Code Design for Non-Coherent MIMO-OFDM Systems*

Helmut Bölcskei and Moritz Borgmann

Communication Technology Laboratory, ETH Zurich
ETH Zentrum, ETF E122, Sternwartstrasse 7, 8092 Zürich, Switzerland
Email: boelcskei{moriborg}@nari.ee.ethz.ch

Abstract

Recently, the use of *coherent* space-frequency coding in Orthogonal Frequency Division Multiplexing (OFDM)-based frequency-selective multiple-input multiple-output (MIMO) antenna fading channels has been proposed [1, 2]. Acquiring knowledge of the fading coefficients in a MIMO channel is already very challenging in the flat-fading case. In the frequency-selective case, estimating the MIMO channel becomes significantly more difficult due to the presence of multiple paths, which results in an increased number of channel coefficients. In this paper, we address code design for *non-coherent frequency-selective MIMO-OFDM fading links*, where neither the transmitter nor the receiver knows the channel. We derive the *design criteria* and quantify the maximum achievable diversity gain. We demonstrate that unlike in the coherent case, non-coherent space-frequency codes designed to achieve full spatial diversity in the flat-fading case can fail completely to exploit not only frequency diversity but also spatial diversity when used in frequency-selective fading environments. Such codes are termed "catastrophic". Finally, we provide *explicit constructions of full-diversity (space and frequency) achieving codes* and assess their performance through simulation results.

1 Introduction and Outline

Space-time coding has emerged as a promising technique for realizing spatial diversity gain in multiple-input multiple-output (MIMO) antenna systems. Previous work on space-time coding has been restricted to coherent [3, 4, 5] and non-coherent [6, 7, 8] narrowband flat-fading MIMO channels, where spatial diversity only is available. Recently, the use of MIMO systems in frequency-selective fading channels in combination with *Orthogonal Frequency Division Multiplexing* (OFDM) and coherent detection has been considered [9, 10, 1, 11, 2, 12, 13]. Code design criteria for the coherent frequency-selective case were derived in [1, 11], and specific code designs have been presented in [2, 14].

While the coherent case is representative for fixed or low mobility wireless systems, future mobile broadband access systems are expected to operate at high vehicle speeds and will hence experience fast frequency-selective fading. Learning the coefficients of a MIMO fading channel is already very challenging in the narrowband frequency-flat fading case. Often this is not possible at all due to fast channel variations. In the broadband case,

*This research was supported in part by ETH Grant TH-6 02-2.

estimating the channel becomes even more challenging due to the presence of multiple paths, which requires the estimation of an increased number of fading coefficients.

Contributions. Considering a strategy which consists of coding across antennas and OFDM tones (*space-frequency coding*), we derive the design criteria for non-coherent space-frequency codes, where neither the transmitter nor the receiver knows the channel. The resulting design rules are found to be vastly different from the well-known Hochwald-Marzetta criteria for the flat-fading case [7]. We furthermore demonstrate that employing known non-coherent space-time codes as space-frequency codes (by coding across frequency rather than across time) will in general not exploit the additionally available frequency diversity. Moreover, we show the existence of "catastrophic non-coherent codes", which are non-coherent space-time codes designed to achieve full spatial diversity in the flat-fading case and fail to exploit *both spatial and frequency diversity* when employed in frequency-selective channels. The non-coherent broadband case therefore calls for new code designs. We present *explicit constructions of full (i.e. spatial and frequency) diversity achieving codes* and derive the *maximum-likelihood (ML) decoding rule*. We prove that the maximum achievable diversity order is given by $d_{max} = M_T M_R L$ with M_T and M_R denoting the number of transmit and receive antennas, respectively, and L denoting the channel order. Finally, we assess the performance of the proposed space-frequency codes through simulation results.

Outline. The paper is organized as follows. Sec. 2 introduces the system model and discusses our assumptions. In Sec. 3, we derive the design criteria and discuss performance bounds. Sec. 4 presents our code constructions and simulation results, and Sec. 5 contains conclusions.

2 Non-Coherent MIMO-OFDM Systems

In this section, we introduce the non-coherent broadband fading channel model, and then briefly describe non-coherent MIMO-OFDM systems and the idea of non-coherent space-frequency coding.

2.1 Channel Model

In the following, M_T and M_R denote the number of transmit and receive antennas, respectively. We assume that the channel consists of L matrix-valued taps (each of size M_R by M_T) with the matrix-valued transfer function given by

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

Note that in general there will be a continuum of delays. The channel model (1) is derived from the assumption of having L resolvable paths, where $L = \lfloor B\tau \rfloor$ with B and τ denoting the signal bandwidth and delay spread, respectively. For the sake of simplicity, we restrict ourselves to purely Rayleigh fading channels. The elements of the $\mathbf{H}_l \, (l = 0, 1, ..., L-1)$ are circularly symmetric[1] zero mean uncorrelated complex gaussian random variables with variance σ_l^2. For the sake of simplicity of exposition, throughout the paper we assume uncorrelated taps $\mathbf{H}_l$ and a uniform power delay profile with $\sigma_l^2 = \frac{1}{L} \, (l = 0, 1, ..., L-1)$.

