On the Capacity of Noncoherent Wideband MIMO-OFDM Systems

Moritz Borgmann and Helmut Bölcskei
Communication Technology Laboratory
Swiss Federal Institute of Technology (ETH) Zurich, CH-8092 Zurich, Switzerland
Email: {moriborg|boelcskei}@nari.ee.ethz.ch

***Abstract*—We study the capacity behavior of full-band MIMO-OFDM systems in the absence of channel state information both at the transmitter and the receiver. Based on capacity lower and upper bounds, we quantify the transmission rate penalty due to channel uncertainty as a function of the number of transmit antennas, the number of resolvable taps in the channel, the power delay profile, and the bandwidth. Our analysis reveals that for a given bandwidth and transmit power there is an optimum, capacity-maximizing, number of transmit antennas. Numerical results show that using a large number of transmit antennas in systems employing bandwidths of several GHz (such as in ultrawideband systems) is detrimental from a capacity point of view. Finally, we evaluate the capacity performance of space-frequency unitary codebooks recently introduced in [13].**

I. Introduction

Acquiring reliable channel state information (CSI) at the receiver of fast-fading wideband multiple-input multiple-output (MIMO) systems poses significant challenges due to the large number of channel parameters involved. It is therefore important to understand the fundamental performance limits of wideband MIMO systems in the absence of CSI at both the transmitter and the receiver (termed the *noncoherent* case).

The information-theoretic performance limits of noncoherent MIMO systems in the flat-fading case have been investigated in [1]–[3]. In [4], the sublinear (in SNR) behavior of the low-SNR MIMO capacity in the flat-fading noncoherent case has been analyzed. In this paper, we are interested in the characterization of the capacity of noncoherent MIMO–orthogonal frequency-division multiplexing (OFDM) systems in frequency-selective block-fading channels. Most practical OFDM systems are full band [5], i.e., the transmit signal occupies all time-frequency slots and hence signaling is not "peaky" in the sense of [6]–[10]. We focus on full-band MIMO-OFDM systems throughout and characterize their capacity behavior over the whole range of low bandwidth (wide enough to induce frequency-selective fading) to the infinite bandwidth limit. Before detailing our contributions, we note that the setup in [11], which considers the capacity of noncoherent time-selective MIMO fading channels, is structurally similar to ours. Specifically, the lower bound in Section III-B.1 of this paper parallels [11, eq. (53)].

A. Contributions

Our main contributions can be summarized as follows:

- Using a well-known bounding technique, first proposed in [12], we provide simple lower and upper bounds on the capacity of noncoherent MIMO-OFDM systems in frequency-selective Rayleigh fading. In addition, we derive an upper bound on the achievable rate of space-frequency unitary codebooks introduced in [13].
- Based on our capacity lower bound, we show that with increasing bandwidth, transmit antennas should be gradually switched off. For large enough bandwidths, using only one transmit antenna is found to be the capacity (lower bound)-maximizing strategy.
- We study the impact of the channel's power delay profile (PDP) on the critical bandwidth where "overspreading" sets in and transmit antennas should be switched off.

This work was supported by Nokia Research Center (NRC) Helsinki and by ETH Grant TH-6 02-2.

B. Notation

The superscripts ^{T,H} stand for transpose and conjugate transpose, respectively. $X_{i,j}$ represents the element in the ith row and jth column of the matrix $\boldsymbol{\mathcal{X}}$. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. The notation $\mathrm{diag}_{i=0}^{N-1}\{X_i\}$ stands for the $N \times N$ diagonal matrix with the entries X_i ($i = 0, 1, \ldots, N-1$) on its main diagonal. $\boldsymbol{\mathcal{X}} \otimes \boldsymbol{\mathcal{Y}}$ denotes the Kronecker product of the matrices $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$, $\mathrm{tr}(\boldsymbol{\mathcal{X}})$ stands for the trace of $\boldsymbol{\mathcal{X}}$, $\lambda_i(\boldsymbol{\mathcal{X}})$ ($i = 0, 1, \ldots, N-1$) are the eigenvalues of the $N \times N$ matrix $\boldsymbol{\mathcal{X}}$ sorted in nonincreasing order, and $\mathrm{vec}(\boldsymbol{\mathcal{X}}) = \left[\mathbf{X}_0^T \ \mathbf{X}_1^T \ \cdots \ \mathbf{X}_{N-1}^T\right]^T$, where $\mathbf{X}_j$ is the jth column of $\boldsymbol{\mathcal{X}}$.

