Diversity-Multiplexing Tradeoff in Selective-Fading MIMO Channels

Pedro Coronel and Helmut Bölcskei
Communication Technology Laboratory
ETH Zurich, 8092 Zurich, Switzerland
E-mail: {pco, boelcskei}@nari.ee.ethz.ch

Abstract— **We establish the optimal diversity-multiplexing (DM) tradeoff of coherent time, frequency and time-frequency selective-fading MIMO channels and provide a code design criterion for DM-tradeoff optimality. Our results are based on the analysis of the "Jensen channel" associated to a given selective-fading MIMO channel. While the original problem seems analytically intractable due to the mutual information being a sum of correlated random variables, the Jensen channel is equivalent to the original channel in the sense of the DM-tradeoff and lends itself nicely to analytical treatment. Finally, as a consequence of our results, we find that the classical rank criterion for space-time code design (in selective-fading MIMO channels) ensures optimality in the sense of the DM-tradeoff.**

I. Introduction

The diversity-multiplexing (DM) tradeoff framework introduced by Zheng and Tse [1] allows to efficiently characterize the information-theoretic performance limits of communication over multiple-input multiple-output (MIMO) fading channels. In addition, the results in [1] have triggered significant activity on the design of DM-tradeoff optimal space-time codes. In particular, the *non-vanishing* determinant criterion [2], [3] on codeword difference matrices has been shown to constitute a sufficient condition for DM-tradeoff optimality in flat-fading MIMO channels with two transmit and two or more receive antennas [3]; this criterion has led to the construction of space-time codes based on constellation rotation [3], [4] and cyclic division algebras [5]. In [6] lattice-based space-time codes have been shown to be DM-tradeoff optimal. The DM-tradeoff optimality of *approximately universal* space-time codes was established in [7].

Contributions: While the results mentioned above focus on frequency-flat block-fading channels, extensions to frequency-selective channels can be found in [8], [9]. However, a general characterization of the optimal DM-tradeoff in time, frequency or time-frequency selective-fading MIMO channels, in the following simply referred to as selective-fading MIMO channels, remains an open problem. The present paper resolves this problem for the coherent case (i.e., for perfect channel state information (CSI) at the receiver) and provides a code design criterion guaranteeing DM-tradeoff optimality. Our results are based on exponentially tight (in the sense of exhibiting the same DM-tradeoff behavior) upper and lower bounds on the mutual information of (coherent) selective-fading MIMO channels. In particular, we show that the DM-tradeoff of this class of channels can be obtained by solving the analytically tractable problem of computing the DM-tradeoff curve corresponding to the associated "Jensen channel".

The first author was previously with IBM Research, Zurich Research Laboratory, Switzerland. This work was supported in part by the STREP project No. IST-026905 MASCOT within the Sixth Framework Programme of the European Commission.

Notation: $\mathrm{M_T}$ and $\mathrm{M_R}$ denote the number of transmit and receive antennas, respectively. We define $\mathrm{m} := \min(\mathrm{M_T}, \mathrm{M_R})$ and $\mathrm{M} := \max(\mathrm{M_T}, \mathrm{M_R})$. For $x \in \mathbb{R}$, we let $[x]^+ := \max(0, x)$. The superscripts ^{T}, ^{H} and * stand for transposition, conjugate transposition and complex conjugation, respectively. $\mathbf{I}_n$ is the $n \times n$ identity matrix, $\mathbf{A} \otimes \mathbf{B}$ and $\mathbf{A} \odot \mathbf{B}$ denote, respectively, the Kronecker and Hadamard products of the matrices $\mathbf{A}$ and $\mathbf{B}$, and $\mathbf{A} \succeq \mathbf{B}$ stands for the positive semidefinite ordering. If $\mathbf{A}$ has columns $\mathbf{a}_k$ $(k=1,2,\ldots,m)$, $\mathrm{vec}(\mathbf{A}) = [\mathbf{a}_1^T \ \mathbf{a}_2^T \ \ldots \ \mathbf{a}_m^T]^T$. For the $n \times m$ matrices $\mathbf{A}_k$ $(k=0,1,\ldots,K-1)$, $\mathrm{diag}\{\mathbf{A}_k\}_{k=0}^{K-1}$ denotes the $nK \times mK$ block-diagonal matrix with the kth diagonal entry given by $\mathbf{A}_k$. If $\mathcal{S}$ is a set, $|\mathcal{S}|$ denotes its cardinality. For index sets $\mathcal{S}_1 \subseteq \{1,2,\ldots,n\}$ and $\mathcal{S}_2 \subseteq \{1,2,\ldots,m\}$, $\mathbf{A}(\mathcal{S}_1, \mathcal{S}_2)$ stands for the (sub)matrix consisting of the rows of $\mathbf{A}$ indexed by $\mathcal{S}_1$ and the columns of $\mathbf{A}$ indexed by $\mathcal{S}_2$. The eigenvalues of the $n \times n$ Hermitian matrix $\mathbf{A}$, sorted in ascending order, are denoted by $\lambda_k(\mathbf{A})$, $k=1,2,\ldots,n$. The Kronecker delta function is defined as $\delta(m) = 1$ for $m = 0$ and zero otherwise. If X and Y are random variables (RVs), $X \sim Y$ denotes equality in distribution and $\mathbb{E}_X$ is the expectation operator with respect to (w.r.t.) the RV X. The random vector $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$ is multivariate circularly symmetric zero-mean complex Gaussian with $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{C}$. $f(x)$ and $g(x)$ are said to be exponentially equal, denoted by $f(x) \doteq g(x)$, if $\lim_{x\to\infty} \frac{\log f(x)}{\log x} = \lim_{x\to\infty} \frac{\log g(x)}{\log x}$. Exponential inequality, denoted by $\dot{\geq}$ and $\dot{\leq}$, is defined analogously.

