Impact of the Propagation Environment on the Performance of Space-Frequency Coded MIMO-OFDM

Helmut Bölcskei, *Senior Member, IEEE*, Moritz Borgmann, *Student Member, IEEE*, and Arogyaswami J. Paulraj, *Fellow, IEEE*

***Abstract*—Previous work on space-frequency coded multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) has been restricted to idealistic propagation conditions. In this paper, using a broadband MIMO channel model taking into account Ricean K-factor, transmit and receive angle spread, and antenna spacing, we study the impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM. For a given space-frequency code, we quantify the achievable diversity order and coding gain as a function of the propagation parameters. We find that while the presence of spatial receive correlation affects all space-frequency codes equally, spatial fading correlation at the transmit array can result in widely varying performance losses. High-rate space-frequency codes such as spatial multiplexing are typically significantly more affected by transmit correlation than low-rate codes such as space-frequency block codes.We show that in the MIMO Ricean case the presence of frequency-selectivity typically results in improved performance compared to the frequency-flat case.**

***Index Terms*—Channel modeling, diversity, multiple-input multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), space-frequency coding, spatial multiplexing.**

I. Introduction and Outline

WO OF THE major impairments of wireless communications systems are *fading* caused by destructive addition of multipaths in the propagation medium 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. 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, *space–time coding* or equivalently the use of multiple antennas on the transmit side combined with signal processing and coding has become an active area of research [2]–[4]. Most of the previous work on space–time coding has been restricted to single-carrier systems operating over narrowband channels, where no frequency diversity is available. Broadband multiple-input multiple-output (MIMO) channels offer spatial diversity due to multiple antennas, as well as frequency diversity due to delay spread. Orthogonal frequency division multiplexing (OFDM) [5], [6] *significantly reduces receiver complexity* in wireless broadband systems. The use of MIMO technology in combination with OFDM, i.e., MIMO-OFDM [7], [8], therefore, seems to be an attractive solution for future broadband wireless systems. Space-frequency coded MIMO-OFDM [9]–[14] is a technique which basically consists of coding across transmit antennas and OFDM tones and exploits both spatial and frequency diversity without requiring channel knowledge at the transmitter.

A. Contributions

Previous work on space-frequency coding assumes idealistic propagation conditions. In this paper, using a physically motivated broadband MIMO channel model taking into account Ricean K-factor, transmit and receive angle spread, and antenna spacing, we study the impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM. Our discussion incorporates space-frequency codes [9]–[14], as well as MIMO-OFDM-based spatial multiplexing [8]. We shall use the term space-frequency coding to refer to both signaling techniques. Our detailed contributions are summarized as follows.

- We derive the *error rate performance of space-frequency coded* MIMO-OFDM as a function of Ricean K-factor, transmit and receive angle spread, antenna spacing, and power delay profile.
- For a given space-frequency code, we quantify the *achievable diversity order* and *coding gain* as a function of the propagation parameters.
- It is shown that the presence of *receive correlation affects all space-frequency codes equally*, whereas *transmit correlation results in widely varying performance losses*. In particular, we find that high-rate space-frequency codes such as spatial multiplexing are typically significantly more affected by transmit correlation than low rate codes such as space-frequency block codes.
- We show that in the *Ricean MIMO case* the *presence of frequency-selectivity typically results in improved performance* in terms of error probability with the improvement

Manuscript received April 22, 2002; revised November 30, 2002. This paper was presented in part at the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, October 2000, and in part at the European Signal Processing Conference (EUSIPCO), Toulouse, France, September 2002.

H. Bölcskei and M. Borgmann are with the Communication Technology Laboratory, ETH Zurich, CH-8092 Zürich, Switzerland (e-mail: boelcskei@nari.ee.ethz.ch).

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).

Digital Object Identifier 10.1109/JSAC.2003.809723

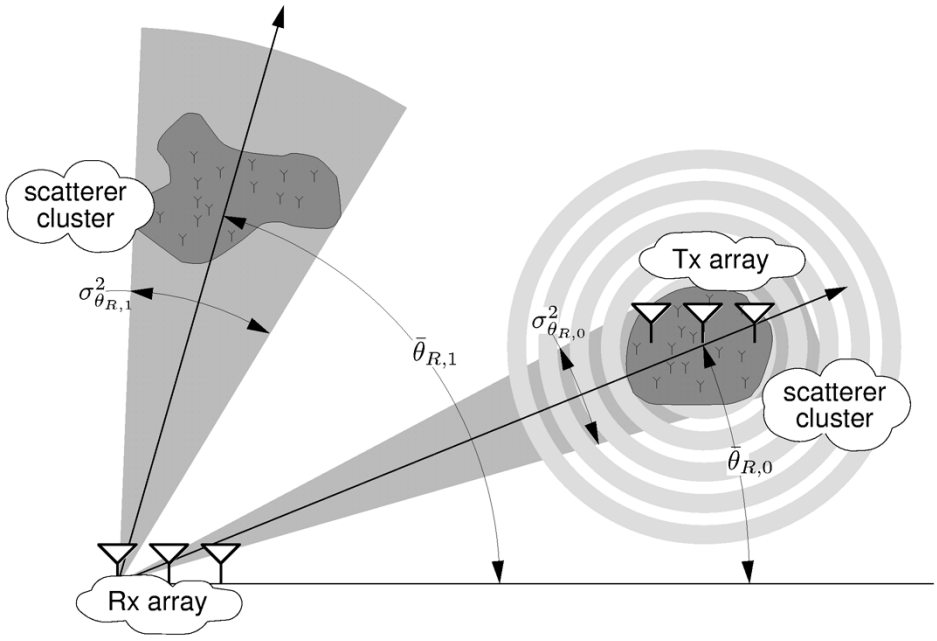

Fig. 1. Schematic representation of MIMO broadband channel composed of multiple clustered paths. For simplicity, only the relevant angles for the receive array are shown—the transmit array situation is reciprocal. The two clusters correspond to different delay taps.

being more pronounced in the case of high-rate space-frequency codes.

B. Organization of the Paper

The rest of this paper is organized as follows. In Section II, we introduce the channel model and briefly review space-frequency coded MIMO-OFDM. In Section III, we derive the error rate performance of space-frequency codes as a function of the propagation parameters. In Section IV, we find the maximum achievable diversity order and coding gain for various propagation scenarios. In Section V, we investigate the impact of the propagation parameters on the performance of space-frequency coded MIMO-OFDM. Section VI contains simulation results, and Section VII presents our conclusions.

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 $\mathbf{B}$, $\mathrm{Tr}(\mathbf{A})$ stands for the trace of $\mathbf{A}$, $r(\mathbf{A})$ is the rank of $\mathbf{A}$, $\sigma(\mathbf{A})$ denotes the eigenvalue spectrum of $\mathbf{A}$, $\mathbf{a}_i$ stands for the ith 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} stand for 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 and identically distributed (i.i.d.) $\mathcal{N}(0, \sigma^2/2)$.

II. Channel Model and Space-Frequency Coding

In this section, we shall first introduce a new model for MIMO broadband fading channels and then briefly review space-frequency coded MIMO-OFDM.

A. Channel Model

Our channel model is an extension of the model used in [8] and [9] to incorporate the Ricean case and the presence of transmit correlation.

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 given by

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

where the $M_R \times M_T$ complex-valued random matrix $\mathbf{H}_l$ represents the lth tap. One can think of each of the taps as representing 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

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

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 uncorrelated*, i.e.,

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

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).

1) 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 $\Delta_T = d_T/\lambda$ and $\Delta_R = d_R/\lambda$, 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.

2) Fading Statistics: We assume that 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 lth tap according to

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

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. For a detailed discussion on the implications of this model, the interested reader is referred to [15].