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

2.2 Non-Coherent MIMO-OFDM Systems

In an OFDM-based MIMO system the individual data streams are OFDM-modulated before transmission (Fig. 1 shows the schematic of a MIMO-OFDM system). In the following, N denotes the number of OFDM tones. The OFDM modulator applies an N-point IFFT to N consecutive data symbols and then prepends the cyclic prefix (CP) (which is a copy of the last L samples of the symbol) to the symbol. In the receiver, the individual signals are passed through an OFDM demodulator, which first discards the CP and then applies an N-point FFT. Organizing the transmitted data symbols into frequency vectors $\mathbf{c}_k = [c_k^{(0)} \ c_k^{(1)} \ \ldots \ c_k^{(M_T-1)}]^T$ $(k = 0, 1, ..., N-1)$ 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 [9, 10]

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

where $\mathbf{w}_k$ is an $M_R \times 1$ complex-valued additive white gaussian noise vector satisfying[2]

$$\mathcal{E}\{\mathbf{w}_k \mathbf{w}_l^H\} = \mathbf{I}_{M_R}\delta[k-l] \tag{3}$$

with $\mathbf{I}_{M_R}$ denoting the identity matrix of size M_R.

Throughout this paper, we employ a generalization of the block-fading model used in [6] to the broadband case. We assume that the delay-spread channel remains constant for the duration of one OFDM symbol and then changes to a new realization in an independent fashion. We emphasize that due to the presence of multipath delay-spread the individual channel gain matrices $\mathbf{H}(e^{j\frac{2\pi}{N}k})$ $(k = 0, 1, ..., N-1)$ will be correlated. Note that in the narrowband case considered in [6, 7, 8], where coding is performed across space and time rather than across space and frequency, the channel is flat in the sense that the channel gain matrix $\mathbf{H}$ remains constant within the entire coherence interval.

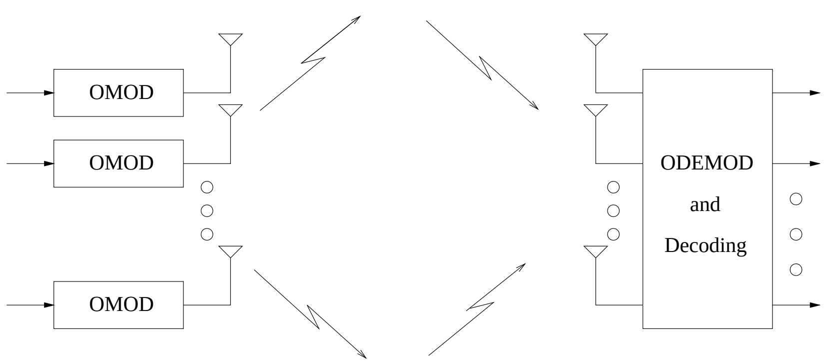

Figure 1: Schematic of a MIMO-OFDM system.

3 Non-Coherent Space-Frequency Codes: Design Criteria and Performance Bounds

In this section, we derive the design criteria for non-coherent space-frequency codes, and we compute ultimate performance limits by quantifying the maximum achievable

[2]The superscripts ^{T,H,*} stand for transpose, conjugate transpose, and elementwise conjugation, respectively. $\mathcal{E}$ denotes the expectation operator.

diversity order and coding gain. Moreover, we establish the existence of “catastrophic” non-coherent space-frequency codes.

Assumptions. The bit stream to be transmitted is encoded by the space-frequency encoder into blocks of size $N \times M_T$. One data burst therefore consists of M_T vectors of size $N \times 1$. We adopt the following notation. The $N \times M_T$ codeword matrix corresponding to the transmitted vectors $\mathbf{c}_k$ $(k = 0, 1, ..., N-1)$ is denoted as $\mathbf{C} = [\mathbf{c}_0 \;\; \mathbf{c}_1 \;\; ... \;\; \mathbf{c}_{N-1}]^T$, the corresponding received $N \times M_R$ matrix is given by $\mathbf{R} = [\mathbf{r}_0 \;\; \mathbf{r}_1 \;\; ... \;\; \mathbf{r}_{N-1}]^T$ with $\mathbf{r}_k$ defined in (2). Furthermore, we have $\mathbf{R} = \mathbf{Y} + \mathbf{W}$, where

$$\mathbf{Y} = \sqrt{E_s}[\mathbf{H}(e^{j2\pi\frac{0}{N}})\mathbf{c}_0 \;\; \mathbf{H}(e^{j2\pi\frac{1}{N}})\mathbf{c}_1 \;\; ... \;\; \mathbf{H}(e^{j2\pi\frac{N-1}{N}})\mathbf{c}_{N-1}]^T$$

and $\mathbf{W} = [\mathbf{w}_0 \;\; \mathbf{w}_1 \;\; ... \;\; \mathbf{w}_{N-1}]^T$ with $\mathbf{w}_k$ defined in (3). Now, since the individual elements of the $\mathbf{H}_l$ satisfy $[\mathbf{H}_l]_{m,n} \sim \mathcal{CN}(0, \frac{1}{L})$, it follows that the $\mathbf{H}(e^{j\frac{2\pi}{N}k})$ $(k = 0, 1, ..., N-1)$ are circularly symmetric complex gaussian with $[\mathbf{H}(e^{j2\pi\frac{k}{N}})]_{m,n} \sim \mathcal{CN}(0, 1)$.

