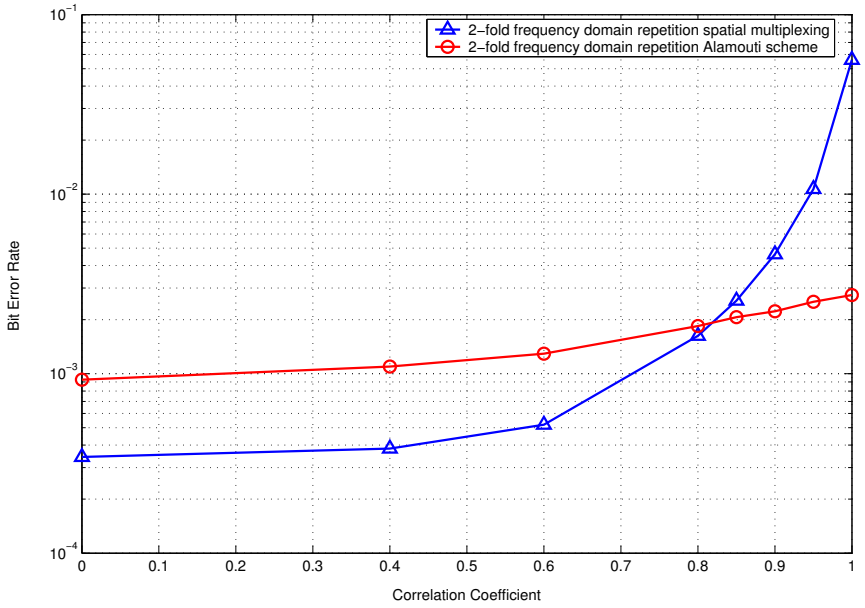


Figure 8: Impact of varying transmit antenna correlation on performance of space-frequency codes.

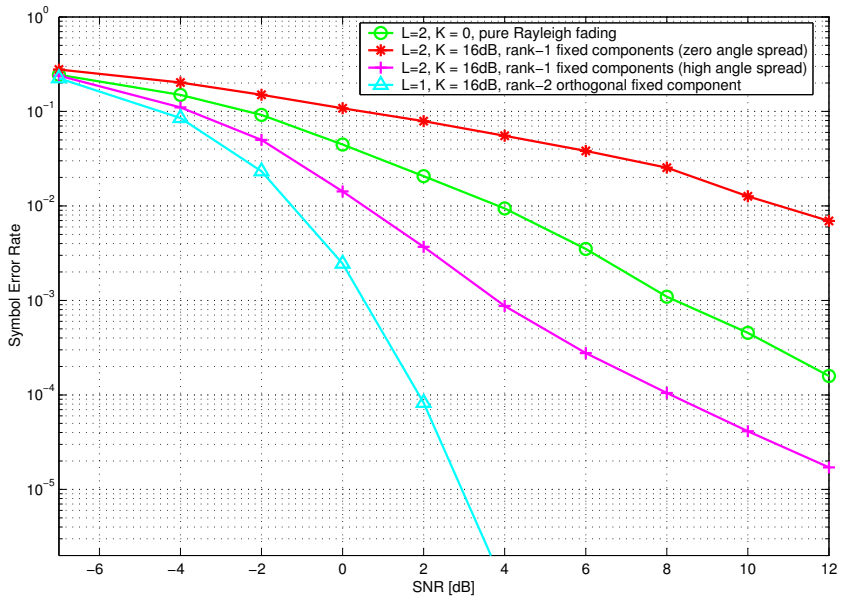


Figure 9: Impact of K-factor ($K = K_0 = K_1$) and geometry of Ricean component on performance of MIMO-OFDM spatial multiplexing.

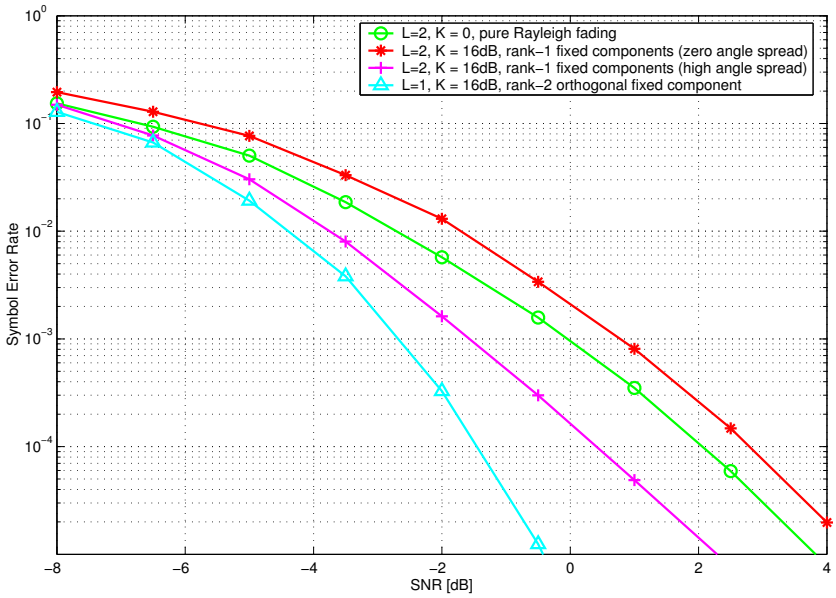


Figure 10: Impact of K-factor ($K = K_0 = K_1$) and geometry of Ricean component on performance of delay diversity combined with convolutional code.