Note that the power delay profile σ_l^2 has been incorporated into the matrices $\widetilde{\mathbf{H}}_{w,l}$. From (2), we have $\mathcal{E}\{\mathbf{vec}(\widetilde{\mathbf{H}}_{w,l})(\mathbf{vec}(\widetilde{\mathbf{H}}_{w,l'}))^H\} = \mathbf{0}_{M_RM_T\times M_RM_T}$ 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. More specifically, we have, for $0 \le l \le L-1$, $\rho(s\Delta_T,\bar{\theta}_{T,l},\sigma_{\theta_{T,l}}) = \mathcal{E}\{[\widetilde{\mathbf{H}}_l]_{k,r}[\widetilde{\mathbf{H}}_l]^*_{k,r+s}\}$ $(0 \le k \le M_R-1, 0 \le r \le M_T-s-1)$ and $\rho(s\Delta_R,\bar{\theta}_{R,l},\sigma_{\theta_{R,l}}) = \mathcal{E}\{[\widetilde{\mathbf{H}}_l]_{r,k}[\widetilde{\mathbf{H}}_l]^*_{r+s,k}\}$ $(0 \le k \le M_T-1, 0 \le r \le M_R-s-1)$, respectively. The correlation matrices $\mathbf{R}_l$ and $\mathbf{S}_l$ are consequently given by

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

and

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

Let us next assume that the actual angle of departure for the lth 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 it is shown in [16] that

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

which implies that the correlation function is essentially 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, which is intuitively appealing. It must be stressed, however, that this is an approximation which becomes inaccurate when the mean angle of arrival is close to zero 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

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

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

3) Ricean Component: The Ricean component of the lth tap is modeled as

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

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 ith component of $\bar{\mathbf{H}}_l$ and $\beta_{l,i}$ is the corresponding complex path amplitude. We can, furthermore, associate a Ricean K-factor with each of the taps by defining

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

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

4) Comments on the Channel Model: For the sake of simplicity of presentation, we assumed that different scatterer clusters can be resolved in time. In practice, this is not necessarily the case. We emphasize, however, that allowing different scatterer clusters to have the same delay does not yield significant new insights into the impact of the propagation conditions on space-frequency code performance.

B. Space-Frequency Coded MIMO-OFDM

In the following, N denotes the number of subcarriers or tones in the OFDM system. Throughout the paper, we assume that the length of the cyclic prefix (CP) is greater than or equal to L. Organizing the transmitted data symbols into 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 ith antenna on the kth tone, it can be shown that the reconstructed data vector for the kth tone is given by

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

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

In a space-frequency coded OFDM system, one data burst consists of N vectors of size $M_T \times 1$ or equivalently one spatial OFDM symbol. Throughout the paper, we assume that the channel is constant over at least one OFDM symbol, the transmitter has no channel knowledge, and the receiver knows the channel perfectly. The maximum-likelihood (ML) decoder computes the vector sequence $\hat{\mathbf{c}}_k$ $(k = 0,1,\ldots,N-1)$ according to

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

where $\mathbf{C} = [\mathbf{c}_0\, \mathbf{c}_1\, \ldots\, \mathbf{c}_{N-1}]$ and $\hat{\mathbf{C}} = [\hat{\mathbf{c}}_0\, \hat{\mathbf{c}}_1\, \ldots\, \hat{\mathbf{c}}_{N-1}]$. For a more detailed treatment of MIMO-OFDM and space-frequency coding, the interested reader is referred to [7]–[10].

III. Error Rate Performance

In this section, we derive the pairwise error probability (PEP) for space-frequency codes taking into account the channel model introduced in Section II-A.

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 probability that the receiver decides erroneously in favor of the signal $\mathbf{E}$ assuming that $\mathbf{C}$ was transmitted is given by [17, p. 258], [3]

$$P(\mathbf{C} \to \mathbf{E}|\mathbf{H}) = Q\left(\sqrt{\frac{E_s}{2\sigma_n^2} d^2\left(\mathbf{C}, \mathbf{E}|\,\mathbf{H}\right)}\right)$$

with $d^2(\mathbf{C}, \mathbf{E}|\mathbf{H}) = \sum_{k=0}^{N-1} \|\mathbf{H}(e^{j(2\pi/N)k})(\mathbf{c}_k - \mathbf{e}_k)\|^2$. Defining $\mathbf{y}_k = \mathbf{H}(e^{j(2\pi/N)k})(\mathbf{c}_k - \mathbf{e}_k)$ for $k = 0, 1, \ldots, N-1$ and $\mathbf{Y} = [\mathbf{y}_0^T\ \mathbf{y}_1^T\ \ldots\ \mathbf{y}_{N-1}^T]^T$, we obtain $d^2(\mathbf{C}, \mathbf{E}|\mathbf{H}) = \|\mathbf{Y}\|^2$. Using the Chernoff bound $Q(x) \le e^{-x^2/2}$, we therefore get

$$P(\mathbf{C} \to \mathbf{E}|\mathbf{H}) \le e^{-\frac{E_s}{4\sigma_n^2}\|\mathbf{Y}\|^2}. \tag{8}$$

Since the $\mathbf{H}_l$ were assumed to be complex Gaussian it follows from (1) that the $\mathbf{H}(e^{j(2\pi/N)k})$ $(k = 0, 1, \ldots, N-1)$ are complex Gaussian as well, and hence, the $M_R N$-dimensional vector $\mathbf{Y}$ is complex Gaussian. The average of (8) over the random channel follows from the characteristic function of $\mathbf{Y}$, which is fully specified by the mean vector $\bar{\mathbf{Y}} = \mathcal{E}\{\mathbf{Y}\} = [\bar{\mathbf{y}}_0^T\ \bar{\mathbf{y}}_1^T\ \ldots\ \bar{\mathbf{y}}_{N-1}^T]^T$ with $\bar{\mathbf{y}}_k = \bar{\mathbf{H}}(e^{j(2\pi/N)k})(\mathbf{c}_k - \mathbf{e}_k)$ and $\bar{\mathbf{H}}(e^{j(2\pi/N)k}) = \sum_{l=0}^{L-1} \bar{\mathbf{H}}_l e^{-j(2\pi/N)kl}$, and the covariance matrix of $\mathbf{Y}$ defined as $\mathbf{C}_Y = \mathcal{E}\{(\mathbf{Y} - \bar{\mathbf{Y}})(\mathbf{Y} - \bar{\mathbf{Y}})^H\}$. In the following, in order to keep the presentation simple, we assume that $N > M_T L$. Averaging (8) over the random channel and denoting $P(\mathbf{C} \to \mathbf{E}) = \mathcal{E}\{P(\mathbf{C} \to \mathbf{E}|\mathbf{H})\}$, we obtain

$$\begin{aligned} P(\mathbf{C} \to \mathbf{E}) \le\ & \exp\left(-\frac{E_s}{4\sigma_n^2}\|\bar{\mathbf{Y}}\|^2\right) \\ & \cdot \prod_{i=0}^{r(\mathbf{C}_Y)-1} \exp\left(\left(\frac{E_s}{4\sigma_n^2}\right)^2 \frac{|b_i|^2 \lambda_i^2(\mathbf{C}_Y)}{1 + \frac{E_s}{4\sigma_n^2}\lambda_i(\mathbf{C}_Y)}\right) \\ & \cdot \frac{1}{1 + \frac{E_s}{4\sigma_n^2}\lambda_i(\mathbf{C}_Y)} \end{aligned} \tag{9}$$

where

$$\mathbf{C}_Y = \sum_{l=0}^{L-1} \sigma_l^2 \left[\mathbf{D}^l (\mathbf{C} - \mathbf{E})^T \mathbf{S}_l (\mathbf{C} - \mathbf{E})^* \mathbf{D}^{-l}\right] \otimes \mathbf{R}_l \tag{10}$$

with $\mathbf{D} = \mathrm{diag}\{e^{-j(2\pi/N)k}\}_{k=0}^{N-1}$, $\mathbf{C}_Y = \mathbf{U}\mathbf{\Phi}\mathbf{U}^H$ with $\mathbf{\Phi} = \mathrm{diag}\{\lambda_0(\mathbf{C}_Y), \ldots, \lambda_{r(\mathbf{C}_Y)-1}(\mathbf{C}_Y), 0, \ldots, 0\}$, and $\mathbf{b} = [b_0\ b_1\ \ldots\ b_{M_R N-1}]^T = \mathrm{diag}\{1/\sqrt{\lambda_0(\mathbf{C}_Y)}, \ldots, 1/\sqrt{\lambda_{r(\mathbf{C}_Y)-1}(\mathbf{C}_Y)}, 0, \ldots, 0\}\mathbf{U}^H\bar{\mathbf{Y}}$. We note that the assumption $N > M_T L$ made above implies that $\mathbf{C}_Y$ is always rank-deficient. For full-rank $\mathbf{C}_Y$, we get a slightly different expression for the PEP upper bound. Due to space limitations this case will, however, not be discussed here. The derivation of (9) is rather lengthy but standard and follows from an extension of results presented in [18] to the complex-valued case, a detailed account of which can be found in [19, Appendix].