The conditional received signal density. The columns of the $N \times M_R$ matrix $\mathbf{R}$ are independent and have the same covariance matrix. Assuming that the codeword matrix $\mathbf{C}_i$ was transmitted, this covariance matrix is given by

$$\mathbf{V}(\mathbf{C}_i) = \mathbf{I}_N + \rho \sum_{l=0}^{L-1} \mathbf{D}^l \mathbf{C}_i \mathbf{C}_i^H \mathbf{D}^{-l}, \tag{4}$$

where $\mathbf{D} = \mathrm{diag}\{e^{-j\frac{2\pi}{N}k}\}_{k=0}^{N-1}$ and $\rho = \frac{E_s}{L}$. (Note that $\rho = E_s/L$ does not denote a normalized SNR. The SNR per receive antenna is given by $E_s M_T$.) Defining the stacked $N \times M_T L$ matrix

$$\mathbf{G}(\mathbf{C}_i) = [\mathbf{C}_i \;\; \mathbf{D}\mathbf{C}_i \;\; ... \;\; \mathbf{D}^{L-1}\mathbf{C}_i], \tag{5}$$

we can rewrite (4) as

$$\mathbf{V}(\mathbf{C}_i) = \mathbf{I}_N + \rho \, \mathbf{G}(\mathbf{C}_i)\mathbf{G}^H(\mathbf{C}_i).$$

Noting that the entries of $\mathbf{R}$ are circularly symmetric complex gaussian and using the fact that the individual columns of $\mathbf{R}$ are independent, the conditional received signal density is obtained as[3]

$$p(\mathbf{R}|\mathbf{C}_i) = \frac{\exp\left(-\mathrm{Tr}\left(\mathbf{R}^H \mathbf{V}^{-1}(\mathbf{C}_i)\mathbf{R}\right)\right)}{\pi^{NM_R}|\mathbf{V}(\mathbf{C}_i)|^{M_R}}. \tag{6}$$

Maximum likelihood (ML) decoding. Assuming a space-frequency constellation of K codeword matrices $\mathbf{C}_i$ $(i = 0, 1, ..., K-1)$, and using (6), the ML decoding rule follows as

$$\widehat{\mathbf{C}}_{ML} = \arg \min_{\mathbf{C}_i \in \{\mathbf{C}_0, \mathbf{C}_1, ..., \mathbf{C}_{K-1}\}} \left(\mathrm{Tr}\left(\mathbf{R}^H \mathbf{V}^{-1}(\mathbf{C}_i)\mathbf{R}\right) + M_R \ln|\mathbf{V}(\mathbf{C}_i)|\right).$$

In the following, we assume that the codeword matrices are such that $\mathbf{G}^H(\mathbf{C}_i)\mathbf{G}(\mathbf{C}_i) = \mathbf{I}_{M_T L}$ $(i = 0, 1, ..., K-1)$. We note that unitarity of the $\mathbf{G}(\mathbf{C}_i)$ has to be ensured explicitly by code design. Applying the matrix inversion lemma to $\mathbf{V}^{-1}(\mathbf{C}_i)$ and noting that $\ln|\mathbf{V}(\mathbf{C}_i)|$ becomes a constant independent of i, the ML decoding rule for unitary $\mathbf{G}(\mathbf{C}_i)$ simplifies to

$$\begin{aligned}\widehat{\mathbf{C}}_{ML} &= \arg \max_{\mathbf{C}_i \in \{\mathbf{C}_0, \mathbf{C}_1, ..., \mathbf{C}_{K-1}\}} \mathrm{Tr}\left(\mathbf{R}^H \mathbf{G}(\mathbf{C}_i)\mathbf{G}^H(\mathbf{C}_i)\mathbf{R}\right) \\ &= \arg \max_{\mathbf{C}_i \in \{\mathbf{C}_0, \mathbf{C}_1, ..., \mathbf{C}_{K-1}\}} \mathrm{Tr}\left(\mathbf{R}^H \left(\sum_{l=0}^{L-1} \mathbf{D}^l \mathbf{C}_i \mathbf{C}_i^H \mathbf{D}^{-l}\right) \mathbf{R}\right).\end{aligned}$$

[3] $|\mathbf{A}|$ and $\mathrm{Tr}(\mathbf{A})$ stand for the determinant and the trace of the matrix $\mathbf{A}$, respectively.