II. Channel and Signal Model

The input-output relation for the class of MIMO channels considered in this paper is given by

$$\mathbf{y}_n = \sqrt{\frac{\mathsf{SNR}}{\mathrm{M_T}}}\mathbf{H}_n\mathbf{x}_n + \mathbf{z}_n, \quad n = 0,1,\ldots,N-1 \tag{1}$$

where the index n corresponds to a time, frequency or time-frequency slot and SNR denotes the signal-to-noise ratio at each receive antenna. The vectors $\mathbf{y}_n$, $\mathbf{x}_n$ and $\mathbf{z}_n$ denote, respectively, the corresponding $\mathrm{M_R} \times 1$ receive signal vector, $\mathrm{M_T} \times 1$ transmit signal vector, and $\mathrm{M_R} \times 1$ zero-mean circularly symmetric complex Gaussian noise vector satisfying $\mathbb{E}\{\mathbf{z}_n \mathbf{z}_{n'}^H\} = \delta(n-n')\,\mathbf{I}_{\mathrm{M_R}}$. We restrict our analysis to spatially uncorrelated Rayleigh fading channels so that, for a given n, $\mathbf{H}_n$ has i.i.d. $\mathcal{CN}(0,1)$ entries. We do allow, however, for correlation across n, assuming, for simplicity, that each scalar subchannel has the same correlation function, i.e., $\mathbb{E}\{(\mathbf{H}_n(i,j))^* \mathbf{H}_{n-m}(i,j)\} = r_{\mathbb{H}}(m)$, $(i=1,2,\ldots,\mathrm{M_R}, j=1,2,\ldots,\mathrm{M_T})$. Defining $\mathbf{H} = [\mathbf{H}_0 \mathbf{H}_1 \ldots \mathbf{H}_{N-1}]$, we therefore have

$$\mathbb{E}\{\mathrm{vec}(\mathbf{H})\,(\mathrm{vec}(\mathbf{H}))^H\} = \mathbf{R}_{\mathbb{H}}^T \otimes \mathbf{I}_{\mathrm{M_T M_R}} \tag{2}$$

where the covariance matrix $\mathbf{R}_{\mathbb{H}}(i,j) = r_{\mathbb{H}}(i-j)$ $(i,j=0,1,\ldots,N-1)$ follows from the channel's scattering function [10]. In the purely frequency-selective case, e.g., assuming an orthogonal frequency-division multiplexing (OFDM) system [11] with N tones and hence $\mathbf{H}_n = \sum_{l=0}^{L-1} \mathbf{H}(l) e^{-j\frac{2\pi}{N}ln}$, where the uncorrelated (across l) matrix-valued taps $\mathbf{H}(l)$ have i.i.d. $\mathcal{CN}(0,\sigma_l^2)$ entries, we obtain $r_{\mathbb{H}}(m) = \sum_{l=0}^{L-1} \sigma_l^2 e^{-j\frac{2\pi}{N}lm}$ $(m=0,1,\ldots,N-1)$. In the remainder of the paper, we use the definition $\rho := \mathrm{rank}(\mathbf{R}_{\mathbb{H}})$.

III. Diversity-multiplexing tradeoff

A. Preliminaries

Assuming perfect CSI in the receiver, the mutual information of the channel in (1) is given by

$$\mathrm{I}(\mathsf{SNR}) = \frac{1}{N}\sum_{n=0}^{N-1} \log\det\left(\mathbf{I}_{\mathrm{M_R}} + \frac{\mathsf{SNR}}{\mathrm{M_T}} \mathbf{H}_n \mathbf{C}_n \mathbf{H}_n^H\right) \tag{3}$$

where the transmit signal vectors are uncorrelated across n and satisfy $\mathbf{x}_n \sim \mathcal{CN}(\mathbf{0},\mathbf{C}_n)$ with power constraint $\mathrm{Tr}(\mathbf{C}_n) \leq \mathrm{M_T}$, $n=0,1,\ldots,N-1$. The DM-tradeoff realized by a family (w.r.t. SNR) of codes $\mathcal{C}_r$ with rate $R(\mathsf{SNR}) = r\log\mathsf{SNR}$, where $r \in [0,\mathrm{m}]$, is given by the function

$$d_{\mathcal{C}}(r) = -\lim_{\mathsf{SNR}\to\infty} \frac{\log P_e(r,\mathsf{SNR})}{\log\mathsf{SNR}}$$

where $P_e(r,\mathsf{SNR})$ is the error probability obtained through ML detection. At a given SNR, the corresponding codebook $\mathcal{C}_r(\mathsf{SNR})$ contains SNR^{Nr} codewords $\mathbf{X} = [\mathbf{x}_0\ \mathbf{x}_1\ \ldots\ \mathbf{x}_{N-1}]$. We say that such a family of codes $\mathcal{C}_r$ operates at multiplexing rate r. The optimal tradeoff curve $d^\star(r) = \sup_{\mathcal{C}_r} d_{\mathcal{C}}(r)$, where the supremum is taken over all families of codes satisfying $R(\mathsf{SNR}) = r\log\mathsf{SNR}$, quantifies the maximum achievable diversity gain as a function of r. Since the outage probability $P_{\mathcal{O}}(r,\mathsf{SNR})$ is a lower bound to the error probability [1], we have

$$d^\star(r) \leq d_{\mathcal{O}}(r) = -\lim_{\mathsf{SNR}\to\infty} \frac{\log P_{\mathcal{O}}(r,\mathsf{SNR})}{\log\mathsf{SNR}}.$$