IV. Maximum Diversity Order and Coding Gain

Based on (9), we are now ready to establish the maximum achievable diversity order and coding gain for various propagation scenarios. Throughout this section, we focus on the high signal-to-noise ratio (SNR) regime, i.e., $(E_s/4\sigma_n^2) \gg 1$.

A. 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, (9) simplifies to

$$\mathbf{P}(\mathbf{C} \to \mathbf{E}) \le \left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_Y)} \prod_{i=0}^{r(\mathbf{C}_Y)-1} \lambda_i^{-1}(\mathbf{C}_Y). \tag{11}$$

We recall the following definitions [3]. 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 $(\prod_{i=0}^{d-1} \lambda_i(\mathbf{C}_Y))^{1/d}$ 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 $\mathcal{L} = \{i : \sigma_i^2 > 0\}$ 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 (10) specializes to

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

Let us next set $(\mathbf{C} - \mathbf{E})^T (\mathbf{C} - \mathbf{E})^* = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$ and $\mathbf{R}_l = \mathbf{U}_l \mathbf{\Sigma}_l \mathbf{U}_l^H (l = 0, 1, \ldots, L-1)$, where $\mathbf{V}$ is an $N \times M_T$ matrix with orthonormal columns, i.e., $\mathbf{V}^H\mathbf{V} = \mathbf{I}_{M_T}$, and the $\mathbf{U}_l$ are unitary $M_R \times M_R$ matrices. Furthermore, $\mathbf{\Lambda}$ is an $M_T \times M_T$ diagonal matrix containing the eigenvalues of $(\mathbf{C} - \mathbf{E})^*(\mathbf{C} - \mathbf{E})^T$, and the $\mathbf{\Sigma}_l$ are $M_R \times M_R$ diagonal matrices containing the eigenvalues of $\mathbf{R}_l$. With these definitions, we can rewrite $\mathbf{C}_Y$ as

$$\mathbf{C}_Y = \sum_{l=0}^{L-1} \sigma_l^2 \left[(\mathbf{D}^l \mathbf{V}) \mathbf{\Lambda} (\mathbf{V}^H \mathbf{D}^{-l})\right] \otimes \left[\mathbf{U}_l \mathbf{\Sigma}_l \mathbf{U}_l^H\right]$$

which, with the following property of Kronecker products [20]

$$\begin{aligned} & (\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) \ldots (\mathbf{A}_k \otimes \mathbf{B}_k) \\ & \qquad = (\mathbf{A}_1 \mathbf{A}_2 \ldots \mathbf{A}_k) \otimes (\mathbf{B}_1 \mathbf{B}_2 \ldots \mathbf{B}_k) \end{aligned} \tag{12}$$

yields

$$\mathbf{C}_Y = \sum_{l=0}^{L-1} \sigma_l^2 \left((\mathbf{D}^l\mathbf{V}) \otimes \mathbf{U}_l\right) (\mathbf{\Lambda} \otimes \mathbf{\Sigma}_l) \left((\mathbf{V}^H\mathbf{D}^{-l}) \otimes \mathbf{U}_l^H\right).$$

Hence, we can write

$$\mathbf{C}_Y = \mathbf{M}\mathbf{L}^{1/2}\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2}\mathbf{L}^{1/2}\mathbf{M}^H \tag{13}$$

where

$$\mathbf{M} = \left[\sqrt{\sigma_0^2}(\mathbf{V} \otimes \mathbf{U}_0) \quad \sqrt{\sigma_1^2}(\mathbf{D}\mathbf{V}) \otimes \mathbf{U}_1 \quad \cdots \sqrt{\sigma_{L-1}^2}(\mathbf{D}^{L-1}\mathbf{V}) \otimes \mathbf{U}_{L-1}\right]$$

$$\mathbf{\Sigma}^{1/2} = \operatorname{diag}\left\{\mathbf{I}_{M_T} \otimes \mathbf{\Sigma}_l^{1/2}\right\}_{l=0}^{L-1}$$

and

$$\mathbf{L}^{1/2} = \operatorname{diag}\{\mathbf{\Lambda}^{1/2} \otimes \mathbf{I}_{M_R}\}_{l=0}^{L-1}.$$

Here, we used the decomposition $\mathbf{\Lambda} \otimes \mathbf{\Sigma}_l = (\mathbf{\Lambda}^{1/2} \otimes \mathbf{I}_{M_R})(\mathbf{I}_{M_T} \otimes \mathbf{\Sigma}_l^{1/2})(\mathbf{I}_{M_T} \otimes \mathbf{\Sigma}_l^{1/2})(\mathbf{\Lambda}^{1/2} \otimes \mathbf{I}_{M_R})$. We can draw a number of important conclusions from (13). Let us organize the corresponding discussion into two different cases.

Case I: Assume that all the correlation matrices $\mathbf{R}_l$ with $l \in \mathcal{L}$ have full rank. Exploiting the fact that $\sigma(\mathbf{A}^H\mathbf{A}) = \{\sigma(\mathbf{A}\mathbf{A}^H), 0, \ldots, 0\}$ with $\mathbf{A}$ a $P \times Q$ matrix, where $Q > P$ and the number of additional eigenvalues equal to zero is $Q - P$, it follows from (13) that $\sigma(\mathbf{C}_Y) = \{\sigma(\widetilde{\mathbf{C}}_Y), 0, \ldots, 0\}$, where

$$\widetilde{\mathbf{C}}_Y = \mathbf{\Sigma}^{1/2}\mathbf{L}^{1/2}\mathbf{M}^H\mathbf{M}\mathbf{L}^{1/2}\mathbf{\Sigma}^{1/2}. \tag{14}$$

Since $r(\widetilde{\mathbf{C}}_Y) = r(\mathbf{C}_Y)$ and the nonzero eigenvalues of $\mathbf{C}_Y$ are equal to the nonzero eigenvalues of $\widetilde{\mathbf{C}}_Y$, it follows from (11) that the PEP can be inferred from $r(\widetilde{\mathbf{C}}_Y)$ and the nonzero eigenvalues of $\widetilde{\mathbf{C}}_Y$. In the following, we shall need Ostrowski's Theorem, which we state for completeness.

Theorem 1 (Ostrowski): Let $\mathbf{A}$ and $\mathbf{S}$ be $n \times n$ matrices with $\mathbf{A}$ Hermitian and $\mathbf{S}$ nonsingular. Let the eigenvalues of $\mathbf{A}$ and $\mathbf{S}\mathbf{S}^H$ be arranged in increasing order $\lambda_{\min} = \lambda_1 \le \lambda_2 \le \ldots \le \lambda_{n-1} \le \lambda_n = \lambda_{\max}$. For each $k = 1, 2, \ldots, n$ there exists a nonnegative real number θ_k such that $0 < \lambda_1(\mathbf{S}\mathbf{S}^H) \le \theta_k \le \lambda_n(\mathbf{S}\mathbf{S}^H)$ and $\lambda_k(\mathbf{S}\mathbf{A}\mathbf{S}^H) = \theta_k\lambda_k(\mathbf{A})$.