$\mathcal{E}\{\cdot\}$ stands for the expectation operator. The covariance matrix of a random vector $\mathbf{X}$ is denoted by $\mathfrak{C}_{\mathbf{X}}$. A circularly symmetric zero-mean complex Gaussian random variable (RV) is a RV $Z = X + jY \sim \mathcal{CN}(0, \sigma^2)$, where the real-valued RVs X and Y are i.i.d. $\mathcal{N}(0, \sigma^2/2)$. The discrete Kronecker delta function $\delta_{i,j}$ is 1 if $i = j$ and 0 else. Logarithms are to the base e. The notation $u(t) = \mathcal{O}(v(t))$ denotes that $|u(t)/v(t)|$ remains bounded as $t \to \infty$.

II. System and Channel Model

In this section, we describe our channel and signal models.

A. Wideband MIMO Channel Model

In the following, M_T and M_R denote the number of transmit and receive antennas, respectively. We assume that the channel impulse response consists of L matrix-valued taps $\boldsymbol{\mathcal{H}}_l$ of dimension $M_R \times M_T$, $l = 0, 1, \ldots, L-1$, with the matrix-valued transfer function given by

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

Note that in general there will be a continuum of delays. The channel model (1) is derived from the assumption of having at

most L resolvable paths, where $L = \lfloor W\tau \rfloor + 1$ with W and τ denoting the signal bandwidth and delay spread, respectively.

We employ a *block fading* channel model, i.e., we assume that the channel impulse response remains constant for the duration of the *coherence time* T_{c} and then changes independently to a new realization. Coding is performed over an infinite number of independent channel uses.

We restrict ourselves to purely Rayleigh fading channels, so that the elements of $\boldsymbol{\mathcal{H}}_l$ ($l = 0, 1, \ldots, L-1$) are i.i.d. circularly symmetric zero-mean complex Gaussian RVs. We do allow, however, for correlations across taps so that

$$\mathfrak{C}_{\mathbf{H}} = \mathcal{E}_{\mathbf{H}}\{\mathbf{H}\mathbf{H}^H\} = \mathbf{I}_{M_R} \otimes \boldsymbol{\Sigma} \otimes \mathbf{I}_{M_T}$$

with $\mathbf{H} = \mathrm{vec}(\overline{\boldsymbol{\mathcal{H}}})$ and

$$\overline{\boldsymbol{\mathcal{H}}} = \begin{bmatrix} \boldsymbol{\mathcal{H}}_0 & \boldsymbol{\mathcal{H}}_1 & \cdots & \boldsymbol{\mathcal{H}}_{L-1} \end{bmatrix}^T.$$

The $L \times L$ covariance matrix $\boldsymbol{\Sigma}$ describes the intertap correlation; the eigenvalues of $\boldsymbol{\Sigma}$ are denoted as σ_l^2 ($l = 0, 1, \ldots, L-1$). Note that we assume the intertap correlation to be the same for all transmit and receive antennas. For uncorrelated scattering (US) channels, we have $\boldsymbol{\Sigma} = \mathrm{diag}_{l=0}^{L-1}\{\sigma_l^2\}$. Throughout the paper, we employ the normalization $\mathrm{tr}(\boldsymbol{\Sigma}) = 1$.

B. Space-Frequency Coded MIMO-OFDM Systems

In an OFDM-based MIMO system, the individual transmit signals corresponding to the M_T transmit antennas are OFDM-modulated before transmission. The total system bandwidth W is spanned by N *tones* with a *tone spacing* of $F = W/N$. The OFDM modulator applies an N-point IFFT to N consecutive (frequency domain) data symbols and then prepends the cyclic prefix (CP) of length $L_{\mathrm{CP}} \geq L$ to the parallel-to-serial-converted (time domain) OFDM symbol. The receiver discards the CP and then applies an N-point FFT to each of the M_R received signals. The CP serves as a guard interval and turns the linear convolution describing the action of the channel into a cyclic convolution. Throughout the paper, we assume that an OFDM symbol spans one block of length T_{c} and that $N = WT_{\mathrm{c}}$; thus, the loss in spectral efficiency due to the presence of a CP is neglected.