Extending the arguments that lead to [1, Eq. (9)] to the case $N>1$, we can conclude that setting $\mathbf{C}_n = \mathbf{I}_{\mathrm{M_T}}$ $(n=0,1,\ldots,N-1)$ in (3) does not alter the exponential behavior of mutual information. Hence

$$P_{\mathcal{O}}(r,\mathsf{SNR}) \doteq \mathbb{P}\left(\frac{1}{N}\sum_{n=0}^{N-1} \log\det\left(\mathbf{I}_{\mathrm{M_R}} + \mathsf{SNR}\,\mathbf{H}_n\mathbf{H}_n^H\right) < r\log\mathsf{SNR}\right) \tag{4}$$

where we used the fact that the factor $1/\mathrm{M_T}$ in (3) can be neglected in the scale of interest. Let $\boldsymbol{\mu}(n) := [\mu_1(n)\ \mu_2(n)\ \ldots\ \mu_{\mathrm{m}}(n)]$ $(n=0,1,\ldots,N-1)$, with the singularity levels defined as

$$\mu_k(n) = -\frac{\log\lambda_k(\mathbf{H}_n\mathbf{H}_n^H)}{\log\mathsf{SNR}}, \quad k=1,2,\ldots,\mathrm{m}$$

and note that [1]

$$P_{\mathcal{O}}(r,\mathsf{SNR}) \doteq \mathbb{P}(\mathcal{O}(r)) \tag{5}$$

where

$$\mathcal{O}(r) = \left\{\boldsymbol{\mu}(n) \in \mathbb{R}_+^{\mathrm{m}}, n=0,1,\ldots,N-1 : \frac{1}{N}\sum_{n=0}^{N-1}\sum_{k=1}^{\mathrm{m}} [1-\mu_k(n)]^+ < r\right\} \tag{6}$$

and $\mathbb{R}_+^{\mathrm{m}}$ denotes the nonnegative orthant. Unlike the frequency-flat fading case treated in [1], characterizing $d_{\mathcal{O}}(r)$ for the selective-fading case seems analytically intractable with the main difficulty stemming from the fact that one has to deal with the sum of correlated (recall that the $\mathbf{H}_n$ are correlated across n) terms in (4). It turns out, however, that one can find lower and upper bounds on $\mathrm{I}(\mathsf{SNR})$ which are exponentially tight (and, hence, preserve the DM-tradeoff behavior) and analytically tractable. The next section formalizes this idea.

B. Jensen channel and Jensen outage event

We start by noting that applying Jensen's inequality yields

$$\mathrm{I}(\mathsf{SNR}) = \frac{1}{N}\sum_{n=0}^{N-1} \log\det\left(\mathbf{I}_{\mathrm{M_R}} + \frac{\mathsf{SNR}}{\mathrm{M_T}}\mathbf{H}_n\mathbf{H}_n^H\right) \leq \log\det\left(\mathbf{I}_{\mathrm{m}} + \frac{\mathsf{SNR}}{\mathrm{M_T}N}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H\right) := \mathrm{J}(\mathsf{SNR}) \tag{7}$$

where the "Jensen channel" is defined as

$$\boldsymbol{\mathcal{H}} = \begin{cases} [\mathbf{H}_0\ \mathbf{H}_1\ \ldots\ \mathbf{H}_{N-1}], & \text{if } \mathrm{M_R} \leq \mathrm{M_T}, \\ [\mathbf{H}_0^H\ \mathbf{H}_1^H\ \ldots\ \mathbf{H}_{N-1}^H], & \text{if } \mathrm{M_R} > \mathrm{M_T}. \end{cases}$$

In the following, we say that a Jensen outage event occurs if the Jensen channel $\boldsymbol{\mathcal{H}}$ is in outage w.r.t. the rate $R(\mathsf{SNR}) = r\log\mathsf{SNR}$, i.e., if $\mathrm{J}(\mathsf{SNR}) < R(\mathsf{SNR})$. The corresponding outage probability will be denoted as $P_{\mathcal{J}}(r,\mathsf{SNR})$ and clearly satisfies $P_{\mathcal{J}}(r,\mathsf{SNR}) \leq P_{\mathcal{O}}(r,\mathsf{SNR})$. The operational significance of a Jensen outage will be established at the end of this section. We shall first focus on characterizing the Jensen outage event analytically. Using (2), it is readily seen that $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}_w(\mathbf{R}_{\mathbb{H}}^{1/2} \otimes \mathbf{I}_{\mathrm{M}})$, where $\boldsymbol{\mathcal{H}}_w$ is an i.i.d. $\mathcal{CN}(0,1)$ matrix with the same dimensions as $\boldsymbol{\mathcal{H}}$. Noting that $\boldsymbol{\mathcal{H}}_w\mathbf{U} \sim \boldsymbol{\mathcal{H}}_w$ for $\mathbf{U}$ unitary and using the eigendecomposition $\mathbf{R}_{\mathbb{H}} \otimes \mathbf{I}_{\mathrm{M}} = \mathbf{U}(\boldsymbol{\Lambda} \otimes \mathbf{I}_{\mathrm{M}})\mathbf{U}^H$, where $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1(\mathbf{R}_{\mathbb{H}}), \lambda_2(\mathbf{R}_{\mathbb{H}}), \ldots, \lambda_\rho(\mathbf{R}_{\mathbb{H}}), 0, \ldots, 0\}$, it follows that