Applying Ostrowski's Theorem to (14) yields

$$\lambda_i(\widetilde{\mathbf{C}}_Y) = \theta_i\lambda_i(\mathbf{L}^{1/2}\mathbf{M}^H\mathbf{M}\mathbf{L}^{1/2}) \tag{15}$$

where $0 < \lambda_{\min}(\mathbf{R}) \le \theta_i \le \lambda_{\max}(\mathbf{R})$ with $\mathbf{R} = \operatorname{diag}\{\mathbf{R}_l\}_{l\in\mathcal{L}}$. Next, we note that $\sigma(\mathbf{M}\mathbf{L}\mathbf{M}^H) = \{\sigma(\mathbf{L}^{1/2}\mathbf{M}^H\mathbf{M}\mathbf{L}^{1/2}), 0, \ldots, 0\}$ and $\mathbf{M}\mathbf{L}\mathbf{M}^H = \sum_{l=0}^{L-1} \sigma_l^2 ((\mathbf{D}^l\mathbf{V}) \otimes \mathbf{U}_l)(\mathbf{\Lambda} \otimes \mathbf{I}_{M_R})((\mathbf{V}^H\mathbf{D}^{-l}) \otimes \mathbf{U}_l^H) = \mathbf{C}_u \otimes \mathbf{I}_{M_R}$, where

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

Therefore, we can conclude that the eigenvalues of $\mathbf{C}_Y$ can be expressed as

$$\lambda_i(\mathbf{C}_Y) = \theta_i\lambda_i(\mathbf{C}_u \otimes \mathbf{I}_{M_R}). \tag{17}$$

Finally, we remark that in the i.i.d. case $\mathbf{C}_Y = \mathbf{C}_u \otimes \mathbf{I}_{M_R}$, and each eigenvalue of $\mathbf{C}_u$ is an eigenvalue of $\mathbf{C}_u \otimes \mathbf{I}_{M_R}$ with multiplicity M_R.

Inserting (17) into (11) and using $r(\mathbf{C}_u \otimes \mathbf{I}_{M_R}) = M_R r(\mathbf{C}_u)$ and $\theta_i > 0$ (follows from the assumption that all the $\mathbf{R}_l$ are full rank), we obtain

$$P(\mathbf{C} \to \mathbf{E}) \le \left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_u)M_R} r(\mathbf{C}_u)M_R - 1 \prod_{i=0}^{r(\mathbf{C}_u)M_R-1} \theta_i^{-1} \prod_{i=0}^{r(\mathbf{C}_u)-1} \lambda_i^{-M_R}(\mathbf{C}_u).$$

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 $\mathbf{R}_l$ 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 $(\prod_{i=0}^{r(\mathbf{C}_u)-1} \lambda_i(\mathbf{C}_u))^{1/r(\mathbf{C}_u)}$ multiplied by $(\prod_{i=0}^{r(\mathbf{C}_u)M_R-1} \theta_i)^{1/(r(\mathbf{C}_u)M_R)}$.

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 $P(\mathbf{C} \to \mathbf{E}) \le (E_s/4\sigma_n^2)^{-M_T M_R|\mathcal{L}|} 1/\det(\widetilde{\mathbf{C}}_Y)$. Using $\det(\widetilde{\mathbf{C}}_Y) = \det(\mathbf{\Sigma}) \det(\mathbf{L}^{1/2}\mathbf{M}^H\mathbf{M}\mathbf{L}^{1/2})$, and exploiting the fact that $\det(\mathbf{L}^{1/2}\mathbf{M}^H\mathbf{M}\mathbf{L}^{1/2}) = \prod_i \lambda_i(\mathbf{M}\mathbf{L}\mathbf{M}^H)$ with $\lambda_i(\mathbf{M}\mathbf{L}\mathbf{M}^H)$ denoting the nonzero eigenvalues of $\mathbf{M}\mathbf{L}\mathbf{M}^H = \mathbf{C}_u \otimes \mathbf{I}_{M_R}$, we get

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

We can, therefore, conclude 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 [11]. The coding gain is given by the coding gain achieved in the i.i.d. case multiplied by $(\prod_{l\in\mathcal{L}} \det(\mathbf{R}_l))^{1/(M_R|\mathcal{L}|)}$. Using the normalization $\operatorname{Tr}(\mathbf{R}) = M_R|\mathcal{L}|$, it is easily seen that $\prod_{l\in\mathcal{L}} \det(\mathbf{R}_l) \le 1$, 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 $\mathbf{R}$.

Case II: If $\mathbf{R}$ 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 $d \le M_T \sum_{l\in\mathcal{L}} r(\mathbf{R}_l)$. This can be seen by noting that $r(\mathbf{C}_Y) = r(\mathbf{L}^{1/2}\mathbf{\Sigma}^{1/2}\mathbf{M}^H\mathbf{M}\mathbf{\Sigma}^{1/2}\mathbf{L}^{1/2}) = r(\mathbf{M}\mathbf{\Sigma}\mathbf{M}^H) = r(\sum_{l=0}^{L-1} \sigma_l^2(\mathbf{D}^l\mathbf{V}\mathbf{V}^H\mathbf{D}^{-l}) \otimes \mathbf{R}_l) \le M_T \sum_{l\in\mathcal{L}} r(\mathbf{R}_l)$, where we exploited the diagonality of $\mathbf{L}$ and $\mathbf{\Sigma}$ and the fact that $\mathbf{L}$ has full rank for a space-frequency code achieving full diversity gain in the i.i.d. case.

Retaining the assumption of $\mathbf{R}$ being rank-deficient and assuming that the space-frequency code does not achieve

full space-frequency diversity in the i.i.d. case, focusing on a minimum rank error event $\mathbf{C}-\mathbf{E}$, it follows that:

$$\begin{aligned} r(\mathbf{\Lambda})M_R|\mathcal{L}| + r\left(\sum_{l=0}^{L-1}\sigma_l^2(\mathbf{D}^l\mathbf{V}\mathbf{V}^H\mathbf{D}^{-l})\otimes\mathbf{R}_l\right) \\ - M_TM_R|\mathcal{L}| \le d \\ \le \min\Bigg(r(\mathbf{\Lambda})M_R|\mathcal{L}|, \\ r\left(\sum_{l=0}^{L-1}\sigma_l^2(\mathbf{D}^l\mathbf{V}\mathbf{V}^H\mathbf{D}^{-l})\otimes\mathbf{R}_l\right)\Bigg). \end{aligned}$$

The impact of receive correlation on the coding gain is in general difficult to quantify in Case II.

Transmit Correlation Only: In the case of transmit correlation only $\mathbf{R}_l = \mathbf{I}_{M_R}$ for $l = 0, 1, \ldots, L-1$ and

$$\mathbf{C}_Y = \underbrace{\sum_{l=0}^{L-1}\sigma_l^2\left[\mathbf{D}^l(\mathbf{C}-\mathbf{E})^T\mathbf{S}_l(\mathbf{C}-\mathbf{E})^*\mathbf{D}^{-l}\right]}_{\mathbf{C}_s}\otimes\mathbf{I}_{M_R}. \quad (19)$$

Focusing on a minimum-rank codeword pair $\{\mathbf{C},\mathbf{E}\}$, the average PEP can be upper-bounded by

$$P(\mathbf{C}\to\mathbf{E}) \le \left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_s)M_R}\prod_{i=0}^{r(\mathbf{C}_s)-1}\lambda_i^{-M_R}(\mathbf{C}_s). \quad (20)$$