Pairwise error probability. We are now in a position to derive the pairwise error probability (PEP) for non-coherent unitary space-frequency codes. Assume that the matrix $\mathbf{C}_i$ was transmitted and that ML decoding is performed. The Chernoff upper bound on the average (with respect to the channel) probability of mistaking $\mathbf{C}_i$ for another codeword, say $\mathbf{C}_j$, is given by [15]

$$P(\mathbf{C}_i \rightarrow \mathbf{C}_j|\mathbf{C}_i) \leq \prod_{r=0}^{M_T L-1} \left(\frac{1}{1+\frac{\rho^2}{4(1+\rho)}(1-\lambda_r^2(i,j))} \right)^{M_R}, \tag{7}$$

where $\lambda_r(i,j)$ denotes the r-th singular value of the $M_T L \times M_T L$ matrix

$$\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i) = \begin{bmatrix} \mathbf{C}_j^H\mathbf{C}_i & \mathbf{C}_j^H\mathbf{D}\mathbf{C}_i & \dots & \mathbf{C}_j^H\mathbf{D}^{L-1}\mathbf{C}_i \\ \mathbf{C}_j^H\mathbf{D}^{-1}\mathbf{C}_i & \mathbf{C}_j^H\mathbf{C}_i & \dots & \mathbf{C}_j^H\mathbf{D}^{L-2}\mathbf{C}_i \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{C}_j^H\mathbf{D}^{-L+1}\mathbf{C}_i & \mathbf{C}_j^H\mathbf{D}^{-L+2}\mathbf{C}_i & \dots & \mathbf{C}_j^H\mathbf{C}_i \end{bmatrix}.$$

From (7) it follows that the PEP upper bound is lowest when all singular values $\lambda_r(i,j)$ $(r = 0, 1, ..., M_T L-1)$ are equal to zero and highest when all singular values $\lambda_r(i,j)$ $(r = 0, 1, ..., M_T L-1)$ are equal to one. The ideal constellation therefore has all the columns of $\mathbf{G}(\mathbf{C}_l)$ orthogonal to all the columns of $\mathbf{G}(\mathbf{C}_{l'})$ for $l \neq l'$. Note, however, that since all the columns within the $\mathbf{G}(\mathbf{C}_l)$ are orthogonal to each other, all the pairwise $\lambda_r(i,j)$ cannot all be made zero if $K > \frac{N}{M_T L}$. Especially for large L, satisfying this condition will yield very low code rates.

PEP time-domain formulation. Set[4] $\mathbf{C} = \mathbf{F}\mathbf{C}_t$ with $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{N}}\, e^{-j2\pi\frac{mn}{N}}$ denoting the $N \times N$ FFT matrix. Next, note that $\mathbf{F}^H\mathbf{D}^l\mathbf{F}\mathbf{C}_t = \mathbf{C}_{t-l}$, where $\mathbf{C}_{t-l}$ denotes the matrix obtained by cyclically shifting the columns of $\mathbf{C}_t$ by l positions down. It then follows that

$$\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i) = \mathbf{G}^H(\mathbf{C}_{j,t})\mathbf{G}(\mathbf{C}_{i,t})$$

with

$$\mathbf{G}(\mathbf{C}_{i,t}) = [\mathbf{C}_{i,t} \;\; \mathbf{C}_{i,t-1} \;\; \dots \;\; \mathbf{C}_{i,t-L+1}].$$

This result suggests an interpretation of the additional taps inducing "virtual antennas", which transmit delayed versions of $\mathbf{C}_{i,t}$. In the frequency-flat fading case the PEP is determined by the distance between the subspaces spanned by the columns of the codeword matrices $\mathbf{C}_{i,t}$. In the frequency-selective case, this distance is replaced by the distance between the subspaces spanned by the pseudo-codewords $\mathbf{G}(\mathbf{C}_{i,t})$.

Diversity order and coding gain. In the following, we shall focus on the high SNR case, where the PEP upper bound reduces to

$$P(\mathbf{C}_i \rightarrow \mathbf{C}_j|\mathbf{C}_i) \leq \left(\frac{\rho}{4}\right)^{-M_R s(i,j)} \prod_{r=0}^{s(i,j)-1} \frac{1}{(1-\lambda_r^2(i,j))^{M_R}}$$

with $s(i,j)$ denoting the number of singular values of $\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)$ that satisfy $\lambda_r(i,j) < 1$. Setting $s = \min_{i \neq j} s(i,j)$, we conclude that the diversity gain achieved by the code is given by $d = sM_R$, and the coding gain is obtained as the minimum of

$$\left(\prod_{r=0}^{s(i,j)-1} (1-\lambda_r^2(i,j) \right)^{\frac{1}{s(i,j)}}$$

[4]The subscript t indicates that $\mathbf{C}_t$ denotes the time-domain version of $\mathbf{C}$.

with the minimum taken over all $\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)$ that satisfy $s(i,j) = s$. Since $s(i,j) \leq M_T L$, the maximum achievable diversity order is given by $d_{max} = M_T M_R L$, and the coding gain is maximized if $\lambda_r(i,j) = 0$ for $r = 0, 1, ..., s(i,j) - 1$ for all pairs $\{i,j\}$ with $s(i,j) = s$.