$$\begin{aligned}\mathrm{J}(\mathsf{SNR}) &= \log\det\left(\mathbf{I}_{\mathrm{m}} + \frac{\mathsf{SNR}}{\mathrm{M_T}N}\boldsymbol{\mathcal{H}}_w(\mathbf{R}_{\mathbb{H}} \otimes \mathbf{I}_{\mathrm{M}})\boldsymbol{\mathcal{H}}_w^H\right)\\ &\sim \log\det\left(\mathbf{I}_{\mathrm{m}} + \frac{\mathsf{SNR}}{\mathrm{M_T}N}\boldsymbol{\mathcal{H}}_w(\boldsymbol{\Lambda} \otimes \mathbf{I}_{\mathrm{M}})\,\boldsymbol{\mathcal{H}}_w^H\right).\end{aligned}$$

Next, observe that the following positive semidefinite ordering holds

$$\lambda_1(\mathbf{R}_{\mathbb{H}})\,\mathrm{diag}\{\mathbf{I}_{\rho\mathrm{M}}, \mathbf{0}\} \;\preceq\; \boldsymbol{\Lambda} \otimes \mathbf{I}_{\mathrm{M}} \;\preceq \lambda_\rho(\mathbf{R}_{\mathbb{H}})\,\mathrm{diag}\{\mathbf{I}_{\rho\mathrm{M}}, \mathbf{0}\}\,. \tag{8}$$

Since $f(\mathbf{A}) = \log\det(\mathbf{I} + \mathbf{A})$ is increasing over the cone of positive semidefinite matrices [12], we get the following bounds on the Jensen outage probability

$$\begin{aligned}&\mathbb{P}\left(\log\det\left(\mathbf{I}_{\mathrm{m}} + \lambda_\rho(\mathbf{R}_{\mathbb{H}})\frac{\mathsf{SNR}}{\mathrm{M_T}N}\overline{\boldsymbol{\mathcal{H}}}_w\overline{\boldsymbol{\mathcal{H}}}_w^H\right) < r\log\mathsf{SNR}\right)\\ &\le P_{\mathcal{J}}(r, \mathsf{SNR})\\ &\le \mathbb{P}\left(\log\det\left(\mathbf{I}_{\mathrm{m}} + \lambda_1(\mathbf{R}_{\mathbb{H}})\frac{\mathsf{SNR}}{\mathrm{M_T}N}\overline{\boldsymbol{\mathcal{H}}}_w\overline{\boldsymbol{\mathcal{H}}}_w^H\right) < r\log\mathsf{SNR}\right)\end{aligned} \tag{9}$$

where $\overline{\boldsymbol{\mathcal{H}}}_w = \boldsymbol{\mathcal{H}}_w([1:\mathrm{m}], [1:\rho\mathrm{M}])$. Taking the exponential limit (in SNR) in (9), it follows readily that

$$P_{\mathcal{J}}(r, \mathsf{SNR}) \doteq \mathbb{P}\left(\log\det\left(\mathbf{I}_{\mathrm{m}} + \mathsf{SNR}\,\overline{\boldsymbol{\mathcal{H}}}_w\overline{\boldsymbol{\mathcal{H}}}_w^H\right) < r\log\mathsf{SNR}\right). \tag{10}$$

For later use, we define $\boldsymbol{\alpha} := [\alpha_1\ \alpha_2\ \ldots\ \alpha_{\mathrm{m}}]$ with the singularity levels

$$\alpha_k = -\frac{\log\lambda_k(\overline{\boldsymbol{\mathcal{H}}}_w\overline{\boldsymbol{\mathcal{H}}}_w^H)}{\log\mathsf{SNR}}, \quad k = 1, 2, \ldots, \mathrm{m} \tag{11}$$

and note that $P_{\mathcal{J}}(r, \mathsf{SNR}) \doteq \mathbb{P}\left(\mathcal{J}(r)\right)$, where

$$\mathcal{J}(r) = \left\{\boldsymbol{\alpha} \in \mathbb{R}_+^{\mathrm{m}}\colon \alpha_1 \ge \alpha_2 \ge \ldots \ge \alpha_{\mathrm{m}}, \sum_{k=1}^{\mathrm{m}}[1 - \alpha_k]^+ < r\right\}.$$

It is now natural to define the Jensen outage curve as

$$d_{\mathcal{J}}(r) = -\lim_{\mathsf{SNR}\to\infty}\frac{\log P_{\mathcal{J}}(r, \mathsf{SNR})}{\log\mathsf{SNR}}.$$

Based on (10), we can conclude that $d_{\mathcal{J}}(r)$ is nothing but the DM-tradeoff curve of an effective MIMO channel with $\rho\mathrm{M}$ transmit and m receive antennas. We can therefore directly apply the results in [1] to infer that the Jensen outage curve is the piecewise linear function connecting the points $(r, d_{\mathcal{J}}(r))$ for $r = 0, 1, \ldots, \mathrm{m}$, with

$$d_{\mathcal{J}}(r) = (\rho\mathrm{M} - r)(\mathrm{m} - r). \tag{12}$$

Since, as already noted, $P_{\mathcal{J}}(r, \mathsf{SNR}) \le P_{\mathcal{O}}(r, \mathsf{SNR})$, we obtain

$$d_{\mathcal{C}}(r) \le d^\star(r) \le d_{\mathcal{O}}(r) \le d_{\mathcal{J}}(r), \quad r \in [0, \mathrm{m}], \tag{13}$$

for any family of codes $\mathcal{C}_r$. The optimal DM-tradeoff curve $d^\star(r)$ will be established in the next section by showing that codes satisfying $d_{\mathcal{C}}(r) = d_{\mathcal{J}}(r)$ do exist and hence $d^\star(r) = d_{\mathcal{J}}(r)$.