Assuming that $\mathbf{S}_l$ has full rank for $l\in\mathcal{L}$, we shall next make the dependence of the average PEP on the eigenvalues of the $\mathbf{S}_l$ $(l = 0, 1, \ldots, L-1)$ explicit. For this define

$$\mathbf{F}(\mathbf{C},\mathbf{E}) = \left[\sqrt{\sigma_0^2}(\mathbf{C}-\mathbf{E})^T \ \cdots \ \sqrt{\sigma_{L-1}^2}\mathbf{D}^{L-1}(\mathbf{C}-\mathbf{E})^T\right]$$

and let the eigenvalues of $\mathbf{F}^H(\mathbf{C},\mathbf{E})\mathbf{F}(\mathbf{C},\mathbf{E})$ and $\mathbf{S} = \mathbf{S}^{1/2}\mathbf{S}^{1/2} = \mathrm{diag}\{\mathbf{S}_l\}_{l\in\mathcal{L}}$ be arranged in increasing order. With these definitions we have $\mathbf{C}_s = \mathbf{F}(\mathbf{C},\mathbf{E})\mathbf{S}\mathbf{F}^H(\mathbf{C},\mathbf{E})$. Next, using Ostrowski's Theorem, we get

$$\begin{aligned}\lambda_i(\mathbf{F}(\mathbf{C},\mathbf{E})\mathbf{S}\mathbf{F}^H(\mathbf{C},\mathbf{E})) &= \theta_i\lambda_i(\mathbf{F}(\mathbf{C},\mathbf{E})\mathbf{F}^H(\mathbf{C},\mathbf{E})) \\ &= \theta_i\lambda_i(\mathbf{C}_u)\end{aligned}$$

where $0 < \lambda_{\min}(\mathbf{S}) \le \theta_i \le \lambda_{\max}(\mathbf{S})$ and $\mathbf{F}(\mathbf{C},\mathbf{E})\mathbf{F}^H(\mathbf{C},\mathbf{E}) = \mathbf{C}_u$ was used. Consequently, we obtain the following bound on PEP for nonsingular $\mathbf{S}$

$$\begin{aligned}&P(\mathbf{C}\to\mathbf{E}) \\ &\le \left(\frac{E_s}{4\sigma_n^2}\right)^{-M_Rr(\mathbf{C}_u)}\prod_{i=0}^{r(\mathbf{C}_u)-1}\theta_i^{-M_R}\prod_{i=0}^{r(\mathbf{C}_u)-1}\lambda_i^{-M_R}(\mathbf{C}_u) \quad (21)\end{aligned}$$

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 correlated case is equal to the diversity order achieved by the space-frequency code in the i.i.d. case. For general $\mathbf{S}$, focusing on a minimum-rank error event $\mathbf{C}-\mathbf{E}$, we get

$$M_R\left(r(\mathbf{C}_u)+r(\mathbf{S})-M_T|\mathcal{L}|\right) \le d \le M_R\min\left\{r(\mathbf{C}_u), r(\mathbf{S})\right\}. \quad (22)$$

We note that (22) shows that if $\mathbf{S}$ 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 $M_R(r(\mathbf{C}_u)+r(\mathbf{S})-M_T|\mathcal{L}|)$. 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 (19) 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. Simulation results in Section VI will offer more and in particular quantitative insights along these lines. In the presence of receive correlation only, the situation is fundamentally different as the performance degradation depends solely on the amount of receive correlation. This can be seen by investigating (18), for example, where it is obvious that the loss due to receive correlation is quantified solely by the term $\prod_{l\in\mathcal{L}}\det^{-M_T}(\mathbf{R}_l)$.

Joint Transmit-Receive Correlation: In the case of joint transmit-receive correlation, it follows from (10) 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 lth 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 offered by the channel.

B. Ricean Fading

So far, we have restricted our attention to Rayleigh fading. Let us next consider the Ricean case. Again assuming high SNR $((E_s/4\sigma_n^2)\gg 1)$, the PEP upper bound in (9) simplifies to

$$\begin{aligned}&\mathbf{P}(\mathbf{C}\to\mathbf{E}) \\ &\le \exp\left\{\frac{E_s}{4\sigma_n^2}\left(\sum_{i=0}^{r(\mathbf{C}_Y)-1}|b_i|^2\lambda_i(\mathbf{C}_Y) - \|\bar{\mathbf{Y}}\|^2\right)\right\} \\ &\quad\cdot\left(\frac{E_s}{4\sigma_n^2}\right)^{-r(\mathbf{C}_Y)}\prod_{i=0}^{r(\mathbf{C}_Y)-1}\lambda_i^{-1}(\mathbf{C}_Y). \quad (23)\end{aligned}$$

We can see that the PEP upper bound consists of an exponential term multiplied by the right-hand side (RHS) of (11), which is just the PEP upper bound in the Rayleigh-fading case. The behavior of the second and third terms of the RHS in (23) was studied in Section IV-A. We shall, therefore, focus on the first term of the RHS of (23), which represents the contribution due to the Ricean component, and start by noting that

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

which after application of the Rayleigh–Ritz Theorem yields

$$e^{-\frac{E_s}{4\sigma_n^2}\|\bar{\mathbf{Y}}\|^2} \leq \exp\left\{\frac{E_s}{4\sigma_n^2}\left(\sum_{i=0}^{r(\mathbf{C}_Y)-1}|b_i|^2\lambda_i(\mathbf{C}_Y)-\|\bar{\mathbf{Y}}\|^2\right)\right\} \leq 1. \quad (24)$$

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 (24) is achieved, we get

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

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 $\bar{\mathbf{H}}(e^{j(2\pi/N)k})\,(k=0,1\ldots,N-1)$ and, hence, the matrices $\bar{\mathbf{H}}_l$. In particular, for high-rate codes such as spatial multiplexing the matrices $\bar{\mathbf{H}}(e^{j(2\pi/N)k})$ are required to have high rank in order to ensure good performance in terms of error probability. Let us first investigate the flat-fading case, where $\bar{\mathbf{H}}(e^{j(2\pi/N)k}) = \bar{\mathbf{H}}_0$ 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 spread is sufficiently large. We can, therefore, expect that large angle spread will lead to small error probability and vice versa. In the frequency-selective case, the matrices $\bar{\mathbf{H}}(e^{j(2\pi/N)k})$ can have full rank even if the individual matrices $\bar{\mathbf{H}}_l$ are rank-deficient, and the performance will depend on the angle spread measured over all the paths (i.e. the total angle spread). This observation will be confirmed in the simulation example in Section VI-D. For high K-factor, we note that in the MIMO Ricean case the situation is fundamentally different from the single-input single-output case, where frequency-selectivity causes dips in the frequency response, and the performance degrades compared to the flat-fading case.

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

V. Impact of Propagation Parameters

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.

A. 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 a Töplitz matrix, we can invoke a result from [21] to obtain the limiting $(M_T\to\infty)$ distribution of the eigenvalues of $\mathbf{S}_l$ as

$$\widehat{\lambda}_l(\nu) = \sum_{s=-\infty}^{\infty}\rho(s\Delta_T,\bar{\theta}_{T,l},\sigma_{\theta_{T,l}})e^{-j2\pi s\nu}, \qquad 0\leq\nu<1$$

which upon using (4) yields $\widehat{\lambda}_l(\nu)=\vartheta_3(\pi(\nu-\Delta_T\cos(\bar{\theta}_{T,l})), e^{-(1/2)(2\pi\Delta_T\sin(\bar{\theta}_{T,l})\sigma_{\theta_{T,l}})^2})$ with the third-order theta function given by $\vartheta_3(\nu,q)=\sum_{n=-\infty}^{\infty}q^{n^2}\,e^{2jn\nu}$ [22]. Although this expression yields the exact eigenvalue distribution only in the limiting case $M_T\to\infty$, in the finite M_T case good approximations of the eigenvalues of $\mathbf{S}_l$ can be obtained by sampling $\widehat{\lambda}_l(\nu)$ uniformly [21], which allows us to assume that the eigenvalue distribution in the finite case follows $\widehat{\lambda}_l(\nu)$. Noting that the correlation function $\rho(s\Delta_T,\bar{\theta}_{T,l},\sigma_{\theta_{T,l}})$ as a function of s is essentially a modulated Gaussian function with its spread inversely proportional to the product of antenna spacing and cluster angle spread, it follows that $\widehat{\lambda}_l(\nu)$ 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 angle spread (i.e., high spatial fading correlation) $\widehat{\lambda}_l(\nu)$ will be peaky. Figs. 2(a) and (b) show the limiting eigenvalue distribution of $\mathbf{S}_l$ for high and low spatial fading correlation, respectively.