Denoting the M_T data symbols transmitted on the kth tone as $\mathbf{X}_k = \begin{bmatrix} X_{k,0} & X_{k,1} & \cdots & X_{k,M_T-1} \end{bmatrix}^T$ ($k = 0, 1, \ldots, N-1$), the M_R-dimensional received signal vectors are given by

$$\mathbf{Y}_k = \sqrt{\rho}\, \boldsymbol{\mathcal{H}}\big(e^{j2\pi \frac{k}{N}}\big)\mathbf{X}_k + \mathbf{W}_k, \qquad k = 0, 1, \ldots, N-1$$

where ρ is an energy normalization factor, and $\mathbf{W}_k$ is complex-valued zero-mean circularly symmetric Gaussian noise satisfying $\mathcal{E}\{\mathbf{W}_k \mathbf{W}_{k'}^H\} = \mathbf{I}_{M_R}\delta_{k,k'}$. The input-output relation for an entire OFDM symbol can now be written as $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{S}} + \boldsymbol{\mathcal{W}}$, where

$$\boldsymbol{\mathcal{S}} = \sqrt{\rho}\Big[\boldsymbol{\mathcal{H}}\big(e^{j2\pi \frac{0}{N}}\big)\mathbf{X}_0 \quad \boldsymbol{\mathcal{H}}\big(e^{j2\pi \frac{1}{N}}\big)\mathbf{X}_1 \quad \cdots \quad \cdots \quad \boldsymbol{\mathcal{H}}\big(e^{j2\pi \frac{N-1}{N}}\big)\mathbf{X}_{N-1}\Big]^T \tag{2}$$

and $\boldsymbol{\mathcal{Y}} = \begin{bmatrix} \mathbf{Y}_0 & \mathbf{Y}_1 & \cdots & \mathbf{Y}_{N-1} \end{bmatrix}^T$ is the $N \times M_R$ matrix of received signals, and $\boldsymbol{\mathcal{W}} = \begin{bmatrix} \mathbf{W}_0 & \mathbf{W}_1 & \cdots & \mathbf{W}_{N-1} \end{bmatrix}^T$ is the $N \times M_R$ additive noise matrix. Defining the $N \times M_T$ matrix $\boldsymbol{\mathcal{X}} = \begin{bmatrix} \mathbf{X}_0 & \mathbf{X}_1 & \cdots & \mathbf{X}_{N-1} \end{bmatrix}^T$, we can rewrite (2) in terms of the matrix-valued channel taps $\boldsymbol{\mathcal{H}}_l$ ($l = 0, 1, \ldots, L-1$) as

$$\boldsymbol{\mathcal{S}} = \sqrt{\rho} \sum_{l=0}^{L-1} \boldsymbol{\Delta}^l\, \boldsymbol{\mathcal{X}}\, \boldsymbol{\mathcal{H}}_l^T = \sqrt{\rho}\, \boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\, \overline{\boldsymbol{\mathcal{H}}} \tag{3}$$

with the diagonal matrix $\boldsymbol{\Delta} = \mathrm{diag}_{k=0}^{N-1}\{e^{-j2\pi k/N}\}$ and the stacked $N \times M_T L$ matrix

$$\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}}) = \begin{bmatrix} \boldsymbol{\mathcal{X}} & \boldsymbol{\Delta}\boldsymbol{\mathcal{X}} & \cdots & \boldsymbol{\Delta}^{L-1}\boldsymbol{\mathcal{X}} \end{bmatrix}.$$

In the remainder of the paper, we shall use the input-output relation $\mathbf{Y} = \mathbf{S} + \mathbf{W}$ with $\mathbf{Y} = \mathrm{vec}(\boldsymbol{\mathcal{Y}})$, $\mathbf{W} = \mathrm{vec}(\boldsymbol{\mathcal{W}})$, and $\mathbf{S} = \mathrm{vec}(\boldsymbol{\mathcal{S}})$. We shall also need $\mathbf{X} = \mathrm{vec}(\boldsymbol{\mathcal{X}})$. With (3), we can write $\mathbf{S} = \sqrt{\rho}\big(\mathbf{I}_{M_R} \otimes \boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\big)\mathbf{H}$.

Throughout the paper, we assume uniform power allocation across frequency and transmit antennas, so that

$$\mathcal{E}\{|X_{k,m}|^2\} = \frac{1}{M_T} \tag{4}$$

for $k = 0, 1, \ldots, N-1$, $m = 0, 1, \ldots, M_T - 1$. Unless otherwise stated, we consider *i.i.d. Gaussian codebooks*, i.e., the entries of $\boldsymbol{\mathcal{X}}$ are assumed i.i.d. $\mathcal{CN}(0, 1/M_T)$. We shall frequently make use of the fact that conditioned on $\mathbf{X}$, the vector $\mathbf{S}$ is jointly complex Gaussian with covariance matrix

$$\mathfrak{C}_{\mathbf{S}|\mathbf{X}} = \rho\, \mathbf{I}_{M_R} \otimes \big(\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\, \boldsymbol{\Gamma}\, \boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\big) \tag{5}$$

where $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} \otimes \mathbf{I}_{M_T}$. Hence, the signal power per receive antenna is given by $P = (\rho/T_{\mathrm{c}})\, \mathcal{E}_{\boldsymbol{\mathcal{X}}}\big\{\mathrm{tr}\big(\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\, \boldsymbol{\Gamma}\, \boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\big)\big\} = W\rho$, and the receive SNR is $\mathrm{SNR} = P/W = \rho$. Throughout the paper, we assume that P is constant, so that the receive SNR approaches zero as $W \to \infty$.