IV. Jensen-optimal code design criterion

The goal of this section is to derive a sufficient condition for a family of codes to achieve $d_{\mathcal{J}}(r)$, and hence, by virtue of (13), to be DM-tradeoff optimal.

A. Code design criterion

Theorem 1: Consider a family of codes $\mathcal{C}_r$ with block length $N \ge \rho\mathrm{M_T}$ that operates over the channel (1). If, for any codebook $\mathcal{C}_r(\mathsf{SNR}) \in \mathcal{C}_r$ and any two codewords $\mathbf{X}$, $\mathbf{X}' \in \mathcal{C}_r(\mathsf{SNR})$, the codeword difference matrix $\mathbf{E} = \mathbf{X} - \mathbf{X}'$ is such that

$$\mathrm{rank}\big(\mathbf{R}_{\mathbb{H}} \odot \mathbf{E}^H\mathbf{E}\big) = \rho\mathrm{M_T} \tag{14}$$

then the error probability (for ML decoding) satisfies

$$P_e(r, \mathsf{SNR}) \doteq \mathsf{SNR}^{-d_{\mathcal{J}}(r)}.$$

Proof: We start by deriving an upper bound on the average (w.r.t. the random channel) pairwise error probability (PEP). Assuming that $\mathbf{X}$ was transmitted, the probability of the ML decoder mistakenly deciding in favor of codeword $\mathbf{X}'$ can be upper-bounded in terms of the codeword difference vectors $\mathbf{e}_n = \mathbf{x}_n - \mathbf{x}'_n$ $(n = 0, 1, \ldots, N-1)$ as

$$\begin{aligned}\mathbb{P}\left(\mathbf{X} \to \mathbf{X}'\right) &\le \mathbb{E}_{\mathbf{H}}\left\{\exp\left(-\frac{\mathsf{SNR}}{4\mathrm{M_T}}\sum_{n=0}^{N-1}||\mathbf{H}_n\mathbf{e}_n||^2\right)\right\}\\ &= \mathbb{E}_{\mathbf{H}}\left\{\exp\left(-\frac{\mathsf{SNR}}{4\mathrm{M_T}}\mathrm{Tr}\left(\mathbf{H}_w\boldsymbol{\Upsilon}\mathbf{H}_w^H\right)\right)\right\}\end{aligned}$$

where

$$\boldsymbol{\Upsilon} = (\mathbf{R}_{\mathbb{H}}^{1/2} \otimes \mathbf{I}_{\mathrm{M_T}})\,\mathrm{diag}\big\{\mathbf{e}_n\mathbf{e}_n^H\big\}_{n=0}^{N-1}\,(\mathbf{R}_{\mathbb{H}}^{1/2} \otimes \mathbf{I}_{\mathrm{M_T}})$$

and $\mathbf{H}_w$ denotes an $\mathrm{M_R} \times \mathrm{M_T}N$ i.i.d. $\mathcal{CN}(0,1)$ matrix. Straightforward manipulations reveal that $\mathrm{rank}(\boldsymbol{\Upsilon}) = \mathrm{rank}\big(\mathbf{R}_{\mathbb{H}} \odot \mathbf{E}^H\mathbf{E}\big)$ so that the assumption (14) implies $\mathrm{rank}(\boldsymbol{\Upsilon}) = \rho\mathrm{M_T}$. With the eigendecomposition $\boldsymbol{\Upsilon} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, we have $\mathrm{Tr}\left(\mathbf{H}_w\boldsymbol{\Upsilon}\mathbf{H}_w^H\right) \sim \mathrm{Tr}\left(\mathbf{H}_w\boldsymbol{\Lambda}\mathbf{H}_w^H\right)$, and hence

$$\mathbb{P}\left(\mathbf{X} \to \mathbf{X}'\right) \le \mathbb{E}_{\mathbf{H}}\left\{\exp\left(-\frac{\mathsf{SNR}}{4\mathrm{M_T}}\mathrm{Tr}\left(\mathbf{H}_w\boldsymbol{\Lambda}\mathbf{H}_w^H\right)\right)\right\}.$$

Setting $\overline{\mathbf{H}}_w = \mathbf{H}_w([1{:}\mathrm{M_R}], [1{:}\rho\mathrm{M_T}])$ and denoting the smallest nonzero eigenvalue of $\boldsymbol{\Upsilon}$ as λ, we note that

$$\mathrm{Tr}\left(\mathbf{H}_w\boldsymbol{\Lambda}\mathbf{H}_w^H\right) \ge \lambda\;\mathrm{Tr}\left(\overline{\mathbf{H}}_w\overline{\mathbf{H}}_w^H\right) \tag{15}$$

and thus

$$\mathbb{P}\left(\mathbf{X} \to \mathbf{X}'\right) \le \mathbb{E}_{\overline{\mathbf{H}}_w}\left\{\exp\left(-\frac{\lambda\,\mathsf{SNR}}{4\mathrm{M_T}}\mathrm{Tr}\left(\overline{\mathbf{H}}_w\overline{\mathbf{H}}_w^H\right)\right)\right\}. \tag{16}$$

