The Space Frontier: Physical Limits of Multiple Antenna Information Transfer

Romain Couillet[1,2], Sebastian Wagner[1], Mérouane Debbah[2], and Alonso Silva[2]

[1]NXP Semiconductors, 06560 Sophia Antipolis, France

[2]Chaire Alcatel Lucent, Supelec, 91192 Gif sur Yvette, France

ABSTRACT

In this paper, we study the capacity limits of dense multi-antenna systems. We derive asymptotic capacity expressions for point-to-point and broadcast channels by applying recent tools from random matrix theory. In the case of broadcast channels, we focus on linear precoding techniques. Interestingly, the asymptotic capacity depends only on the ratio between the size of the antenna array and the wavelength. This provides useful guidelines on the achievable limits of information transfer. In particular, it is shown that the total capacity grows unbounded if the transmitter has perfect knowledge on the channel, while the capacity saturates in case of no channel knowledge at the transmitter. We provide numerical results supporting the theoretical derivations.

1. INTRODUCTION

Sixty years ago, Shannon [1] provided a mathematical framework to analyze fundamental limits of information transfer in the case of single-input single-output channels. He introduced the *channel capacity* as the maximum rate at which information can be reliably transmitted. From a purely theoretical point of view, there is no bound on the capacity as both bandwidth and power can be arbitrarily high. However, in practice, we can only transmit with finite power and over a restricted frequency band for physical reasons. Recently multiple-input multiple-output (MIMO) systems have been extensively studied since significant growth in terms of capacity has been predicted in [2], [3]. More specifically, in a system with n_{T} transmit and n_{R} receive antennas the capacity scales linearly with $\min(n_{\mathrm{T}}, n_{\mathrm{R}})$ for i.i.d (independent and identically distributed) Gaussian channels, at high signal-to-noise ratio (SNR). Again, MIMO systems suggest that the capacity can increase to infinity if the number of antennas grows large at both transmitter and receiver.

However, recent works [18] have shown that the capacity, even for an increasing number of antennas, is limited by the amount of scatterers in the environment. In other words, the number of antennas should be less than the number of degrees of freedom (modes) provided by the medium. Our goal is to show that, even when the medium offers an infinite number of modes, the capacity is mainly limited by the ratio between the size of the antenna array and the wavelength, which we call the *space frontier*. Indeed, in general, for a given space, increasing n_{T} or n_{R} decreases the relative distances between the antennas. Once the distance is less than half the transmit signal wavelength λ the antennas become correlated [16] and the capacity does not grow linearly anymore. In case of a circular antenna array it has been demonstrated by Pollock [4] that the capacity saturates if the number of antennas increases. In this work, we aim to extend Pollock's contribution to one- and two-dimensional antenna arrays. We study the capacity limits for point-to-point MIMO channels as well as for MIMO Gaussian broadcast channels (MIMO-GBC) with linear precoding. In the latter we assume a single transmitter modeled as a dense line of antennas which transmits to many independent single-antenna receivers. The general capacity solutions for those schemes are mathematically involved [9] and require the application of recent results from random matrix theory (RMT) and free probability [14].

This paper is organized as follows: Section 2 briefly introduces important tools from random matrix theory. Section 3 presents the capacity limits for the different MIMO channels. Section 4 gives simulation results validating our theoretical claims. Section 5 discusses the theoretical and practical implications of our results. Finally, section 6 states our conclusions.

Notation: In the following, boldface lower-case symbols represent vectors, capital boldface characters denote matrices ($\mathbf{I}_N$ is the $N \times N$ identity matrix). The Hermitian transpose is denoted $(\cdot)^{\mathsf{H}}$. The set of $N \times M$ matrices over the algebra $\mathcal{A}$ is denoted $\mathcal{M}(\mathcal{A}, N, M)$. The operators $\det(\mathbf{X})$ and $\mathrm{tr}(\mathbf{X})$ represent the determinant and the trace of matrix $\mathbf{X}$, respectively. The symbol $\mathrm{E}[\cdot]$ denotes expectation. The derivative of a function f of a single variable x is denoted $\frac{\mathrm{d}}{\mathrm{d}x} f$.

2. RANDOM MATRIX THEORY TOOLS

Since the pioneering work of Wigner [19] on the asymptotic empirical eigenvalue distribution of random hermitian matrices, random matrix theory has grown into a new field of research in theoretical physics and applied probability. The main application to communications lies in the derivation of asymptotic results for large matrices. Specifically, the eigenvalue distribution of large Hermitian matrices converges, in many practical cases, to a definite probability distribution, called *empirical distribution*. For instance, if $\mathbf{X} \in \mathcal{M}(\mathbb{C}, N, L)$ is a *Gaussian* matrix (i.e. a matrix with Gaussian i.i.d. entries), the eigenvalue distribution of the matrix $\frac{1}{L}\mathbf{X}\mathbf{X}^{\mathsf{H}}$ is known to converge, when $N, L \to \infty$ and $N/L \to c$, towards the *Marčhenko-Pastur law* μ_c [14].

RMT provides many tools to handle the empirical distribution of large random matrices. Among those tools, the *Stieltjes transform* $\mathcal{S}_{\mathbf{X}}$ of a large Hermitian matrix $\mathbf{X}$, defined on the half complex space $\{z \in \mathbb{C}, \mathrm{Im}(z) > 0\}$, is

$$\mathcal{S}_{\mathbf{X}}(z) = \int_{-\infty}^{+\infty} \frac{1}{\lambda - z} f(\lambda) \mathrm{d}\lambda \tag{1}$$

where f is the empirical distribution of $\mathbf{X}$.