III. Bounds on Capacity

Before stating our main results, we start with a brief summary of capacity results for the perfect receive CSI (*coherent*) frequency-selective MIMO-OFDM Rayleigh-fading channel.

A. Perfect Receive CSI MIMO-OFDM Capacity

The (ergodic) capacity of the coherent MIMO-OFDM channel is given by

$$C_{\mathrm{coh}} = \frac{1}{T_{\mathrm{c}}} \sup_{p(\mathbf{X})} I(\mathbf{Y}; \mathbf{X} \mid \mathbf{H})$$

with the supremum taken over all input distributions $p(\mathbf{X})$ that satisfy the power constraint (4). It is well known that i.i.d. Gaussian codebooks achieve the capacity of the coherent MIMO-OFDM channel [14], [15]. For US channels, we have

$$C_{\mathrm{coh}} = W\, \mathcal{E}_{\boldsymbol{\mathcal{H}}_{\mathrm{w}}}\left\{ \log\det\Big(\mathbf{I}_{M_R} + \frac{P}{M_T W}\boldsymbol{\mathcal{H}}_{\mathrm{w}}\boldsymbol{\mathcal{H}}_{\mathrm{w}}^H\Big)\right\} \tag{6}$$

where $\boldsymbol{\mathcal{H}}_{\mathrm{w}}$ is an $M_R \times M_T$ matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries. In the high-SNR regime, we have [16]

$$C_{\mathrm{coh}} = W \min(M_T, M_R)\, \log\left(\frac{P}{M_T W}\right) + \mathcal{O}(1).$$

In the large-bandwidth limit, the capacity converges to the capacity of M_R parallel infinite-bandwidth AWGN channels:

$$\lim_{W \to \infty} C_{\mathrm{coh}} = M_R P.$$

Finally, we note that in the case of nonzero intertap correlation, the capacity of the coherent MIMO-OFDM channel is given by (6) with an SNR penalty that depends on the covariance matrix $\mathbf{\Sigma}$.

B. Noncoherent MIMO-OFDM Capacity

The capacity of the noncoherent MIMO-OFDM channel is given by

$$C = \frac{1}{T_{\mathrm{c}}} \sup_{p(\mathbf{X})} I(\mathbf{Y};\mathbf{X}). \tag{7}$$

Following [12], we decompose $I(\mathbf{Y};\mathbf{X})$ using the chain rule:

$$I(\mathbf{Y};\mathbf{X}) = I(\mathbf{Y};\mathbf{X},\mathbf{H}) - I(\mathbf{Y};\mathbf{H}\,|\,\mathbf{X}). \tag{8}$$

1) Lower and Upper Bounds on Capacity: In order to obtain a lower bound on $I(\mathbf{Y};\mathbf{X})$, we follow [17] and start by noting that $I(\mathbf{Y};\mathbf{X},\mathbf{H}) \geq I(\mathbf{Y};\mathbf{X}\,|\,\mathbf{H})$.

The second term on the right-hand side of (8) can be evaluated by noting that since the channel $\mathbf{H}$ is Gaussian, $\mathbf{Y}$ conditioned on $\mathbf{X}$ is Gaussian, so that

$$\begin{aligned} &I(\mathbf{Y};\mathbf{H}\,|\,\mathbf{X}) \\ &\quad = M_R\,\mathcal{E}_{\boldsymbol{\mathcal{X}}}\Big\{\log\det\Big(\mathbf{I}_N + \rho\,\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\,\mathbf{\Gamma}\,\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\Big)\Big\} &(9)\\ &\quad \leq M_R\log\det\Big(\mathbf{I}_{M_TL} + \rho\;\mathcal{E}_{\boldsymbol{\mathcal{X}}}\Big\{\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\Big\}\,\mathbf{\Gamma}\Big) &(10)\end{aligned}$$

where (10) follows from Jensen's inequality. For i.i.d. symbols, $\mathcal{E}_{\boldsymbol{\mathcal{X}}}\{\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\} = (N/M_T)\mathbf{I}_{M_TL}$, which yields

$$I(\mathbf{Y};\mathbf{H}\,|\,\mathbf{X}) \leq M_TM_R\sum_{l=0}^{L-1}\log\Big(1+\rho\,\frac{N}{M_T}\sigma_l^2\Big). \tag{11}$$

Since any specific input distribution provides a lower bound on C, the i.i.d. Gaussian input distribution results in the general bound on capacity

$$C \geq C_{\mathrm{coh}} - \frac{M_TM_R}{T_{\mathrm{c}}}\sum_{l=0}^{L-1}\log\Big(1+\frac{PT_{\mathrm{c}}}{M_T}\sigma_l^2\Big). \tag{12}$$

Thus, we can see that the capacity of the noncoherent MIMO-OFDM channel is lower-bounded by the capacity of the corresponding coherent channel up to a "penalty term", which results from channel uncertainty.