Finally, 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 $\widehat{\lambda}_l(\nu)$. In particular, defining

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

we have that [23]

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

This result allows to provide an upper bound on the PEP in terms of $\widehat{\lambda}_l(\nu)$ and, hence, makes the dependence of the PEP on 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 (21) that

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

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. Similar conclusions can be drawn for most of the other cases discussed in the previous section.

B. 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 $\mathbf{R}_l=\mathbf{a}(\bar{\theta}_{R,l})\mathbf{a}^H(\bar{\theta}_{R,l})$ $(l=$

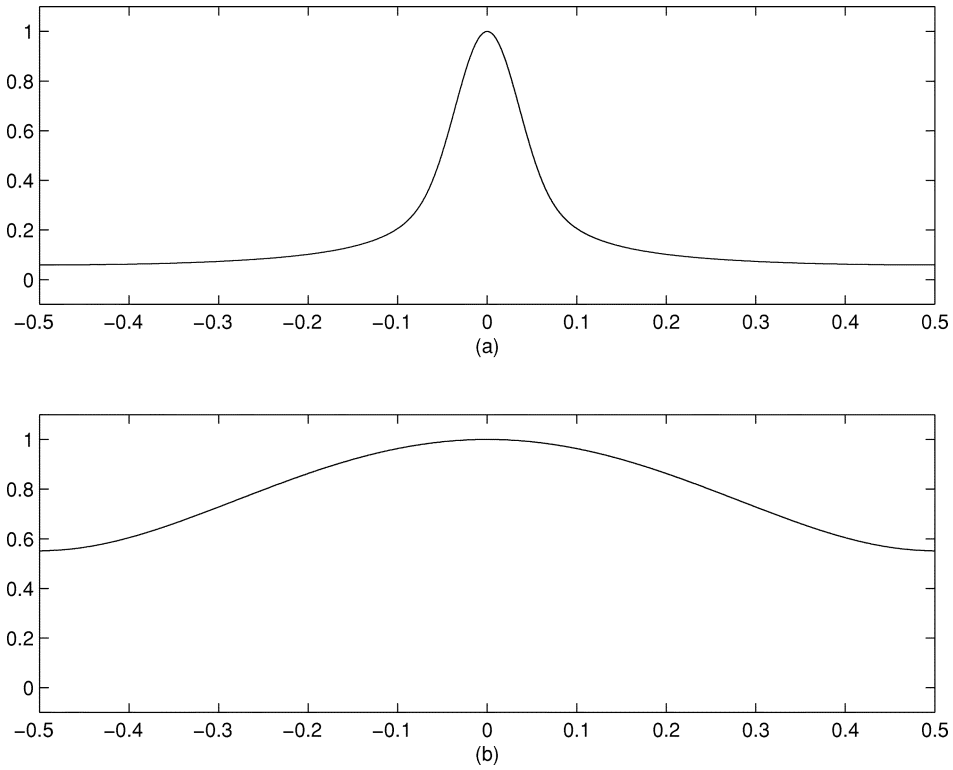

Fig. 2. Limiting eigenvalue distribution of the correlation matrix $\mathbf{S}_l$ for the cases of (a) high-spatial fading correlation and (b) low-spatial fading correlation.

$0, 1, \ldots, L-1$). 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, we get

$$\widetilde{\mathbf{C}}_Y = \operatorname{diag}\left\{\sigma_l^2[(\mathbf{C}-\mathbf{E})^*(\mathbf{C}-\mathbf{E})^T] \otimes \mathbf{R}_l\right\}_{l\in\mathcal{L}}. \quad (25)$$

The block diagonality of $\widetilde{\mathbf{C}}_Y$ implies several interesting properties. First, we note that (25) 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 (25) that the space-frequency code design criteria reduce to the well-known space–time code design criteria first reported in [3]. 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 applied to the stacked matrix $[\sqrt{\sigma_0^2}(\mathbf{C}-\mathbf{E})^T \ \cdots \ \sqrt{\sigma_{L-1}^2}\mathbf{D}^{L-1}(\mathbf{C}-\mathbf{E})^T]$. 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 $\mathbf{F}$ 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 columns of $(\mathbf{C}_t - \mathbf{E}_t)$ by l positions downwards. 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 $(\mathbf{C}_t - \mathbf{E}_t)$ 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 such as for example in [11]. Now, in the case of receive correlation only where the correlation matrices 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 channel orthogonalizes the delayed versions of the transmitted signal, and space-frequency code design for the broadband case reduces to classical space–time code design for the narrowband case.

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 ($L = 1$), 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

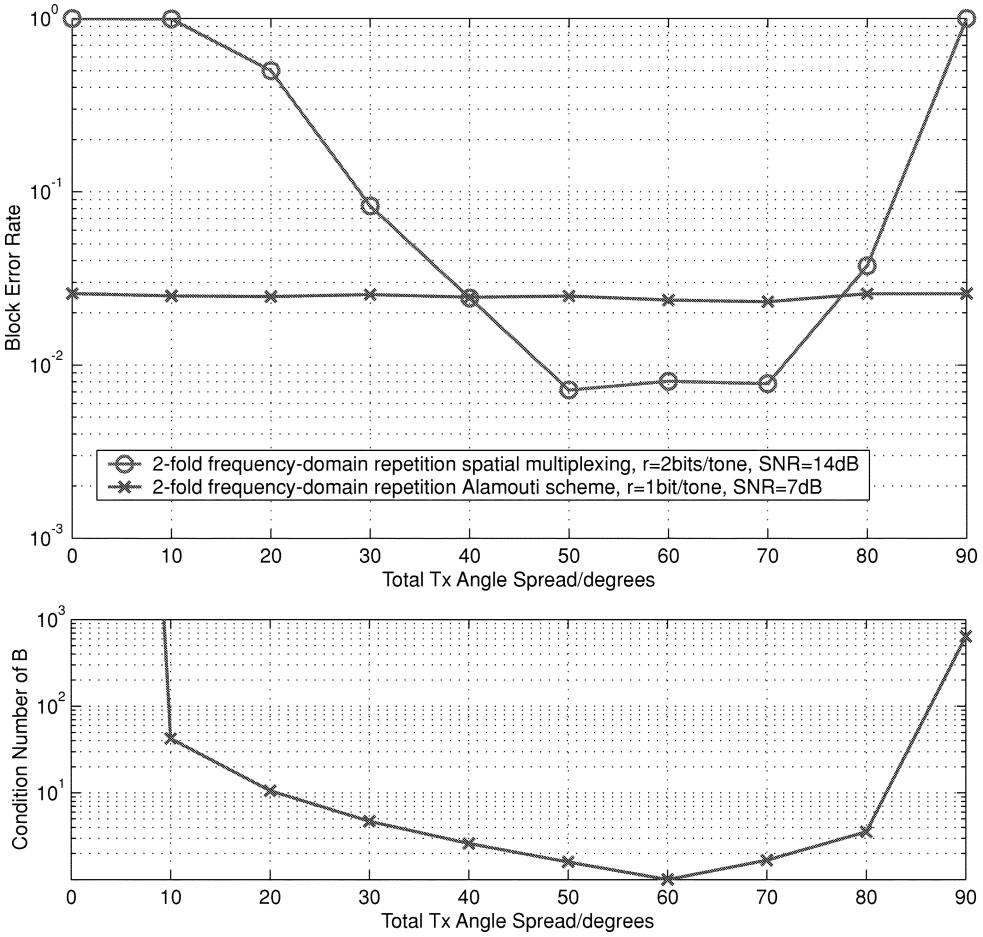

Fig. 3. Impact of total transmit angle spread on performance of space-frequency codes.

large PEP. Next, consider the case of multipath propagation, where $L > 1$, and retain the assumption of small cluster angle spreads and/or small transmit antenna spacing. If the total angle spread is large, it follows that even though the individual correlation matrices $\mathbf{S}_l$ are rank-deficient, the $\mathbf{S}_l$ will span different subspaces. Consequently, if $(\mathbf{C}-\mathbf{E})^*$ lies in the null space of one of the $\mathbf{S}_l$, it can still lie in the range space of one of the other $\mathbf{S}_l$. Using (19) it hence follows that the PEP performance may still be good. This effect will be studied further and quantified in the simulation example in Section VI-A.