Silverstein [13] derived a fixed-point expression of the Stieltjes transform for a particular random matrix structure in the following theorem,

THEOREM 1. *Let the entries of the $N \times K$ matrix $\mathbf{W}$ be i.i.d. (independent and identically distributed) with zero mean and variance $1/N$. Let $\mathbf{X}$ be an $N \times N$ Hermitian random matrix with an empirical eigenvalue distribution function converging weakly to $P_{\mathbf{X}}(x)$ almost surely.*

Moreover, let $\mathbf{Y}$ be a $K \times K$ real diagonal random matrix with an empirical distribution function converging almost surely in distribution to a probability distribution function $P_{\mathbf{Y}}(x)$ as $K \to \infty$. Then almost surely, the empirical eigenvalue distribution of the random matrix:

$$\mathbf{H} = \mathbf{X} + \mathbf{W}\mathbf{Y}\mathbf{W}^{\mathsf{H}} \tag{2}$$

converges weakly, as $K, N \to \infty$ but $K/N \to \alpha$ fixed, to the unique distribution function whose Stieltjes transform satisfies:

$$\mathcal{S}_{\mathbf{H}}(z) = \mathcal{S}_{\mathbf{X}}\left(z - \alpha \int \frac{y}{1 + y\mathcal{S}_{\mathbf{H}}(z)} \mathrm{d}P_{\mathbf{Y}}(y)\right) \tag{3}$$

This theorem is generalized by Girko [10] who derived a fixed-point equation for the Stieltjes transform of large Hermitian matrices $\mathbf{H} = \mathbf{W}\mathbf{W}^{\mathsf{H}}$ when $\mathbf{W}$ has independent entries w_{ij} with variance σ_{ij}^2/N such that the set $\left\{\sigma_{ij}^2\right\}_{i,j}$ is uniformly upper-bounded. In the following, we will extensively use this result to derive the asymptotic MIMO capacity.

3. FUNDAMENTAL CAPACITY LIMITS

3.1 Dense MIMO capacity

We first consider a MIMO system consisting of n_{T} transmitters and n_{R} receivers. We use the linear model in which the received vector $\mathbf{y} \in \mathbb{C}^{n_{\mathrm{R}}}$ depends on the transmitted vector $\mathbf{x} \in \mathbb{C}^{n_{\mathrm{T}}}$ via

$$\mathbf{y} = \sqrt{\frac{\rho}{n_{\mathrm{T}}}} \mathbf{H}\mathbf{x} + \mathbf{n} \tag{4}$$

where $\mathbf{H} \in \mathcal{M}(\mathbb{C}, n_{\mathrm{R}}, n_{\mathrm{T}})$, $\mathbf{n}$ is a zero mean circularly symmetric complex Gaussian noise with unit variance and ρ is the average SNR.

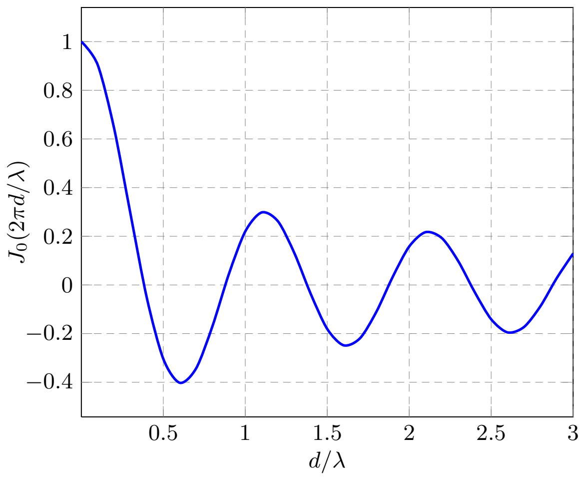

Figure 1: Spatial correlation vs. d/λ

Let the elements of the transmitted vector $\mathbf{x}$ be Gaussian with covariance matrix $\mathrm{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}] = \mathbf{\Phi}$. The ergodic channel capacity is given by [1]

$$C(n_{\mathrm{R}}, n_{\mathrm{T}}) = \mathrm{E}\left[\log\det\left(\mathbf{I}_{n_{\mathrm{R}}} + \frac{\rho}{n_{\mathrm{T}}}\mathbf{H}\mathbf{\Phi}\mathbf{H}^{\mathsf{H}}\right)\right] \tag{5}$$

Following Jakes' model [12], the spatial autocorrelation functions of fading processes $h_1(kd)$ and $h_2((k+1)d)$ experienced by two antennas on a line at position respectively kd and $(k+1)d$ is

$$\mathrm{E}[h_1(kd)h_2((k+1)d)^*] = J_0(2\pi d/\lambda) \tag{6}$$

where $\lambda = c/f_c$ denotes the carrier wavelength, and $J_0(x)$ is the zero-order Bessel function of the first kind. Thus the most immediate effect that results from locating various antennas in close proximity is that their signals tend to be, to some extent, correlated. The correlation function of Jake's model is depicted in figure 1.

In [4], Pollock et al. considered an increasing number of antennas in a uniform circular array of fixed radius. Using bounds on the Bessel function, Pollock derived an approximation of the channel capacity and shows that the capacity bound is independent of $(n_{\mathrm{R}}, n_{\mathrm{T}})$. In the following, we will extend these results using RMT. For a given $\beta \in \mathbb{R}^+$, we will consider that $n_{\mathrm{T}}/n_{\mathrm{R}} \to \beta$ when n_{T} and n_{R} grow large.

The entries of $\mathbf{H}$ represent the fading coefficients between each transmit and each receive antenna normalized such that

$$\mathrm{E}\left[\mathrm{tr}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)\right] = n_{\mathrm{R}} n_{\mathrm{T}} \tag{7}$$

while

$$\mathrm{E}[\|\mathbf{x}\|^2] = n_{\mathrm{T}} \tag{8}$$

It is useful to decompose the input covariance matrix $\mathbf{\Phi} = \mathrm{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}]$ in its eigenvectors and eigenvalues,

$$\mathbf{\Phi} = \mathbf{V}\mathbf{P}\mathbf{V}^{\mathsf{H}} \tag{9}$$

According to the *maximum entropy principle* [5], the most *uninformative* density function for $\mathbf{H}$, given n_R, n_T, l and λ, is the classical separable (also termed Kronecker or product) correlation model [6].

$$\mathbf{H} = \mathbf{\Theta}_\mathrm{R}^{1/2}\mathbf{H}_w\mathbf{\Theta}_\mathrm{T}^{1/2} \tag{10}$$

where the deterministic matrices $\mathbf{\Theta}_\mathrm{T}$ and $\mathbf{\Theta}_\mathrm{R}$ represent the correlation between the transmit antennas and the receive antennas, respectively. The entries of $\mathbf{H}_w$ are i.i.d. standard Gaussian (i.e. with zero-mean and unit variance).