A trivial upper bound on the capacity of the noncoherent MIMO-OFDM channel is the capacity of the same channel with perfect CSI at the receiver, i.e.,

$$C \leq C_{\mathrm{coh}}. \tag{13}$$

We will see that this bound is useful only at high SNR.

2) Fourthegy Upper Bound: The following upper bound on $I(\mathbf{Y};\mathbf{X})$ is analogous to the upper bound for the single-input single-output (SISO) case derived in [9, Appendix A]. We start by noting that $I(\mathbf{Y};\mathbf{X},\mathbf{H})$ is the mutual information of an AWGN channel with input signal $\mathbf{S} = \sqrt{\rho}\,\big(\mathbf{I}_{M_R}\otimes\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\big)\mathbf{H}$, where the transmitter has complete control over $\mathbf{S}$. Consequently, $I(\mathbf{Y};\mathbf{X},\mathbf{H})$ is maximized by a jointly complex Gaussian $\mathbf{S}$ with covariance matrix $\mathfrak{C}_{\mathbf{S}} = \rho\,\mathbf{I}_{M_R}\otimes\mathcal{E}_{\boldsymbol{\mathcal{X}}}\{\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\,\mathbf{\Gamma}\,\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\}$, so that

$$\begin{aligned} &I(\mathbf{Y};\mathbf{X},\mathbf{H}) \\ &\qquad \leq M_R\log\det\Big(\mathbf{I}_N + \rho\;\mathcal{E}_{\boldsymbol{\mathcal{X}}}\Big\{\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\,\mathbf{\Gamma}\,\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\Big\}\Big). &(14)\end{aligned}$$

Applying Hadamard's inequality and using $\log(1+x)\leq x$, we can further upper-bound the expression in (14), to obtain

$$\begin{aligned} I(\mathbf{Y};\mathbf{X},\mathbf{H}) &\leq M_RN\log(1+\rho\,) &(15)\\ &\leq M_RN\rho\,. &(16)\end{aligned}$$

Denoting the M_TL eigenvalues of $\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}})\,\mathbf{\Gamma}\,\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})$ by $\mu_r(\boldsymbol{\mathcal{X}})$, $r = 0,1,\ldots,M_TL-1$, and using the bound $\log(1+x)\geq x - x^2/2$, we obtain from (9):

$$\begin{aligned} I(\mathbf{Y};\mathbf{H}\,|\,\mathbf{X}) &= M_R\,\mathcal{E}_{\boldsymbol{\mathcal{X}}}\Bigg\{\sum_{r=0}^{M_TL-1}\log(1+\rho\,\mu_r(\boldsymbol{\mathcal{X}}))\Bigg\} &(17)\\ &\geq M_R\,\mathcal{E}_{\boldsymbol{\mathcal{X}}}\Bigg\{\sum_{r=0}^{M_TL-1}\Big(\rho\,\mu_r(\boldsymbol{\mathcal{X}}) - \frac{1}{2}\rho^{\,2}\mu_r^2(\boldsymbol{\mathcal{X}})\Big)\Bigg\}. \end{aligned}$$

Combined with (16), this yields the *fourthegy upper bound* [9] on the achievable rate R according to

$$R \leq \frac{1}{2T_{\mathrm{c}}}\,\mathcal{E}_{\boldsymbol{\mathcal{X}}}\{J_C(\boldsymbol{\mathcal{X}})\} \tag{18}$$

where

$$J_C(\boldsymbol{\mathcal{X}}) = M_R\,\frac{P^2}{W^2}\sum_{r=0}^{M_TL-1}\mu_r^2(\boldsymbol{\mathcal{X}})$$

is the fourthegy of $\boldsymbol{\mathcal{X}}$. Note that the fourthegy upper bound (18) is applicable for arbitrary (not just Gaussian) codebooks. We shall see later that (18) is tight only in the very low-SNR regime, where it approaches zero for i.i.d. Gaussian codebooks.