Next, note that

$$\begin{aligned}\mathrm{Tr}\left(\overline{\mathbf{H}}_w\overline{\mathbf{H}}_w^H\right) &= \mathrm{Tr}\left(\overline{\boldsymbol{\mathcal{H}}}_w\overline{\boldsymbol{\mathcal{H}}}_w^H\right)\\ &= \sum_{k=1}^{\mathrm{m}}\lambda_k(\overline{\boldsymbol{\mathcal{H}}}_w\overline{\boldsymbol{\mathcal{H}}}_w^H)\\ &= \sum_{k=1}^{\mathrm{m}}\mathsf{SNR}^{-\alpha_k}\end{aligned} \tag{17}$$

where (17) follows from (11). We can now write the PEP upper-bound in (16) in terms of the singularity levels α_k ($k = 1, 2, \ldots, \mathrm{m}$) characterizing the Jensen outage event:

$$\mathbb{P}\left(\mathbf{X} \to \mathbf{X}'\right) \leq \mathbb{E}_{\boldsymbol{\alpha}}\left\{\exp\left(-\frac{\lambda}{4\mathrm{M_T}}\sum_{k=1}^{\mathrm{m}}\mathsf{SNR}^{1-\alpha_k}\right)\right\}. \quad (18)$$

Next, consider a realization of the random vector $\boldsymbol{\alpha}$ and let $\mathcal{S} = \{k : \alpha_k \leq 1\}$. We have

$$\begin{aligned}\sum_{k=1}^{\mathrm{m}}\mathsf{SNR}^{1-\alpha_k} &\geq \sum_{k\in\mathcal{S}}\mathsf{SNR}^{1-\alpha_k}\\ &\overset{\text{(i)}}{\geq} |\mathcal{S}|\,\mathsf{SNR}^{\frac{1}{|\mathcal{S}|}\sum_{k\in\mathcal{S}}(1-\alpha_k)}\\ &\overset{\text{(ii)}}{=} |\mathcal{S}|\,\mathsf{SNR}^{\frac{1}{|\mathcal{S}|}\sum_{k=1}^{\mathrm{m}}[1-\alpha_k]^+} \quad (19)\end{aligned}$$

where (i) follows from the arithmetic-geometric mean inequality and (ii) follows from the definition of $\mathcal{S}$. Using (19) in (18), we get

$$\mathbb{P}\left(\mathbf{X} \to \mathbf{X}'\right) \leq \mathbb{E}_{\boldsymbol{\alpha}}\left\{\exp\left(-\frac{\lambda\,|\mathcal{S}|}{4\mathrm{M_T}}\,\mathsf{SNR}^{\frac{1}{|\mathcal{S}|}\sum_{k=1}^{\mathrm{m}}[1-\alpha_k]^+}\right)\right\}. \quad (20)$$

The dependence of the PEP upper bound (20) on the singularity levels characterizing the Jensen outage event suggests to split up the overall error probability according to

$$\begin{aligned}P_e(r,\mathsf{SNR}) &= \mathbb{P}\left(\text{error}, \boldsymbol{\alpha}\in\mathcal{J}(r)\right) + \mathbb{P}\left(\text{error}, \boldsymbol{\alpha}\notin\mathcal{J}(r)\right)\\ &= \mathbb{P}\left(\boldsymbol{\alpha}\in\mathcal{J}(r)\right)\mathbb{P}\left(\text{error}|\boldsymbol{\alpha}\in\mathcal{J}(r)\right)\\ &\quad + \mathbb{P}\left(\boldsymbol{\alpha}\notin\mathcal{J}(r)\right)\mathbb{P}\left(\text{error}|\boldsymbol{\alpha}\notin\mathcal{J}(r)\right)\\ &\leq \mathbb{P}\left(\boldsymbol{\alpha}\in\mathcal{J}(r)\right)\\ &\quad + \mathbb{P}\left(\boldsymbol{\alpha}\notin\mathcal{J}(r)\right)\mathbb{P}\left(\text{error}|\boldsymbol{\alpha}\notin\mathcal{J}(r)\right). \quad (21)\end{aligned}$$

For any $\boldsymbol{\alpha}\notin\mathcal{J}(r)$, we have $\sum_{k=1}^{\mathrm{m}}[1-\alpha_k]^+ \geq r$ and $|\mathcal{S}| \geq 1$, which upon noting that $|\mathcal{C}_r(\mathsf{SNR})| = \mathsf{SNR}^{Nr}$, yields the following union bound based on the PEP in (20)

$$\mathbb{P}\left(\text{error}|\boldsymbol{\alpha}\notin\mathcal{J}(r)\right) \leq \mathsf{SNR}^{Nr}\exp\left(-\frac{\lambda}{4\mathrm{M_T}}\,\mathsf{SNR}^{r/\mathrm{m}}\right)$$

where we used $|\mathcal{S}| \leq \mathrm{m}$. Hence, for any $r > 0$, $\mathbb{P}\left(\text{error}|\boldsymbol{\alpha}\notin\mathcal{J}(r)\right)$ decays exponentially in SNR and we have

$$\begin{aligned}\mathbb{P}\left(\text{error}, \boldsymbol{\alpha}\notin\mathcal{J}(r)\right) &= \underbrace{\mathbb{P}\left(\boldsymbol{\alpha}\notin\mathcal{J}(r)\right)}_{\leq 1}\mathbb{P}\left(\text{error}|\boldsymbol{\alpha}\notin\mathcal{J}(r)\right)\\ &\leq \mathsf{SNR}^{Nr}\exp\left(-\frac{\lambda}{4\mathrm{M_T}}\,\mathsf{SNR}^{r/\mathrm{m}}\right). \quad (22)\end{aligned}$$