With statistical channel-state information at the transmitter (CSIT), capacity is achieved if the eigenvectors of the input covariance $\mathbf{\Phi}$ coincide with those of $\mathbf{\Theta}_\mathrm{T}$. Consequently, denoting by $\mathbf{\Lambda}_\mathrm{T}$ and $\mathbf{\Lambda}_\mathrm{R}$ the diagonal eigenvalue matrices of $\mathbf{\Theta}_\mathrm{T}$ and $\mathbf{\Theta}_\mathrm{R}$ respectively we have

$$\mathcal{C}(\beta,\rho) = \lim_{n_\mathrm{T}\to\infty} \log\det\left(\mathbf{I} + \frac{\rho}{n_\mathrm{T}}\mathbf{\Lambda}_\mathrm{R}^{1/2}\mathbf{H}_w\mathbf{\Lambda}_\mathrm{T}^{1/2}\mathbf{P}\mathbf{\Lambda}_\mathrm{T}^{1/2}\mathbf{H}_w^\mathrm{H}\mathbf{\Lambda}_\mathrm{R}^{1/2}\right) \tag{11}$$

As a direct consequence of theorem 1:

THEOREM 2. *[11] The capacity of a Rayleigh-faded channel with separable transmit and receive correlation matrices $\mathbf{\Theta}_\mathrm{T}$ and $\mathbf{\Theta}_\mathrm{R}$ and statistical CSIT almost surely converges to*

$$\begin{aligned}\frac{\mathcal{C}(\beta,\rho)}{n_\mathrm{R}} \to & \beta\mathrm{E}[\log(1+\rho\cdot\lambda_\mathrm{T}\Gamma(\rho))] \\ & + \mathrm{E}[\log(1+\rho\cdot\lambda_\mathrm{R}\Upsilon(\rho))] \\ & - \beta\cdot\rho\cdot\Gamma(\rho)\Upsilon(\rho)\log e\end{aligned} \tag{12}$$

where

$$\Gamma(\rho) = \frac{1}{\beta}\mathrm{E}\left[\frac{\lambda_\mathrm{R}}{1+\rho\cdot\lambda_\mathrm{R}\Upsilon(\rho)}\right] \tag{13}$$

$$\Upsilon(\rho) = \mathrm{E}\left[\frac{\lambda_\mathrm{T}}{1+\rho\cdot\lambda_\mathrm{T}\Gamma(\rho)}\right] \tag{14}$$

and the dumb random variables λ_R, λ_T are asymptotically distributed as the diagonal elements of $\mathbf{\Lambda}_\mathrm{R}$ and $\mathbf{P}\mathbf{\Lambda}_\mathrm{T}$ respectively.

3.2 Antenna array geometry and correlation

3.2.1 1D scenario

We will consider a line of length l that is sending information to a parallel line of length l at a distance L. We consider that the number of antennas is n_T and n_R in the transmitting and receiving line respectively, separated uniformly. In the transmitting line, the antennas are put in the y-axis with the following positions $T_k, k\in[0, n_\mathrm{T}-1]$:

$$T_0 = (0,0),\ldots,T_i = (0, i\frac{l}{n_\mathrm{T}-1}),\ldots,T_n=(0,l) \tag{15}$$

The line of receivers is at a distance L from this line, parallel and with distribution of the antenna positions $R_k, k\in[0, n_\mathrm{R}-1]$:

$$R_0 = (L,0),\ldots,R_i = (L, i\frac{l}{n_\mathrm{R}-1}),\ldots,R_n=(L,l) \tag{16}$$

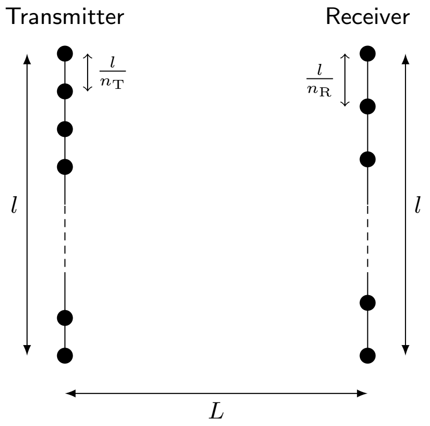

Figure 2: One-dimensional antenna array geometry

The antenna setup is depicted in figure 2.

The autocorrelation matrices $\mathbf{\Theta}_\mathrm{T}$ and $\mathbf{\Theta}_\mathrm{R}$ have the same form,

$$\begin{bmatrix} 1 & J_0(\frac{2\pi}{\lambda}\frac{l}{N-1}) & J_0(\frac{2\pi}{\lambda}\frac{2l}{N-1}) & \cdots & J_0(\frac{2\pi}{\lambda}\frac{(N-1)l}{N-1}) \\ J_0(\frac{2\pi}{\lambda}\frac{l}{N-1}) & 1 & J_0(\frac{2\pi}{\lambda}\frac{l}{N-1}) & \cdots & J_0(\frac{2\pi}{\lambda}\frac{(N-2)l}{N-1}) \\ J_0(\frac{2\pi}{\lambda}\frac{2l}{N-1}) & J_0(\frac{2\pi}{\lambda}\frac{l}{N-1}) & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ J_0(\frac{2\pi}{\lambda}\frac{(N-1)l}{N-1}) & J_0(\frac{2\pi}{\lambda}\frac{(N-1)l}{N-1}) & \cdots & \cdots & 1 \end{bmatrix} \tag{17}$$

with N equal to n_T or n_R respectively for $\mathbf{\Theta}_\mathrm{T}$ or $\mathbf{\Theta}_\mathrm{R}$. The normalized matrices $\frac{1}{n_\mathrm{R}}\mathbf{\Theta}_\mathrm{R}$ and $\frac{1}{n_\mathrm{T}}\mathbf{\Theta}_\mathrm{T}$ are *Wiener class* Toeplitz matrices [20], i.e.