Catastrophic codes. Note that designing a code such that the singular values of $\mathbf{C}_j^H \mathbf{C}_i$ are small will in general not guarantee that the singular values of the $M_T L \times M_T L$ matrix $\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)$ are small as well. In other words, a code that performs well in the frequency-flat fading case (i.e. $L = 1$) will in general not exploit the additionally available frequency diversity. The situation can be even worse as the following example shows. Assume we have a code for $M_T = M_R = 1$ and $L = 1$ such that all the code vectors $\mathbf{c}_i$ are mutually orthogonal, which constitutes the best possible case (but at the same time very low code rate) in the presence of flat-fading. With this assumption, the high SNR PEP upper bound exhibits inverse proportional behavior to ρ, i.e.,

$$P(\mathbf{C}_i \to \mathbf{C}_j | \mathbf{C}_i) \leq \frac{1}{1 + \rho/4}.$$

Using this code in a scenario where $M_T = M_R = 1$ and $L = 2$ (i.e., frequency-selective fading), the PEP upper bound is determined by the singular values of (for the sake of clarity of exposition we work in the time-domain)

$$\mathbf{G}^H(\mathbf{C}_{j,t})\mathbf{G}(\mathbf{C}_{i,t}) = \begin{bmatrix} \mathbf{c}_{j,t}^H \mathbf{c}_{i,t} & \mathbf{c}_{j,t}^H \mathbf{c}_{i,t-1} \\ \mathbf{c}_{j,t-1}^H \mathbf{c}_{i,t} & \mathbf{c}_{j,t}^H \mathbf{c}_{i,t} \end{bmatrix}, \tag{8}$$

where we exploited $\mathbf{c}_{j,t-1}^H \mathbf{c}_{i,t-1} = \mathbf{c}_{j,t}^H \mathbf{c}_{i,t}$. Since the $\mathbf{c}_i$ were assumed to be mutually orthogonal, (8) simplifies to

$$\mathbf{G}^H(\mathbf{C}_{j,t})\mathbf{G}(\mathbf{C}_{i,t}) = \begin{bmatrix} 0 & \mathbf{c}_{j,t}^H \mathbf{c}_{i,t-1} \\ \mathbf{c}_{j,t-1}^H \mathbf{c}_{i,t} & 0 \end{bmatrix}.$$

The singular values of $\mathbf{G}^H(\mathbf{C}_{j,t})\mathbf{G}(\mathbf{C}_{i,t})$ are given by $\lambda_1 = |\mathbf{c}_{j,t}^H \mathbf{c}_{i,t-1}|$ and $\lambda_2 = |\mathbf{c}_{j,t-1}^H \mathbf{c}_{i,t}|$, which results in

$$P(\mathbf{C}_i \to \mathbf{C}_j | \mathbf{C}_i) \leq \frac{1}{1 + \frac{\rho}{4}(1 - |\mathbf{c}_{j,t}^H \mathbf{c}_{i,t-1}|^2)} \frac{1}{1 + \frac{\rho}{4}(1 - |\mathbf{c}_{j,t-1}^H \mathbf{c}_{i,t}|^2)}.$$

Now, if $\mathbf{c}_j$ and $\mathbf{c}_i$ are such that $|\lambda_1| < 1$ and $|\lambda_2| < 1$, the PEP decays as ρ^{-2}, and we achieve second-order diversity. If only one singular value has magnitude less than 1, the PEP is proportional to ρ^{-1}, and hence we get the same diversity order as in the frequency-flat fading case. Finally, if $|\lambda_1| = |\lambda_2| = 1$, the PEP upper bound becomes 1, and the diversity order is 0, i.e., the performance is worse than in the frequency-flat fading case. Any non-coherent code exhibiting such a behavior is termed "catastrophic". An example of a "catastrophic" code-vector pair are the 2-periodic vectors $\mathbf{c}_{i,t} = [0 \;\; c \;\; 0 \;\; c \;\; \ldots \;\; c]^T$ and $\mathbf{c}_{j,t} = \mathbf{c}_{i,t-1}$ with $|c| = \sqrt{\frac{2}{N}}$. It is easily verified that these vectors lead to $|\lambda_1| = |\lambda_2| = 1$.

The existence of catastrophic codes is striking as it shows a fundamental difference between the coherent and the non-coherent frequency-selective cases. In the coherent case, it was shown in [14] that a code designed to achieve a certain diversity order in the frequency-flat case achieves at least the same diversity order in the frequency-selective case. In other words, using a code designed for the frequency-flat fading case does not degrade diversity order in the frequency-selective case (the coding gain may still be smaller though). In the non-coherent case, codes designed to achieve a certain diversity order under flat-fading conditions are not guaranteed to achieve at least the same diversity order in a frequency-selective environment.