Summarizing, we can conclude that increased total angle spread not only leads to an increased ergodic capacity as observed in [8], but generally also to improved performance from an error probability point of view.

VI. Simulation Results

In this section, we provide simulation results that corroborate and quantify some of the analytical results derived in the paper. Unless specified otherwise, we simulated a space-frequency coded MIMO-OFDM system with $M_T = M_R = 2$, $L = 2, \Delta_T = \Delta_R = 1, \sigma_0^2 = \sigma_1^2 = 1/2, N = 32$ tones, and ML decoding. The signal-to-noise ratio is defined as $\mathrm{SNR} = 10\log_{10}(M_T E_s/\sigma_n^2)$. 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 the space-frequency coding scheme proposed in [11], where an arbitrary (inner) space–time code is used to transmit data in the first $N/2$ tones followed by a repetition of this block. This construction ensures that the additionally available frequency diversity is fully exploited.

A. 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 quaternary phase-shift keying (QPSK) based spatial multiplexing [24], [25], which amounts to transmitting independent data symbols on each tone and each antenna, and the Alamouti scheme [4] 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 top diagram in Fig. 3 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. 3. The lower rate Alamouti scheme is virtually unaffected by the total transmit angle spread, which is due to the

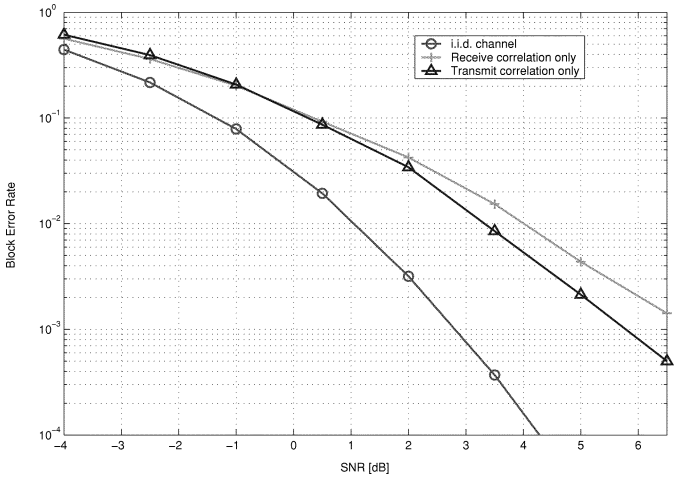

Fig. 4. Impact of transmit and receive antenna correlation on the performance of space-frequency Trellis codes.

fact that it orthogonalizes the channel irrespectively of the channel realization. This results in a performance which is independent of the rank properties of the channel. For spatial multiplexing the situation is different since the performance depends significantly 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. 3). If the total transmit angle spread is small, the two subspaces tend to be aligned so that certain transmit signal vectors which happen to lie in the orthogonal complement of this subspace will result in 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). An alternative interpretation of this phenomenon can be found at the end of Section V.

B. Simulation Example 2

This example investigates the impact of transmit and receive antenna correlation on the performance of space-frequency codes. Fig. 4 displays the block error rate for a system employing the two-antenna 16-state QPSK Trellis code proposed in [3] as inner code. For $M_T = 2$ and $M_R = 3$, we show the cases of no antenna 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. 4 shows that best performance is achieved in the i.i.d. case, 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 is due to the different number of transmit and receive antennas. More specifically, in the i.i.d. case the channel offers diversity order $d = M_T M_R L = 12$. In the following, we make the simplifying assumption that $|\rho(\Delta, \bar{\theta}_l, \sigma_{\theta_l})| = |\rho(2\Delta, \bar{\theta}_l, \sigma_{\theta_l})| = 1$ 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 (22) as $d = M_R(r(\mathbf{S}_0) + r(\mathbf{S}_1)) = 6$, whereas in the case of receive correlation only $d \leq M_T(r(\mathbf{R}_0) + r(\mathbf{R}_1)) = 4$. These differences in the achievable diversity order are reflected in the error rate performance depicted in Fig. 4. 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.

C. Simulation Example 3

This simulation example compares the performance of orthogonal and nonorthogonal space-frequency codes in the presence of transmit correlation.[1] In Section IV, we concluded that the impact of transmit antenna correlation is highly dependent on how the code excites 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 zero to one. The correlation matrix was chosen to be identical for both taps. Fig. 5 shows the resulting bit-error rates assuming Gray encoding. It is clearly seen that while 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 in Section VI-A lead us to the conclusion that space-frequency block codes exhibit superior robustness with respect to varying propagation conditions when compared to high-rate schemes such as spatial multiplexing.

D. 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 binary phase-shift keying (BPSK) code to a simple BPSK-based delay diversity scheme with delay 4 (achieved by premultiplying the signals with a linear phase

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

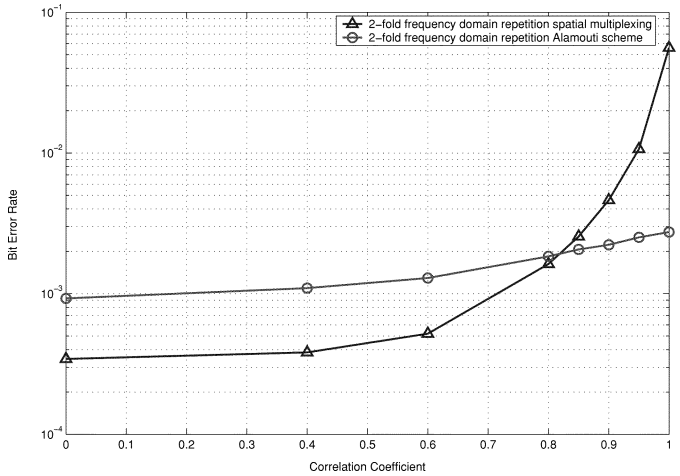

Fig. 5. Impact of varying transmit antenna correlation on performance of space-frequency codes.

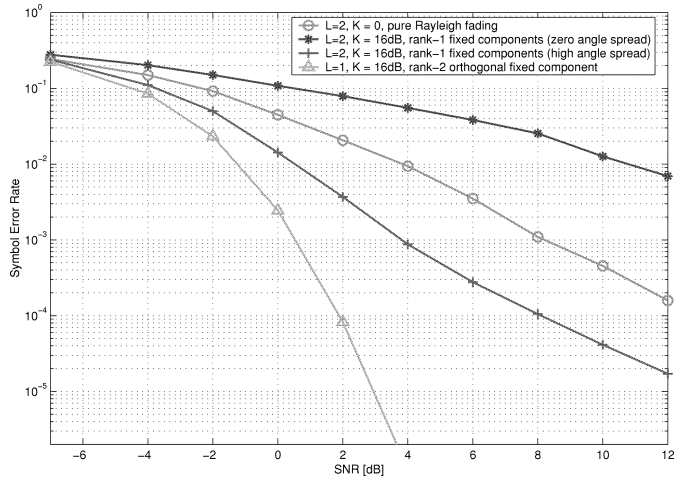

Fig. 6. Impact of K-factor and geometry of Ricean component on performance of MIMO-OFDM spatial multiplexing.

factor in the frequency domain). In both cases, we used the rate 1/2, constraint length-4 convolutional code proposed in [26] 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 (one symbol per vector transmission) is half the spatial rate of spatial multiplexing (two 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 a minimum mean-square error (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. 6 and 7, respectively, are the following.