Consequently, noting that $\mathbb{P}\left(\boldsymbol{\alpha}\in\mathcal{J}(r)\right) \doteq P_{\mathcal{J}}(r,\mathsf{SNR})$ and using (22) in (21), we obtain

$$P_e(r,\mathsf{SNR}) \,\dot{\leq}\, P_{\mathcal{J}}(r,\mathsf{SNR}).$$

Since $P_{\mathcal{J}}(r,\mathsf{SNR}) \leq P_{\mathcal{O}}(r,\mathsf{SNR})$, it follows trivially that $P_{\mathcal{J}}(r,\mathsf{SNR}) \,\dot{\leq}\, P_{\mathcal{O}}(r,\mathsf{SNR})$. In addition, for a specific family of codes $\mathcal{C}_r$, we have $P_{\mathcal{O}}(r,\mathsf{SNR}) \leq P_e(r,\mathsf{SNR})$ and hence $P_{\mathcal{O}}(r,\mathsf{SNR}) \,\dot{\leq}\, P_e(r,\mathsf{SNR})$. Putting the pieces together, we finally obtain

$$P_{\mathcal{O}}(r,\mathsf{SNR}) \,\dot{\leq}\, P_e(r,\mathsf{SNR}) \,\dot{\leq}\, P_{\mathcal{J}}(r,\mathsf{SNR}) \,\dot{\leq}\, P_{\mathcal{O}}(r,\mathsf{SNR})$$

which implies

$$P_e(r,\mathsf{SNR}) \doteq P_{\mathcal{J}}(r,\mathsf{SNR})$$

and hence (by definition of $d_{\mathcal{J}}(r)$)

$$P_e(r,\mathsf{SNR}) \doteq \mathsf{SNR}^{-d_{\mathcal{J}}(r)}.$$

■

As a direct consequence of Theorem 1, a family of codes that satisfies (14) for all codeword difference matrices in any codebook $\mathcal{C}_r(\mathsf{SNR}) \in \mathcal{C}_r$ realizes a DM-tradeoff curve $d_{\mathcal{C}}(r) = d_{\mathcal{J}}(r)$ and hence, by (13)

$$d_{\mathcal{J}}(r) \leq d^{\star}(r) \leq d_{\mathcal{J}}(r)$$

which implies

$$d^{\star}(r) = d_{\mathcal{J}}(r). \quad (23)$$

The optimal DM-tradeoff curve for selective-fading MIMO channels is therefore given by the DM-tradeoff curve of the associated Jensen channel. Put differently, Theorem 1 shows that, even though $\mathcal{J}(r) \subseteq \mathcal{O}(r)$ by definition, we still have

$$\mathbb{P}\left(\mathcal{J}(r)\right) \doteq \mathbb{P}\left(\mathcal{O}(r)\right)$$

which essentially says that the "original" channel has the same high-SNR outage behavior as its associated Jensen channel.

The code design criterion in Theorem 1 provides a sufficient condition for achieving the DM-tradeoff curve. Interestingly, the classical rank criterion [13]–[18], aimed at maximizing the diversity gain for $r = 0$, can be shown [19] to be equivalent to the criterion in Theorem 1. We emphasize, however, that optimality w.r.t. the DM-tradeoff at multiplexing rate r requires that (14) is satisfied for all codeword difference matrices in any codebook $\mathcal{C}_r(\mathsf{SNR}) \in \mathcal{C}_r$, in particular also for $\mathsf{SNR} \to \infty$. We next state a sufficient condition for DM-tradeoff optimality which makes this aspect explicit and establishes a connection to the approximately universal code design criterion in [7].

Corollary 1: A family of codes $\mathcal{C}_r$ of block length $N \geq \rho\mathrm{M_T}$ is DM-tradeoff optimal if there exists an $\epsilon > 0$ such that

$$\lambda^{\mathrm{m}}(\mathsf{SNR}) \,\dot{\geq}\, \mathsf{SNR}^{-(r-\epsilon)} \quad (24)$$

where

$$\lambda(\mathsf{SNR}) = \min_{\substack{\mathbf{E}=\mathbf{X}-\mathbf{X}'\\ \mathbf{X},\mathbf{X}'\in\mathcal{C}_r(\mathsf{SNR})}} \lambda_{N-\rho\mathrm{M_T}+1}(\mathbf{R}_{\mathbb{H}} \odot \mathbf{E}^H\mathbf{E}).$$

Proof: Using (24) in (22), we obtain

$$\mathbb{P}\left(\text{error}, \boldsymbol{\alpha}\notin\mathcal{J}(r)\right) \leq \mathsf{SNR}^{Nr}\exp\left(-\frac{\mathsf{SNR}^{\epsilon/\mathrm{m}}}{4\mathrm{M_T}}\right)$$

which, following the same logic as in the proof of Theorem 1, implies that $P_e(r,\mathsf{SNR}) \doteq \mathsf{SNR}^{-d_{\mathcal{J}}(r)}$. ■

Note that the quantity $\lambda^{\mathrm{m}}(\mathsf{SNR})$ is trivially a lower bound on the product of the m smallest nonzero eigenvalues of any codeword difference matrix in the codebook $\mathcal{C}_r(\mathsf{SNR})$. Consequently, in the case of non-selective fading, where $\mathbf{R}_{\mathbb{H}} \odot \mathbf{E}^H\mathbf{E} = \mathbf{E}^H\mathbf{E}$, any family of codes $\mathcal{C}_r$ satisfying

(24) will also be approximately universal in the sense of [7, Th. 3.1]. Moreover, if $\lambda(\mathsf{SNR})$ remains strictly positive as $\mathsf{SNR} \to \infty$, $\mathcal{C}_r$ fulfills the non-vanishing determinant criterion [2], [3] and will, by (22) and the same arguments as in the proof of Theorem 1, be DM-tradeoff optimal.