Simplified design criterion. The design criteria for non-coherent space-frequency codes discussed above (maximization of diversity order and coding gain) are often difficult to incorporate into the code design procedure. In [8] a simplified design criterion is derived which aims at minimizing

$$\delta = \max_{0\leq j<i\leq K-1} \|\mathbf{C}_j^H \mathbf{C}_i\|_F^2,$$

where $\|\mathbf{A}\|_F^2 = \sum_{i,j} |[\mathbf{A}]_{i,j}|^2$ denotes the Frobenius norm of the matrix $\mathbf{A}$. In the frequency-selective case, this simplified criterion generalizes to a minimization of

$$\delta = \max_{0\leq j<i\leq K-1} \|\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)\|_F^2. \tag{9}$$

Similarly to [8], we can interpret δ as a measure of the distance between the subspaces spanned by $\mathbf{G}(\mathbf{C}_j)$ and $\mathbf{G}(\mathbf{C}_i)$. We emphasize that in the frequency-flat case treated in [8], the distance is measured between the subspaces spanned by the columns of the codewords $\mathbf{C}_j$ and $\mathbf{C}_i$. In the frequency-selective case, the distance is measured by the "induced" subspaces spanned by the pseudo-codewords $\mathbf{G}(\mathbf{C}_j)$ and $\mathbf{G}(\mathbf{C}_i)$. We note furthermore that $\|\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)\|_F^2$ can be expressed directly in terms of the codewords $\mathbf{C}_i$ and $\mathbf{C}_j$ as

$$\|\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)\|_F^2 = L\|\mathbf{C}_j^H\mathbf{C}_i\|_F^2 + \sum_{r=1}^{L-1}(L-r)\left(\|\mathbf{C}_j^H\mathbf{D}^r\mathbf{C}_i\|_F^2 + \|\mathbf{C}_j^H\mathbf{D}^{-r}\mathbf{C}_i\|_F^2\right). \tag{10}$$

It is instructive to specialize (10) to the case $M_T = 1$ transmit antennas and $L = 2$ taps, where

$$\|\mathbf{G}^H(\mathbf{C}_j)\mathbf{G}(\mathbf{C}_i)\|_F^2 = 2|\mathbf{c}_j^H\mathbf{c}_i|^2 + |\mathbf{c}_j^H\mathbf{D}\mathbf{c}_i|^2 + |\mathbf{c}_j^H\mathbf{D}^{-1}\mathbf{c}_i|^2$$

or equivalently in terms of the time-domain code vectors

$$\|\mathbf{G}^H(\mathbf{C}_{j,t})\mathbf{G}(\mathbf{C}_{i,t})\|_F^2 = 2|\mathbf{c}_{j,t}^H\mathbf{c}_{i,t}|^2 + |\mathbf{c}_{j,t}^H\mathbf{c}_{i,t-1}|^2 + |\mathbf{c}_{j,t-1}^H\mathbf{c}_{i,t}|^2. \tag{11}$$

In [8], various approaches for minimizing the maximum of $|\mathbf{c}_{j,t}^H\mathbf{c}_{i,t}|$ over all i, j have been proposed. We can see that in the frequency-selective case minimizing the maximum of $|\mathbf{c}_{j,t}^H\mathbf{c}_{i,t}|$ alone will not minimize δ defined in (9). The frequency-selective nature of the channel has to be taken into account explicitly.

4 Code Design

We shall next present two code constructions that are capable of exploiting both space and frequency diversity. Our first construction is heavily inspired by a construction first proposed in [8]. We emphasize, however, that the construction in [8] does not take into account frequency-selectivity and does not ensure that the $\mathbf{G}(\mathbf{C}_i)$ are unitary. For the sake of simplicity of exposition, in the following, we consider a scenario with $M_T = 2$ transmit antennas and $L = 2$ taps. The codeword matrices $\mathbf{C}_i$ ($i = 0, 1, ..., K-1$) are chosen as

$$\mathbf{C}_i = \mathbf{\Phi}^i[\mathbf{f}_k \;\; \mathbf{f}_{k'}], \quad i = 0, 1, ..., K-1, \tag{12}$$

where $\mathbf{f}_k$ denotes the k-th column of the $N \times N$ FFT matrix $\mathbf{F}$, and

$$\mathbf{\Phi} = \text{diag}\{e^{j2\pi \frac{u_l}{K}}\}_{l=0}^{N-1},$$

where without loss of generality $0 \leq u_l \leq K-1$ for $l = 0, 1, ..., N-1$. Our construction contains two key elements, namely the choice of the parameters k and k' and the choice of the diagonal matrix $\mathbf{\Phi}$. The parameters k and k' have to be chosen such that the stacked matrix