$$\lim_{n_\mathrm{R}\to\infty}\frac{1}{n_\mathrm{R}}\sum_{k=1}^{n_\mathrm{R}}|\left[\mathbf{\Theta}_\mathrm{R}\right]_{1,k}| < \infty \tag{18}$$

There is no exact expression for the eigenvalues like in the case of a circulant matrix. However the eigenvalue distribution of a Wiener class Toeplitz matrix for large $(n_\mathrm{R}, n_\mathrm{T})$ converges to that of the circulant matrix, both with the same first row [20]. The set of the eigenvalues of $\frac{1}{n_\mathrm{R}}\mathbf{\Theta}_\mathrm{R}$ and $\frac{1}{n_\mathrm{T}}\mathbf{\Theta}_\mathrm{T}$ for large $(n_\mathrm{R}, n_\mathrm{T})$ is the image of the function

$$\begin{aligned} F_1: \mathbb{N} & \to \mathbb{R} \\ n & \mapsto \lim_{N\to\infty}\frac{1}{N}\sum_{p=-(N-1)}^{N-1} J_0\left(\frac{2\pi l}{\lambda}\frac{p}{N-1}\right)\cos(2\pi n\frac{p}{N}) \end{aligned} \tag{19}$$

$$= 2\int_{-1}^{1} J_0\left(\frac{2\pi l}{\lambda}x\right)\cos(2\pi nx)\mathrm{d}x \tag{20}$$

Since $F(\mathbb{N})$ is a discrete countable set (and not a continuum), the limit eigenvalue distribution for $\mathbf{\Theta}_\mathrm{T}$ and $\mathbf{\Theta}_\mathrm{R}$ is a sum of Dirac functions

$$p_\nu(\nu) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N}\delta\left(\nu - N\cdot F_1(k)\right) \tag{21}$$

At this point it is important to note that the cumulated surface of both antenna arrays must be constant regardless of n_{R} and n_{T}. Hence increasing the number of antennas must lead to a reduction of the individual antenna surface. As a result, the power per receive antenna must scale with $1/n_{\mathrm{R}}$, hence

$$\rho = \frac{\rho'}{n_{\mathrm{R}}} \tag{22}$$

for a constant total SNR ρ'.

We first consider the case where no CSIT is available, hence a uniform power allocation over the transmit antennas is optimal (i.e. $\mathbf{P} = \mathbf{I}_{n_{\mathrm{T}}}$).

Applying theorem 2 and expanding the expectations for large $(n_{\mathrm{R}}, n_{\mathrm{T}})$, we have

$$\begin{aligned} \mathcal{C}(\beta,\rho') =& \beta n_{\mathrm{R}} \frac{1}{n_{\mathrm{T}}} \sum_{k=0}^{n_{\mathrm{T}}} \log(1 + \frac{\rho'}{n_{\mathrm{R}}} n_{\mathrm{T}} F_1(k) \Gamma) \\ &+ n_{\mathrm{R}} \frac{1}{n_{\mathrm{R}}} \sum_{k=0}^{n_{\mathrm{R}}} \log(1 + \frac{\rho'}{n_{\mathrm{R}}} n_{\mathrm{R}} F_1(k) \Upsilon) \\ &- n_{\mathrm{R}} \beta \cdot \frac{\rho'}{n_{\mathrm{R}}} \cdot \Gamma(\rho') \Upsilon(\rho') \log(e) \end{aligned} \tag{23}$$

with

$$\Gamma(\rho') = \frac{1}{\beta n_{\mathrm{R}}} \sum_{k=0}^{n_{\mathrm{R}}} \frac{n_{\mathrm{R}} F_1(k)}{1 + \rho' F_1(k) \Upsilon} \tag{24}$$

$$\Upsilon(\rho') = \frac{1}{n_{\mathrm{T}}} \sum_{k=0}^{n_{\mathrm{T}}} \frac{n_{\mathrm{T}} F_1(k)}{1 + \rho' \beta F_1(k) \Gamma} \tag{25}$$

In the limits, this is

$$\begin{aligned} \mathcal{C}(\beta,\rho') \rightarrow & \sum_{k=0}^{+\infty} \log(1 + \rho' \beta F_1(k) \Gamma) \\ &+ \sum_{k=0}^{+\infty} \log(1 + \rho' F_1(k) \Upsilon) \\ &- \beta \rho' \Gamma(\rho') \Upsilon(\rho') \log(e) \end{aligned} \tag{26}$$

with

$$\Gamma(\rho') = \frac{1}{\beta} \sum_{k=0}^{+\infty} \frac{F_1(k)}{1 + \rho' F_1(k) \Upsilon} \tag{27}$$

$$\Upsilon(\rho') = \sum_{k=0}^{+\infty} \frac{F_1(k)}{1 + \rho' \beta F_1(k) \Gamma} \tag{28}$$

where $\forall k \in \mathbb{N}$, $F_1(k) \geq 0$ as eigenvalues of a covariance matrix. Also $\sum_{k=0}^{+\infty} F_1(k) = 1$ as the trace of $\Theta_{\mathrm{R}}/n_{\mathrm{R}}$. This implies straightforwardly that Γ and Υ are finite and therefore the total capacity $\mathcal{C}$ is also finite.

Further note that (21) only depends on the system parameters through the ratio l/λ. This leads to the conclusion that the MIMO capacity limit with no CSIT only depends on the ratio l/λ and β.

Consider now the case of perfect CSIT. Here, it is optimal to distribute the power according to the water-filling solution [21]. That is, only *sufficiently strong* eigenmodes of the channel (10) are used for transmission. If we allocate the power constrained by (8) on the dominating channel eigenmodes (i.e. the relevant eigenvalues of $\frac{1}{n_{\mathrm{T}}} \mathbf{\Theta}_{\mathrm{R}}^{1/2} \mathbf{H}_w \mathbf{\Theta}_{\mathrm{T}} \mathbf{H}_w^{\mathrm{H}} \mathbf{\Theta}_{\mathrm{R}}^{1/2}$), then for large n_{T}, the capacity grows unbounded. As a result, increasing the number of antennas at either side of the transmission allows to achieve *arbitrarily high* capacity under the assumption of perfect CSIT. However, CSIT has a cost in terms of rate which can be extremely high in mobile environments.