3) Upper Bound for Space-Frequency Unitary Codebooks: In [13], space-frequency unitary (SFU) codes for noncoherent MIMO-OFDM systems were introduced and analyzed from an error probability point of view. The discrete SFU codebook constructions in [13] can easily be extended to yield a continuous codebook. SFU codebooks do not result in i.i.d. entries of $\boldsymbol{\mathcal{X}}$, but satisfy $\boldsymbol{\mathcal{G}}^H(\boldsymbol{\mathcal{X}})\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{X}}) = (N/M_T)\mathbf{I}_{M_TL}$ *deterministically*. It follows that the $\mu_r(\boldsymbol{\mathcal{X}})$ are simply the scaled eigenvalues $(N/M_T)\sigma_l^2$ with multiplicity M_T each. Hence, from (17) we obtain the exact expression

$$I(\mathbf{Y};\mathbf{H}\,|\,\mathbf{X}) = M_TM_R\sum_{l=0}^{L-1}\log\Big(1+\frac{PT_{\mathrm{c}}}{M_T}\sigma_l^2\Big).$$

With (15), the achievable rate for SFU codebooks satisfies

$$\begin{aligned} R_{\mathrm{SFU}} \leq\; &M_RW\log\Big(1+\frac{P}{W}\Big) \\ &- \frac{M_TM_R}{T_{\mathrm{c}}}\sum_{l=0}^{L-1}\log\Big(1+\frac{PT_{\mathrm{c}}}{M_T}\sigma_l^2\Big). \end{aligned} \tag{19}$$

Note that the upper bound (19) can be interpreted as the capacity of M_R parallel AWGN channels up to the same penalty term as in the capacity lower bound (12).

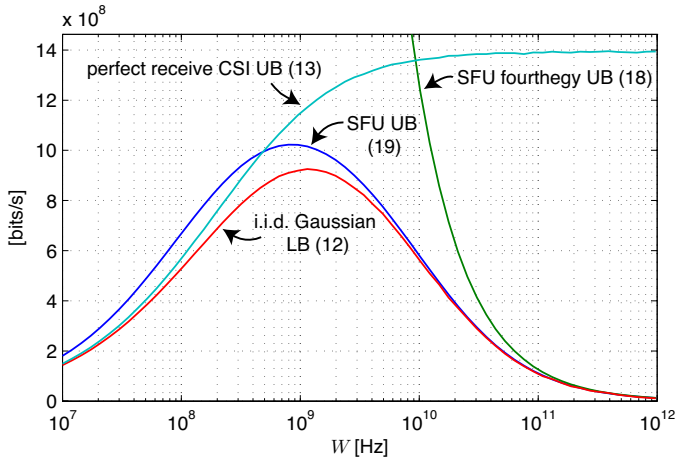

Fig. 1. Perfect receive CSI upper bound (UB) (13), upper bound (19) and fourthegy upper bound (18) for SFU codebooks, and i.i.d. Gaussian codebook lower bound (LB) (12).

IV. Discussion and Numerical Results

In this section, we discuss the implications of the capacity bounds derived above and provide numerical examples.

A. Overspreading

We start by investigating the behavior of the noncoherent MIMO-OFDM capacity for increasing bandwidth. Unless otherwise stated, all numerical results assume an US channel with uniform power delay profile of length $\tau = 20$ μs, a coherence time of $T_{\mathrm{c}} = 6$ ms, $P = -90$ dBm, and $M_T = M_R = 4$, which corresponds to a high-mobility outdoor scenario. Fig. 1 shows the bounds (12), (13), (19), and (18), where (18) is evaluated for SFU codebooks. We note that the behavior of the fourthegy upper bound for i.i.d. Gaussian codebooks is similar to that for SFU codebooks shown in Fig. 1. We can see that the Gaussian input lower bound (12) takes on a maximum at an intermediate bandwidth and then approaches zero. This effect is called "overspreading"—the number of unknown channel coefficients becomes so high that nonpeaky signaling schemes can no longer cope with channel uncertainty. The "overspreading" effect has been described previously for SISO wideband channels in [7]. The upper bound (19) for SFU codebooks follows a similar characteristic as the Gaussian input lower bound (12) and also points at the overspreading effect.