$$\mathbf{B}_{k,k'} = [\mathbf{f}_k \;\; \mathbf{f}_{k'} \;\; \mathbf{D}\mathbf{f}_k \;\; \mathbf{D}\mathbf{f}_{k'}] = [\mathbf{f}_k \;\; \mathbf{f}_{k'} \;\; \mathbf{f}_{k+1} \;\; \mathbf{f}_{k'+1}], \tag{13}$$

is unitary. Here, we exploited the fact that $\mathbf{D}\mathbf{f}_k = \mathbf{f}_{k+1}$. Noting that $\mathbf{D}$ and $\mathbf{\Phi}^i$ ($i = 0, 1, ..., K-1$) commute, it then follows that choosing k and k' such that $\mathbf{B}_{k,k'}$ is orthonormal ensures that $\mathbf{G}(\mathbf{C}_i)$ is orthonormal and hence our space-frequency code is unitary. For the example under consideration, setting $k = 1$ and $k' = 3$, we obtain

$$\mathbf{B}_{k,k'} = [\mathbf{f}_1 \;\; \mathbf{f}_3 \;\; \mathbf{f}_2 \;\; \mathbf{f}_4], \tag{14}$$

which is simply a permutation of the first four columns of the $N \times N$ FFT matrix $\mathbf{F}$, and hence $\mathbf{B}_{k,k'}^H \mathbf{B}_{k,k'} = \mathbf{I}_4$. The parameters u_l in $\mathbf{\Phi}$ have to be chosen such that the PEP upper bound is minimized and have to be determined by random search. At first sight, an optimization seems difficult due to the high number of parameters involved. We note, however, that owing to the special structure of our codes we do not have to perform an optimization involving all $\frac{K(K-1)}{2}$ pairs of codeword matrices. This is so since $\mathbf{G}^H(\mathbf{C}_i)\mathbf{G}(\mathbf{C}_{i'})$ is a function of $i - i'$ only. Thus, we need to take into account $\mathbf{G}^H(\mathbf{C}_0)\mathbf{G}(\mathbf{C}_i)$ with $i = 1, 2, ..., K-1$ only. This drastically simplifies the process of finding the parameters u_l. We emphasize again, however, that optimizing a code for the frequency-flat fading case does not guarantee that it will exploit the additionally available frequency diversity. This statement will be illustrated by means of simulation results later in the paper.

Generalized construction. An alternative and more general construction is obtained by replacing $\mathbf{f}_k$ and $\mathbf{f}_{k'}$ in (12) by more general vectors $\mathbf{g}_k$ and $\mathbf{g}_{k'}$ (recall that $\mathbf{f}_k$ denotes the k-th column of the FFT matrix) to obtain

$$\mathbf{C}_i = \mathbf{\Phi}^i [\mathbf{g}_k \;\; \mathbf{g}_{k'}], \quad i = 0, 1, ..., K-1.$$

This construction exhibits more degrees of freedom as code optimization involves both finding the u_l as well as the vectors $\mathbf{g}_k$ and $\mathbf{g}_{k'}$. The vectors $\mathbf{g}_k$ and $\mathbf{g}_{k'}$ have to be chosen such that the stacked matrix

$$\mathbf{B}_{k,k'} = [\mathbf{g}_k \;\; \mathbf{g}_{k'} \;\; \mathbf{D}\mathbf{g}_k \;\; \mathbf{D}\mathbf{g}_{k'}]$$

satisfies $\mathbf{B}_{k,k'}^H \mathbf{B}_{k,k'} = \mathbf{I}_4$. Again, this choice of $\mathbf{g}_k$ and $\mathbf{g}_{k'}$, along with the unitarity of $\mathbf{\Phi}^i$ and the fact that $\mathbf{D}$ and $\mathbf{\Phi}^i$ commute for $i = 0, 1, ..., K-1$, ensures that $\mathbf{G}(\mathbf{C}_i)$ has orthonormal columns, and hence the code is unitary. Rewriting the vectors $\mathbf{g}_k$ and $\mathbf{g}_{k'}$ in terms of their time-domain versions as $\mathbf{g}_k = \mathbf{F}\mathbf{g}_{k,t}$ and $\mathbf{g}_{k'} = \mathbf{F}\mathbf{g}_{k',t}$, unitarity of $\mathbf{B}_{k,k'}$ is equivalent to unitarity of

$$\mathbf{B}_{k,k',t} = [\mathbf{g}_{k,t} \;\; \mathbf{g}_{k',t} \;\; \mathbf{g}_{k,t-1} \;\; \mathbf{g}_{k',t-1}],$$

where $\mathbf{g}_{k,t-l}$ is the vector resulting from cyclically shifting the elements of $\mathbf{g}_{k,t}$ by l positions down. The first construction above in terms of vectors of the FFT matrix is easily seen to result in an orthonormal $\mathbf{B}_{k,k',t}$ for $k=1$ and $k'=3$ as $\mathbf{f}_{k,t}$ is a vector with a 1 in the k-th position and zeros elsewhere.