3.2.2 2D scenario

The previous scheme can be extended to two dimensions. Here we increase the density of antennas uniformly along each dimension of the surface. Consider a rectangular surface of respective height and width l_x and l_y.

Then, equation (21) has an equivalent version in two dimensions,

$$p_\nu(\nu) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N} \delta(\nu - N \cdot F_2(k)) \tag{29}$$

with

$$\begin{aligned} F_2(k) =& 2 \int_{-l_x}^{l_x} \int_{-l_y}^{l_y} J_0 \left(\frac{2\pi l_x}{\lambda} \sqrt{u_x^2 + \left(\frac{l_y}{l_x}\right)^2 u_y^2} \right) \\ & \times \cos(2\pi\nu(u_x + u_y)) \mathrm{d}u_x \mathrm{d}u_y \end{aligned} \tag{30}$$

From (29), one verifies that the final capacity formulation depends on the two constant values l_x/λ and l_x/l_y, or similarly on the two ratios l_x/λ and l_y/λ. Note that the capacity for a given surface might then differ depending on the shape of the surface.

3.3 MIMO-GBC capacity

We assume a GBC generated by a multi-antenna transmitter and many non-cooperative single-antenna receivers. It has been shown in [8] that the capacity region is achieved by Dirty-Paper Coding (DPC). However, to derive closed-form expressions, we restrict our analysis in the following to suboptimal linear precoding techniques. As the receivers are uncorrelated, the GBC channel model is

$$\mathbf{H} = \mathbf{H}_w \mathbf{\Theta}_{\mathrm{T}}^{1/2} \tag{31}$$

and the transmitted signal $\mathbf{x}$ is obtained by

$$\mathbf{x} = \mathbf{G}\mathbf{u} \tag{32}$$

where the entries of the data vector $\mathbf{u} \in \mathbb{C}^{n_{\mathrm{R}}}$ are of unit power, and $\mathbf{G} \in \mathcal{M}(\mathbb{C}, n_{\mathrm{T}}, n_{\mathrm{R}})$.

3.3.1 ZF-beamforming

Zero-Forcing (ZF) beamforming is a mere channel inversion precoding. If the data vector $\mathbf{u}$ of unit power is intended to be transmitted then, with the same notation as in previous sections, the channel model reads

$$\mathbf{y} = \sqrt{\frac{\rho}{n_{\mathrm{T}}}} \mathbf{H}\mathbf{x} + \mathbf{n} \tag{33}$$

with

$$\begin{cases} \mathbf{x} = \alpha \mathbf{H}'^{-1}\mathbf{u} & \text{, if } \beta = 1 \\ \mathbf{x} = \alpha \left(\mathbf{H}'^{\mathsf{H}}\mathbf{H}'\right)^{-1} \mathbf{H}'^{\mathsf{H}} & \text{, if } \beta > 1 \end{cases} \tag{34}$$

for $\mathbf{H}' = \sqrt{\frac{1}{n_\mathrm{T}}}\mathbf{H}$.

The parameter α is set here to fulfill the transmission power constraint

$$\mathrm{E}[\|\mathbf{x}\|^2] = n_\mathrm{T} \tag{35}$$

which leads to

$$\alpha^2 = \frac{1}{\frac{1}{n_\mathrm{T}}\mathrm{tr}\,(\mathbf{H}'\mathbf{H}'^{\mathsf{H}})^{-1}} \tag{36}$$

where

$$\frac{1}{n_\mathrm{T}}\mathrm{tr}\left(\mathbf{H}'\mathbf{H}'^{\mathsf{H}}\right)^{-1} \to \int \frac{1}{\nu} f(\nu)\mathrm{d}\nu \tag{37}$$

with f the empirical distribution of $\mathbf{H}'\mathbf{H}'^{\mathsf{H}}$.

Contrary to [22] in which no power limitation is imposed to $\mathbf{x}$, no asymptotic expression for α is known when $(n_\mathrm{R}, n_\mathrm{T})$ grow large to the authors' knowledge.

We recognize in (37) the Stieltjes transform of $f(x)$ in $x = 0$.

Recall that $\mathbf{H}\mathbf{H}^{\mathsf{H}} = \mathbf{H}_w \boldsymbol{\Theta}_\mathrm{T} \mathbf{H}_w^{\mathsf{H}}$. Thus by diagonalizing $\boldsymbol{\Theta}_\mathrm{T} = \mathbf{U}\boldsymbol{\Lambda}_\mathrm{T}\mathbf{U}^{\mathsf{H}}$ with $\mathbf{U}$ a unitary matrix, we have

$$\mathbf{H}'\mathbf{H}'^{\mathsf{H}} = \left(\frac{1}{\sqrt{n_\mathrm{T}}}\mathbf{H}_w\mathbf{U}\right)\boldsymbol{\Lambda}_\mathrm{T}\left(\frac{1}{\sqrt{n_\mathrm{T}}}\mathbf{U}^{\mathsf{H}}\mathbf{H}_w^{\mathsf{H}}\right) \tag{38}$$

in which the entries of $\frac{1}{\sqrt{n_\mathrm{T}}}\mathbf{H}_w\mathbf{U}$ are i.i.d. with zero mean and variance $\frac{1}{n_\mathrm{T}}$, and $\boldsymbol{\Lambda}_\mathrm{T}$ is distributed as in (21).