We denote the number of nonzero eigenvalues of $\mathbf{\Sigma}$, i.e., the number of stochastic degrees of freedom per SISO channel, by $\tilde{L}$, which is henceforth called the *effective number of channel taps*. In the following, we assume $\sigma_0^2 = \sigma_1^2 = \cdots = \sigma_{\tilde{L}-1}^2 = 1/\tilde{L}$, which, in the US case, corresponds to a uniform PDP. Analogously to [7], we define

$$\tilde{L}_{\mathrm{crit}} = \frac{PT_{\mathrm{c}}}{M_T}. \tag{20}$$

Using (20), we can now rewrite the fourthegy upper bound (18) for SFU codebooks as

$$R_{\mathrm{SFU}} \leq \frac{1}{2} M_R P \frac{\tilde{L}_{\mathrm{crit}}}{\tilde{L}}.$$

Thus R_{SFU} falls significantly below the AWGN capacity for $\tilde{L} \gg \tilde{L}_{\mathrm{crit}}$ (which implies low SNR). The lower bound (12) can be rewritten in terms of $\tilde{L}_{\mathrm{crit}}$ as

$$C \geq C_{\mathrm{coh}} - M_R P \frac{\tilde{L}}{\tilde{L}_{\mathrm{crit}}} \log\left(1 + \frac{\tilde{L}_{\mathrm{crit}}}{\tilde{L}}\right) \tag{21}$$

which shows that the penalty term is small for $\tilde{L} \ll \tilde{L}_{\mathrm{crit}}$. Thus, as in [7], $\tilde{L}_{\mathrm{crit}}$ can be interpreted as the critical parameter delineating the regime where overspreading occurs. From (20) it follows that for fixed PT_{c}, increasing the number of transmit antennas reduces $\tilde{L}_{\mathrm{crit}}$, so that overspreading occurs for smaller bandwidth if M_T is large. This effect is a consequence of the fact that increasing M_T results in an increase in the number of unknown channel parameters (recall that P is fixed).

B. Impact of the Number of Transmit Antennas

We shall now study the impact of the number of transmit antennas on the overspreading effect in detail. Again, for simplicity, we assume a uniform PDP so that

$$C \geq C_{\mathrm{coh}} - \frac{M_T M_R \tilde{L}}{T_{\mathrm{c}}} \log\left(1 + \frac{PT_{\mathrm{c}}}{M_T \tilde{L}}\right). \tag{22}$$

We note that C_{coh} is proportional to $\min(M_T, M_R)$ (at high SNR), whereas the pre-log in the penalty term scales in the *maximum achievable diversity order* [13] $M_T M_R \tilde{L}$, which equals the total number of stochastic degrees of freedom in the channel. From an error probability point of view, the number of transmit antennas M_T and the effective number of taps $\tilde{L}$ play interchangeable roles [13]; both contribute to lowering the error probability by increasing the diversity order (provided that appropriate codes are used). However, from a capacity (lower bound) point of view, increasing $\tilde{L}$ *decreases* the RHS of (22), since C_{coh} is independent of $\tilde{L}$, but the penalty term grows due to the higher number of unknown channel parameters. Increasing M_T also increases the penalty term, but at the same time increases C_{coh} [cf. (6)], so that we have two competing effects. Hence, the stochastic degrees of freedom due to the frequency selectivity of the channel and due to multiple transmit antennas, respectively, play vastly different roles in terms of capacity [at least in terms of the lower bound (22)].

Fig. 2(a) plots the capacity lower bound (12) for varying M_T; Fig. 2(b) shows the number of transmit antennas maximizing the lower bound for a given bandwidth. By switching off transmit antennas accordingly, the envelope of the capacity curves in Fig. 2(a) can be achieved. We can conclude that it is in general not optimal to use all transmit antennas in the overspread regime. In our example, the overspread regime begins in the practically relevant (e.g., for ultrawideband communication) low-GHz range. We finally note that it has been shown previously in [11] that in noncoherent MIMO channels not all transmit antennas should necessarily be used.

C. Impact of the Scaling Behavior of $\tilde{L}$

The fourthegy upper bound (18) can be used to show that for $\tilde{L} \propto W$ (uncorrelated scattering), the achievable rate both for i.i.d. Gaussian and SFU codebooks approaches zero in the wideband limit; hence, the penalty term $I(\mathbf{Y}; \mathbf{H} \mid \mathbf{X})$

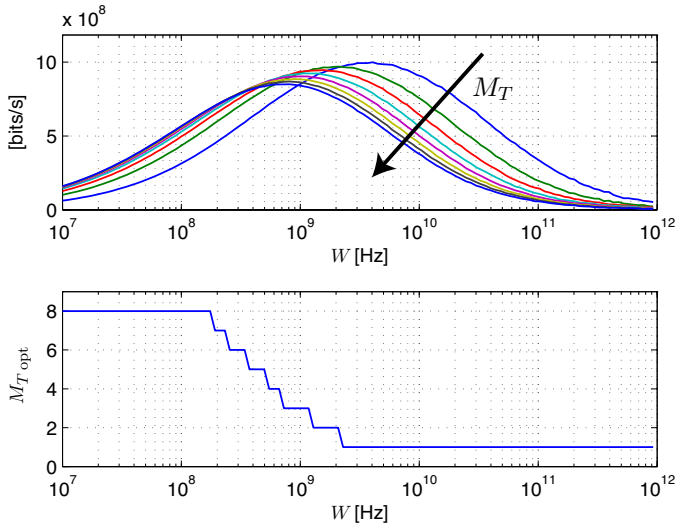

Fig. 2. (a) Lower bound on capacity (12) for varying M_T and $M_R = 4$. (b) Corresponding optimum number of transmit antennas M_T (with a maximum of $M_T = 8$).