- (Scenario 1) Pure Rayleigh fading on both taps, i.e., $K_0 = K_1 = 0$.

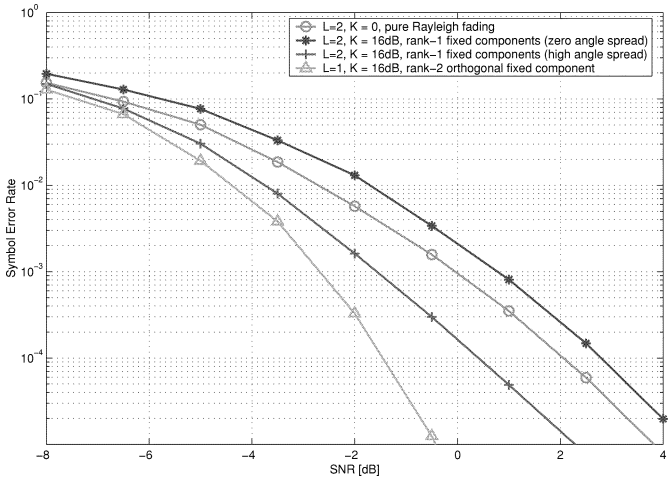

Fig. 7. Impact of K-factor and geometry of Ricean component on performance of delay diversity combined with convolutional code.

- (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 between the paths at the two taps is zero, so $\bar{\mathbf{H}}_0 = \bar{\mathbf{H}}_1$. 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 $\pi/2$ between 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 $\pi/2$ for both the angle of arrival and departure.
- (Scenario 4) Flat-fading scenario $(L = 1)$, 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.

1) Spatial Multiplexing: From Fig. 6 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 $\bar{\mathbf{H}}(e^{j(2\pi/N)k})$ have rank 1 for $k = 0, 1, \ldots, N-1$, thus severely impeding spatial multiplexing performance. Scenario 4 leads to good performance since $\bar{\mathbf{H}}(e^{j(2\pi/N)k}) = \bar{\mathbf{H}}_0$ is orthogonal for $k = 0, 1, \ldots, N-1$. 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 taps results in full-rank $\bar{\mathbf{H}}(e^{j(2\pi/N)k})$, even though $\bar{\mathbf{H}}_0$ and $\bar{\mathbf{H}}_1$ have rank 1. In Scenario 3, if only one tap were present the $\bar{\mathbf{H}}(e^{j(2\pi/N)k})$ $(k = 0, 1, \ldots, N-1)$ 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 single-input single-output Ricean case, where frequency-selectivity leads to a performance degradation when compared to the frequency-flat case.

2) 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 $\bar{\mathbf{H}}(e^{j(2\pi/N)k})$ are less important than in the case of spatial multiplexing.

VII. Conclusion

We studied the impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM systems. In particular, we quantified the achievable diversity order and coding gain as a function of the propagation parameters. The presence of transmit correlation was shown to result in widely varying performance losses while the presence of receive correlation in general affects all space-frequency codes equally. Moreover, we established that orthogonal space-frequency codes exhibit maximum robustness with respect to varying propagation conditions, whereas the performance of MIMO-OFDM spatial multiplexing is very sensitive to changes in the propagation conditions. Finally, we provided simulation results quantifying our analytical results.

References

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

[2] N. Seshadri and J. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Inform. Networks*, vol. 1, no. 1, pp. 49–60, 1994.

[3] 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, pp. 744–765, Mar. 1998.

[4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[5] 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.

[6] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, pp. 665–675, July 1985.

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

[8] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[9] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE Wireless Communications Networking Conf. (WCNC)*, Chicago, IL, Sept. 2000, pp. 1–6.

[10] ——, "Space-frequency codes for broadband fading channels," in *Proc. IEEE ISIT-2001*, Washington, DC, June 2001, p. 219.

[11] 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.

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

[13] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Space–time codes performance criteria and design for frequency selective fading channels," in *Proc. IEEE Int. Conf. Communications (ICC)*, vol. 9, Helsinki, Finland, June 2001, pp. 2800–2804.

[14] S. Zhou and G. B. Giannakis, "Space–time coding with maximum diversity gains over frequency-selective fading channels," *IEEE Signal Processing Lett.*, vol. 8, pp. 269–272, Oct. 2001.

[15] C. N. Chuah, D. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, Mar. 2002.

[16] D. Asztély, "On antenna arrays in mobile communication systems: fast fading and GSM base station receiver algorithms," Royal Inst. Technol., Stockholm, Sweden, Tech. Rep. IR-S3-SB-9611, Mar. 1996.

[17] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[18] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*. New York: Marcel Dekker, 1992.

[19] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Outage performance of space–time block codes for generalized MIMO channels," *IEEE Trans. Inform. Theory*, Mar. 2002, submitted for publication.

[20] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. New York: Cambridge Univ. Press, 1991.

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

[22] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. San Diego, CA: Academic, 1994.

[23] R. M. Gray. (2002, Aug.) Toeplitz and circulant matrices. ISL, Tech. Rep., Stanford Univ., Stanford, CA. [Online]. Available: http://ee-www.stanford.edu/~gray/toeplitz.html

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

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

[26] P. Frenger, P. Orten, and T. Ottosson, "Convolutional codes with optimum distance spectrum," *IEEE Commun. Lett.*, vol. 3, pp. 317–319, Nov. 1999.

Helmut Bölcskei (M'98–SM'02) was born in Austria on May 29, 1970. He received the Dr. techn. degree in electrical engineering from Vienna University of Technology, Vienna, Austria, in 1997.

In 1998, he was with Vienna University of Technology. From 1999 to 2001, he was a Postdoctoral Researcher in the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA. During that period, he was also a consultant for Iospan Wireless Inc., San Jose, CA. From 2001 to 2002, he was an Assistant Professor of electrical engineering, University of Illinois at Urbana-Champaign. Since February 2002, he has been an Assistant Professor of communication theory at ETH Zurich, Zurich, Switzerland. He was a Visiting Researcher at Philips Research Laboratories, Eindhoven, The Netherlands, ENST Paris, France, and the Heinrich Hertz Institute Berlin, Germany. His research interests include communication and information theory with special emphasis on wireless communications.

Dr. Bölcskei received a 2001 IEEE Signal Processing Society Young Author Best Paper Award and was an Erwin Schrödinger Fellow (1999–2001) of the Austrian National Science Foundation (FWF). He serves as an Associate Editor for the IEEE Transactions on Wireless Communications and the IEEE Transactions on Signal Processing.

Moritz Borgmann (S'99) was born in Hamburg, Germany. After undergraduate studies at the Technische Universität München, Munich, Germany, he received the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2000. Since 2002, he has been working toward the Dr. sc. techn. degree at ETH Zurich, Zurich, Switzerland. He held a permanent scholarship from the German National Scholarship Foundation.

He was a Visiting Researcher at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, in 2001. His research interests include coding and modulation for broadband wireless MIMO channels.

Arogyaswami J. Paulraj (F'91) received the Ph.D. degree in 1973. He was educated at the Naval Engineering College and the Indian Institute of Technology, India.

He is a Professor in the Department of Electrical Enginering, Stanford University, Stanford, CA, where he supervises the Smart Antennas Research Group working on applications of space–time techniques for wireless communications. His non-academic positions included Head of Sonar Division, Naval Oceanographic Laboratory, Cochin; Director of the Center for Artificial Intelligence and Robotics, Bangalore; Director of the Center for Development of Advanced Computing; Chief Scientist at Bharat Electronics, Bangalore, and Chief Technical Officer and Founder of Iospan Wireless, Inc. 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 broadband wireless systems. He has won several awards for his engineering and research contributions. He is the author of over 280 research papers and holds eight patents.

Dr. Paulraj is a Member of the Indian National Academy of Engineering.