Applying theorem 1, we then prove the existence of $\mathcal{S}_{\mathbf{H}'\mathbf{H}'^H}$, when $n_\mathrm{T}/n_\mathrm{R} \to \beta$, that satisfies

$$\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(z) = \left(-z + \beta\int \frac{\nu p_\nu(\nu)}{1 + \nu \cdot \mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(z)}\mathrm{d}\nu\right)^{-1} \tag{39}$$

When expanding $\mathbf{x}$ in the system model (4), one obtains parallel AWGN channels with the per user capacity:

$$\mathcal{C}_u(\beta, \rho) = \log(1 + \rho\alpha^2) \tag{40}$$

$$= \log(1 + \rho\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)^{-1}) \tag{41}$$

In our specific correlation scenario, this capacity limit is in fact null. Indeed, if $n_\mathrm{T} = n_\mathrm{R}$,

$$\frac{1}{n_\mathrm{T}}\mathrm{tr}\left(\mathbf{H}'\mathbf{H}'^{\mathsf{H}}\right)^{-1} = \frac{1}{n_\mathrm{T}}\mathrm{tr}\left((\mathbf{U}^{\mathsf{H}}\mathbf{H}_w^{\mathsf{H}}\mathbf{H}_w\mathbf{U})^{-1}\boldsymbol{\Lambda}_\mathrm{T}^{-1}\right) \tag{42}$$

where $\tilde{\mathbf{H}}_w = \mathbf{H}_w\mathbf{U}$ is a Gaussian random matrix with entries of variance $1/n_\mathrm{T}$.

Lemma 3. *For any two Hermitian $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ with eigenvalues $\lambda_i(\mathbf{A})$ and $\lambda_i(\mathbf{B})$ respectively arranged in decreasing order,*

$$\mathrm{tr}(\mathbf{AB}) \geq \sum_{i=1}^{n} \lambda_i(\mathbf{A})\lambda_{n-i+1}(\mathbf{B}) \tag{43}$$

From lemma 3, we have that

$$\mathrm{tr}\left((\tilde{\mathbf{H}}_w^{\mathsf{H}}\tilde{\mathbf{H}}_w)^{-1}\boldsymbol{\Lambda}_\mathrm{T}^{-1}\right) \geq \sum_{i=1}^{n_\mathrm{R}} \lambda_i((\tilde{\mathbf{H}}_w^{\mathsf{H}}\tilde{\mathbf{H}}_w)^{-1})\lambda_{n-i+1}(\boldsymbol{\Lambda}_\mathrm{T}^{-1}) \tag{44}$$

The eigenvalues of $\tilde{\mathbf{H}}_w^{\mathsf{H}}\tilde{\mathbf{H}}_w$ are known [14] to be asymptotically distributed as the *Marčhenko-Pastur* law on a bounded (positive) support excluding zero. Therefore the eigenvalues of $(\tilde{\mathbf{H}}_w^{\mathsf{H}}\tilde{\mathbf{H}}_w)^{-1}$ are also bounded on a finite positive support. Call λ_{min} the minimum of those eigenvalues, we have,

$$\mathrm{tr}\left((\tilde{\mathbf{H}}_w^{\mathsf{H}}\tilde{\mathbf{H}}_w)^{-1}\boldsymbol{\Lambda}_\mathrm{T}^{-1}\right) \geq \lambda_{min}\sum_{i=1}^{n_\mathrm{R}} \lambda_i(\boldsymbol{\Lambda}_\mathrm{T}^{-1}) \tag{45}$$

Observing then that

$$\lambda_{n_\mathrm{R}}(\boldsymbol{\Lambda}_\mathrm{T}) = \sum_{p=-(n_\mathrm{R}-1)}^{n_\mathrm{R}-1} J_0\left(\frac{2\pi l}{\lambda}\frac{p}{n_\mathrm{R}-1}\right)\cos(2\pi\frac{p}{N}) \to 0 \tag{46}$$

we conclude

$$\mathrm{tr}\left((\tilde{\mathbf{H}}_w^{\mathsf{H}}\tilde{\mathbf{H}}_w)^{-1}\boldsymbol{\Lambda}_\mathrm{T}^{-1}\right) \to +\infty \tag{47}$$

Therefore, $\alpha^2 \to 0$ and the ZF capacity goes down to zero for increasing $n_\mathrm{R} = n_\mathrm{T}$. The case $n_\mathrm{T} > n_\mathrm{R}$ solves by dividing $\tilde{\mathbf{H}}_w$ in a block of size $n_\mathrm{R} \times n_\mathrm{R}$ and a block $(n_\mathrm{T} - n_\mathrm{R}) \times n_\mathrm{R}$, and observing that the capacity limited for the former already grows to infinity.

3.3.2 MMSE-beamforming

Let us consider in this section the case of regularized beamforming. We still have the system model in (33) with:

$$\mathbf{x} = \left(\mathbf{H}'^{\mathsf{H}}\mathbf{H}' + \alpha\mathbf{I}_{n_\mathrm{T}}\right)^{-1}\mathbf{H}'^{\mathsf{H}}\mathbf{u} \tag{48}$$

When $\alpha = 0$, we fall back the zero-forcing solution. The parameter α is set so to fulfill the transmission power constraint (8) which leads to

$$1 = \frac{1}{n_\mathrm{T}}\mathrm{tr}\left(\left(\mathbf{H}'^{\mathsf{H}}\mathbf{H}' + \alpha\mathbf{I}\right)^{-1}\mathbf{H}'^{\mathsf{H}}\mathbf{H}'\left(\mathbf{H}'^{\mathsf{H}}\mathbf{H}' + \alpha\mathbf{I}\right)^{-1}\right) \tag{49}$$

$$= \frac{1}{n_\mathrm{T}}\mathrm{tr}\left(\left(\mathbf{H}'^{\mathsf{H}}\mathbf{H}' + \alpha\mathbf{I}\right)^{-2}\mathbf{H}'^{\mathsf{H}}\mathbf{H}'\right) \tag{50}$$

$$\to \int \frac{\nu}{(\nu+\alpha)^2}f(\nu)\mathrm{d}\nu \tag{51}$$

$$= \int \left(\frac{1}{(\nu+\alpha)} - \frac{\alpha}{(\nu+\alpha)^2}\right)f(\nu)\mathrm{d}\nu \tag{52}$$

$$= \mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(-\alpha) + \alpha\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(-\alpha) \tag{53}$$

The received vector signal can be written as:

$$\mathbf{y} = \sqrt{\rho}\cdot\mathbf{H}'\left(\mathbf{H}'^{\mathsf{H}}\mathbf{H} + \alpha\mathbf{I}\right)^{-1}\mathbf{H}'^{\mathsf{H}}\mathbf{u} + \mathbf{n} \tag{54}$$