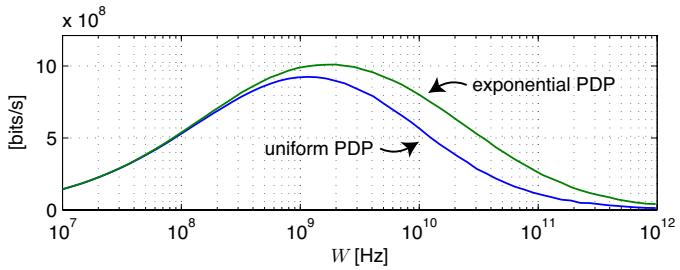

Fig. 3. Lower bound on capacity (12) for uniform and exponential PDPs.

approaches $I(\mathbf{Y};\mathbf{X},\mathbf{H})$. When $\tilde{L} \approx \lfloor W\tau \rfloor$, the parameter $\tilde{L}_{\mathrm{crit}}$ corresponds to a *critical bandwidth* $W_{\mathrm{crit}} \approx (P/M_T)(T_{\mathrm{c}}/\tau)$, which is on the order of 18 GHz for the system and channel parameters chosen in Section IV-A. For increasing T_{c}, with all other parameters fixed, W_{crit} increases and the channel behaves more and more like a coherent channel; as the delay spread τ increases, again all other parameters fixed, W_{crit} decreases and hence the overspread regime moves to lower bandwidths. The overspreading effect can be overcome by employing peaky signaling schemes [6]. In contrast, in the case of specular scattering, where $\tilde{L}$ is bounded as $W \to \infty$, the capacity is bounded away from zero by (21).

D. Impact of the Power Delay Profile

We finally investigate the impact of the PDP (or eigenvalue spread of $\mathbf{\Sigma}$) on capacity. From (12) we can see that a uniform PDP maximizes the penalty term in the capacity lower bound and the SFU codebook upper bound. For an US channel, Fig. 3 shows the lower bound (12) for a uniform PDP and for an exponentially decaying PDP, respectively. We can clearly see that the exponential PDP results in a higher critical bandwidth (beyond which overspreading occurs) than the uniform PDP.

V. Conclusion

We derived upper and lower bounds on the capacity of full-band MIMO-OFDM systems in wideband channels without transmit and receive CSI and discussed the corresponding implications. Beyond a certain critical bandwidth, "overspreading" occurs, i.e., the capacity goes to zero due to the large number of unknown channel parameters. Using multiple transmit antennas contributes to channel uncertainty. Our results show that for MIMO-OFDM systems operating at bandwidths of several GHz (such as ultrawideband MIMO systems), it is not advisable to use a large number of transmit antennas.

References

[1] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[2] 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, no. 1, pp. 139–157, Jan. 1999.

[3] A. Lapidoth and S. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[4] S. Ray, M. Médard, L. Zheng, and J. Abounadi, "On the sublinear behavior of MIMO channel capacity at low SNR," in *Proc. Int. Symp. Inf. Theory and its Applicat. (ISITA 2004)*, Parma, Italy, Oct. 2004.

[5] D. Schafhuber, H. Bölcskei, and G. Matz, "System capacity of wideband OFDM communications over fading channels without channel knowledge," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2004)*, Chicago, IL, June 2004, p. 391.

[6] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[7] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1384–1400, July 2000.

[8] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 840–852, Apr. 2002.

[9] V. G. Subramanian and B. Hajek, "Broad-band fading channels: Signal burstiness and capacity," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 809–827, Apr. 2002.

[10] J. N. Pierce, "Ultimate performance of M-ary transmission on fading channels," *IEEE Trans. Inf. Theory*, vol. 12, no. 1, pp. 2–5, Jan. 1966.

[11] Y. Liang and V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3095–3110, Dec. 2004.

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

[13] M. Borgmann and H. Bölcskei, "Noncoherent space-frequency coded MIMO-OFDM," *IEEE J. Select. Areas Commun.*, 2005, to appear.

[14] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

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

[16] O. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.

[17] S. Shamai (Shitz) and T. L. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 938–942, Apr. 2002.