Simulation example. We conclude with a simulation example where the first construction according to (12) was used with parameters $M_T = 2, M_R = 1, N = 10, K = 64, k = 1$, and $k' = 3$. Two scenarios were considered, namely $L = 1$ and $L = 2$. For both scenarios we found the parameters u_l through random search by maximizing the minimum of $\prod_{r=0}^{M_T L-1}(1-\lambda_r^2(i,j))$ over all pairs (i,j). The best parameter vectors we found were $\mathbf{u} = [1\ 2\ 5\ 12\ 16\ 25\ 29\ 43\ 51\ 59]$ for $L = 1$ and $\mathbf{u} = [5\ 10\ 10\ 22\ 23\ 38\ 39\ 40\ 51\ 53]$ for $L = 2$. From Fig. 2, we can see that for $L = 2$ the code optimized for $L = 2$ indeed achieves the maximum possible diversity order of $d_{max} = M_T M_R L = 4$. In the frequency-flat case (i.e., $L = 1$) the same code yields diversity order 2 and hence again achieves the maximum possible diversity order. However, the code optimized for the frequency-flat fading case ($L = 1$) does not exploit (all of) the additionally available frequency diversity when used in a 2-tap channel (i.e. $L = 2$). On the other hand, it does not exhibit catastrophic behavior either.

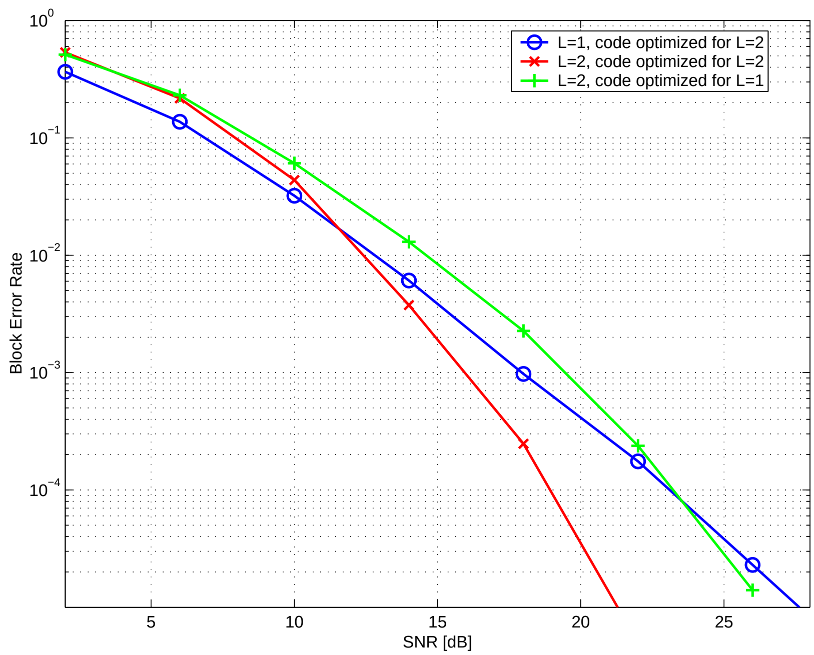

Figure 2: Performance of non-coherent space-frequency codes under frequency-flat fading and frequency-selective fading.

5 Conclusions

We considered space-frequency code design for non-coherent MIMO-OFDM links with frequency-selective fading. Using unitary constellations, we derived the code design criteria, quantified the maximum achievable diversity order and coding gain, and provided two alternative code constructions which can achieve full space-frequency diversity. Furthermore, we demonstrated the existence of "catastrophic" space-frequency codes, and found that non-coherent space-time codes designed for the flat-fading case will in general not exploit the additionally available frequency diversity when used in frequency-selective fading environments. Finally, we assessed the performance of some of our code constructions through simulation results.

References

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

[2] H. Bölcskei and A. J. Paulraj, "Space-frequency codes for broadband fading channels," in *Proc. IEEE ISIT-2001*, (Washington, D.C.), p. 219, June 2001.

[3] J. Guey, M. Fitz, M. Bell, and W. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE VTC*, pp. 136–140, 1996.

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

[5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, March 1998.

[6] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, pp. 139–157, Jan. 1999.

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

[8] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1962–1973, 2000.

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

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

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

[12] D. Agarwal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data rate wireless communication over wideband channels," in *Proc. IEEE VTC 98*, pp. 2232–2236, May 1998.

[13] H. Bölcskei, M. Borgmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Sel. Areas Comm.*, April 2002. submitted.

[14] H. Bölcskei, M. Borgmann, and A. J. Paulraj, "Space-frequency coded MIMO-OFDM with variable diversity-multiplexing tradeoff," *IEEE Trans. Inf. Theory*, 2002. to be submitted.

[15] H. Bölcskei and M. Borgmann, "Non-coherent space-frequency coded MIMO-OFDM," *IEEE Trans. Inf. Theory*, 2003. to be submitted.