Let us denote $\mathbf{H}'^{\mathsf{H}} = [\mathbf{h}_1, \ldots, \mathbf{h}_{n_\mathrm{R}}]$. We will focus on user i without loss of generality. At the output of receiver i,

$$\begin{aligned} y_i &= \sqrt{\rho} \cdot \mathbf{h}_i^{\mathsf{H}} \left(\mathbf{H}'^{\mathsf{H}} \mathbf{H}' + \alpha \mathbf{I} \right)^{-1} \mathbf{h}_i u_i \\ &+ \sum_{k=1, k \neq i}^{n_\mathrm{R}} \mathbf{h}_i^{\mathsf{H}} \left(\mathbf{H}'^{\mathsf{H}} \mathbf{H}' + \alpha \mathbf{I} \right)^{-1} \mathbf{h}_k u_k \\ &+ \mathbf{n} \end{aligned} \tag{55}$$

Lemma 4. *[15] Let $\mathbf{A}$ be a deterministic $N \times N$ complex matrix with uniformly bounded spectral radius for all N. Let $\mathbf{x} = \frac{1}{\sqrt{N}} [x_1, \ldots, x_N]^T$ where the $\{x_i\}$ are i.i.d complex random variables with zero mean, unit variance and finite eighth moment. Then*

$$\mathbb{E}\left[\mid \mathbf{x}^{\mathsf{H}} \mathbf{A} \mathbf{x} - \frac{1}{N} \mathrm{tr} \mathbf{A} \mid^4 \right] \leq \frac{c}{N^2} \tag{56}$$

where c is a constant that does not depend on N or $\mathbf{A}$.

Corollary 5. *This result ensures that*

$$\mathbf{x}^{\mathsf{H}} \mathbf{A} \mathbf{x} - \frac{1}{N} \mathrm{tr} \mathbf{A} \to 0 \tag{57}$$

almost surely.

Henceforth we write $\mathbf{U}_i^{\mathsf{H}} = [\mathbf{h}_1, \ldots, \mathbf{h}_{i-1}, \mathbf{h}_{i+1}, \ldots, \mathbf{h}_{n_\mathrm{R}}]$ (in other words, we remove column i). Applying the matrix inversion lemma yields

$$\mathbf{h}_i^{\mathsf{H}} \left(\mathbf{H}'^{\mathsf{H}} \mathbf{H}' + \alpha \mathbf{I} \right)^{-1} = \frac{\mathbf{h}_i^{\mathsf{H}} \left(\mathbf{U}_i^{\mathsf{H}} \mathbf{U}_i + \alpha \mathbf{I} \right)^{-1}}{1 + \mathbf{h}_i^{\mathsf{H}} \left(\mathbf{U}_i^{\mathsf{H}} \mathbf{U}_i + \alpha \mathbf{I} \right)^{-1} \mathbf{h}_i}$$

We can therefore use lemma 4 as the elements of $\mathbf{h}_i$ are i.i.d. due to the one sided correlation assumption (otherwise, it would not work when correlation is considered on both sides).

Since the removal of a single column in the large matrix $\mathbf{H}'$ does not affect $\mathrm{tr}(\mathbf{H}'^{\mathsf{H}} \mathbf{H}')$, we asymptotically have

$$\mathbf{h}_i^{\mathsf{H}} \left(\mathbf{U}_i^{\mathsf{H}} \mathbf{U}_i + \alpha \mathbf{I} \right)^{-1} \mathbf{h}_i \to \mathbf{h}_i^{\mathsf{H}} \left(\mathbf{H}'^{\mathsf{H}} \mathbf{H}' + \alpha \mathbf{I} \right)^{-1} \mathbf{h}_i \tag{58}$$

hence

$$\mathbf{h}_i^{\mathsf{H}} \left(\mathbf{H}'^{\mathsf{H}} \mathbf{H}' + \alpha \mathbf{I} \right)^{-1} \to \frac{\mathbf{h}_i^{\mathsf{H}} \left(\mathbf{U}_i^{\mathsf{H}} \mathbf{U}_i + \alpha \mathbf{I} \right)^{-1}}{1 + \mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(-\alpha)}$$

Denote $\gamma = (1 + \mathcal{S}_{\mathbf{H}\mathbf{H}^{\mathsf{H}}}(-\alpha))^{-2}$. The SINR expression is therefore given by:

$$\mathsf{SINR}_i = \frac{\rho \gamma \mathbf{h}_i^{\mathsf{H}} \mathbf{W}_i^2 \mathbf{h}_i}{\rho \gamma \mathbf{h}_i^{\mathsf{H}} \mathbf{W}_i \mathbf{U}_i^{\mathsf{H}} \mathbf{U}_i \mathbf{W}_i \mathbf{h}_i + 1} \tag{59}$$

with $\mathbf{W}_i = \left(\mathbf{U}_i^{\mathsf{H}} \mathbf{U}_i + \alpha \mathbf{I} \right)^{-1}$. This leads in the limit to a *user-independent* SINR:

$$\mathsf{SINR} \to \frac{\rho \gamma \mathcal{S}^2_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(-\alpha)}{\rho \gamma \left(\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(-\alpha) + \alpha \frac{\mathrm{d}}{\mathrm{d}x} \mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(-\alpha) \right) + 1} \tag{60}$$

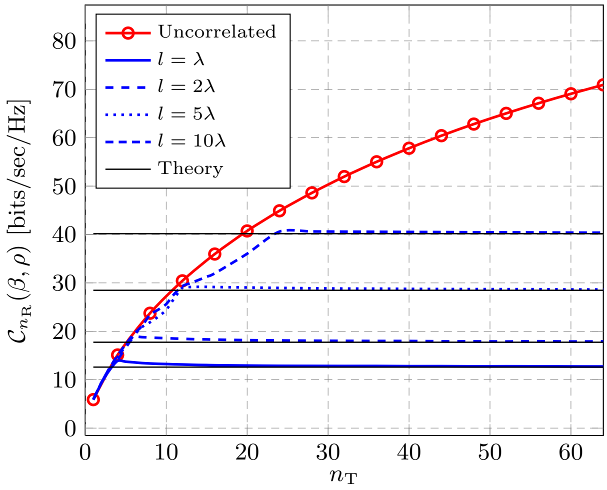

Figure 3: Ergodic MIMO capacity *without* CSIT $\mathcal{C}_{n_\mathrm{R}}(\beta, \rho)$ for different l/λ, $n_\mathrm{R} = n_\mathrm{T}$, $\rho' = 20$ dB