B. Application to the frequency-selective case

As an example, we shall next specialize our results to frequency-selective fading MIMO channels, recovering the results reported previously in [8], [9]. For the sake of simplicity of exposition, we shall employ a cyclic signal model, as obtained in an OFDM system for example. The channel's transfer function is given by

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

where the $\mathbf{H}(l)$ have i.i.d. $\mathcal{CN}\big(0, \sigma_l^2\big)$ entries and satisfy

$$\mathbb{E}\Big\{\mathrm{vec}(\mathbf{H}(l))\,\mathrm{vec}(\mathbf{H}(l'))^H\Big\} = \sigma_l^2\, \delta(l-l')\, \mathbf{I}_{\mathrm{M_T M_R}}.$$

With $\mathbf{H}_n = \mathbf{H}(e^{j2\pi\frac{n}{N}})$, $n = 0, 1 \ldots, N-1$, the channel's covariance matrix follows as

$$\mathbf{R}_{\mathbb{H}} = \mathbf{F}\,\mathrm{diag}\big\{\sigma_0^2, \sigma_1^2, \ldots, \sigma_{L-1}^2, 0, \ldots, 0\big\}\,\mathbf{F}^H$$

where $\mathbf{F}$ is the $N \times N$ FFT matrix. Since $\mathrm{rank}(\mathbf{R}_{\mathbb{H}}) = L$, inserting $\rho = L$ into (12) and using (23) yields the optimal DM-tradeoff curve as the piecewise linear function connecting the points $(r, d^\star(r))$ for $r = 0, 1, \ldots, \mathrm{m}$, with

$$d^\star(r) = (L\mathrm{M} - r)(\mathrm{m} - r). \tag{25}$$

This is the optimal DM-tradeoff curve for frequency-selective fading MIMO channels reported previously in [9]. Specializing (25) to the case $\mathrm{M_T} = \mathrm{M_R} = 1$ and noting that $d^\star(r) = (L - r)(1-r) = L(1-r)$ for $r = \{0, 1\}$, yields the results reported in [8]. We note that the proof techniques employed in [8], [9] are different from the approach taken in this paper and seem to be tailored to the frequency-selective case. In addition, our approach is not limited to large code lengths as (14) can be guaranteed for any $N \geq L\mathrm{M_T}$.

V. Conclusions

Analyzing the high-SNR outage behavior of the Jensen channel instead of the original channel was found to be an effective tool to establish the DM-tradeoff in selective-fading MIMO channels. We showed that satisfying extensions (to the selective-fading MIMO case) of the approximately universal code design criterion [7] and the non-vanishing determinant criterion [2], [3] results in DM-tradeoff optimal codes. Finally, we note that the concepts introduced in this paper can be extended to multiple-access selective-fading MIMO channels and to the analysis of the DM-tradeoff properties of specific (suboptimal) receivers.

References

[1] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] J.-C. Belfiore and G. Rekaya, "Quaternionic lattices for space-time coding," in *Proc. IEEE Inf. Theory Workshop*, Paris, France, Mar./Apr. 2003, pp. 267–270.

[3] H. Yao and G. W. Wornell, "Achieving the full MIMO diversity-multiplexing frontier with rotation based space-time codes," in *Proc. Allerton Conf. on Commun., Control and Computing*, Monticello, IL, Oct. 2003, pp. 400–409.

[4] P. Dayal and M. K. Varanasi, "An optimal two transmit antenna space-time code and its stacked extensions," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4348–4355, Dec. 2005.

[5] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden code: A 2x2 full rate space-time code with nonvanishing determinants," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[6] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieves the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 968–985, Sept. 2004.

[7] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, July 2006.

[8] L. Grokop and D. N. C. Tse, "Diversity/multiplexing tradeoff in ISI channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Chicago, USA, June/July 2004, p. 96.

[9] A. Medles and D. T. M. Slock, "Optimal diversity vs. multiplexing tradeoff for frequency selective MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Adelaide, Australia, Sept. 2005, pp. 1813–1817.

[10] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. COM-11, pp. 360–393, 1963.

[11] A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, Denver, CO, Apr. 1980, pp. 964–967.

[12] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, United Kingdom: Cambridge University Press, 2004.

[13] 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, no. 2, pp. 744–765, Mar. 1998.

[14] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Comm.*, vol. 47, pp. 199–207, Feb. 1999.

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

[16] H. Bölcskei, R. Koetter, and S. Mallik, "Coding and modulation for underspread fading channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Lausanne, Switzerland, June/July 2002, p. 358.

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

[18] X. Ma, G. Leus, and G. B. Giannakis, "Space-time-Doppler block coding for correlated time-selective fading channels," *IEEE Trans. Sig. Proc.*, vol. 53, no. 6, pp. 2167–2181, June 2005.

[19] P. Coronel and H. Bölcskei, "Diversity-multiplexing tradeoff in selective-fading MIMO channels," 2007, in preparation.