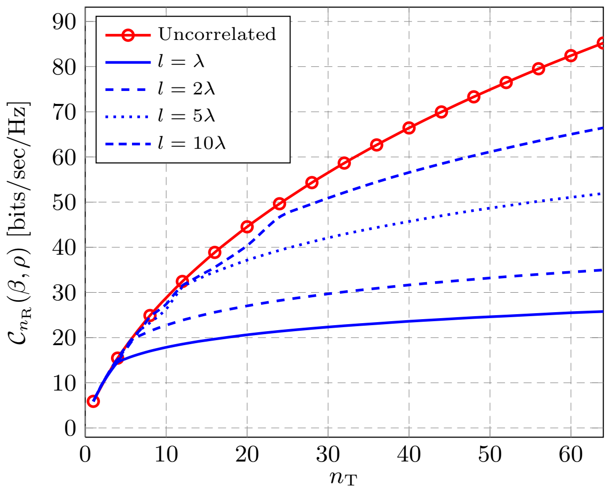

Figure 4: Ergodic MIMO capacity *with* CSIT $\mathcal{C}_{n_\mathrm{R}}(\beta, \rho)$ for different l/λ, $n_\mathrm{R} = n_\mathrm{T}$, $\rho' = 20$ dB

The corresponding achievable rate per user is:

$$\mathcal{C}_u(\beta, \rho) = \log(1 + \mathsf{SINR}) \tag{61}$$

Diagonalizing $\mathbf{U}^{\mathsf{H}}\mathbf{U}$, we observe that the numerator in (59) converge to finite strictly positive values (for the regularization term α ensures that no term diverges). However, as already noted, the strongest eigenvalue for $\mathbf{\Theta}_\mathrm{T}$ grows linearly with n_T, hence, thanks to lemma 3, the denominator grows to infinite for large n_T. This proves that the per-user capacity falls down to zero. Therefore, similarly to the case of ZF-beamforming, MMSE-beamforming does not ensure non-null per-user capacity for large $(n_\mathrm{R}, n_\mathrm{T})$.

As a consequence, it turns out that additional antennas might impair the achievable transmission rate. This is explained by the fact that loading power on more and more correlated antennas, instead of available channel modes, is an inefficient power allocation strategy.

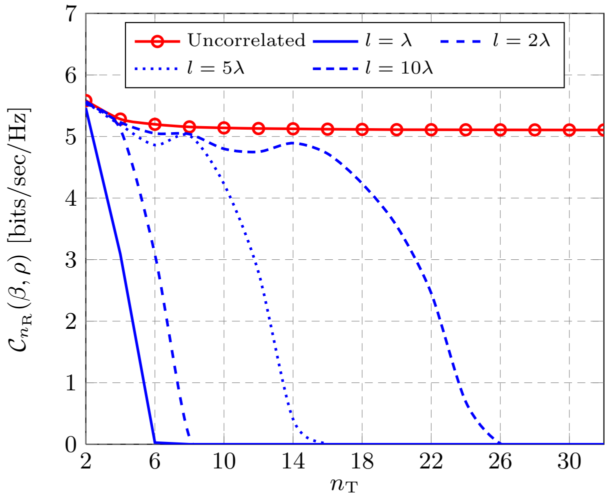

Figure 5: Per-user MIMO-GBC-ZF capacity $\mathcal{C}_{n_{\mathrm{R}}}(\beta, \rho)$ for different l/λ - $n_{\mathrm{T}} = \frac{3}{2}n_{\mathrm{R}}$

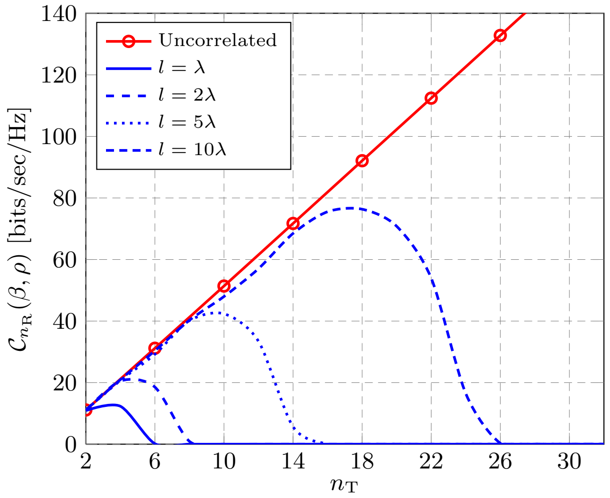

Figure 6: MIMO-GBC-ZF capacity $n_{\mathrm{R}}\mathcal{C}_{n_{\mathrm{R}}}(\beta, \rho)$ for different l/λ - $n_{\mathrm{T}} = \frac{3}{2}n_{\mathrm{R}}$

4. SIMULATION AND RESULTS

Let us first consider the MIMO scenario with dense arrays of antennas both at transmitter and receiver sides. Figures 3 and 4 present the results of ergodic capacities found by numerical simulation. The latter are compared against the theoretical limits derived from (11) (which is obtained by solving numerically (27),(28)). Note that the capacity increases first to a maximum for small $(n_{\mathrm{R}}, n_{\mathrm{T}})$ and then decreases to the limit capacity.

In figure 3, equal power allocation is applied (i.e. $\mathbf{P} = \mathbf{I}_{n_{\mathrm{T}}}$) to the transmitting line model. We observe, as previously concluded, that the capacity saturates for large $(n_{\mathrm{T}}, n_{\mathrm{R}})$. The saturation only depends on the ratios β and l/λ. In figure 4, we apply waterfilling (i.e. loading the transmit power on the dominant eigenmodes of the channel), which leads to a non-saturating capacity.

In the case of MIMO-GBC with correlated transmitters, uncorrelated receivers and linear ZF-precoding, the per-user capacity saturates for large $(n_{\mathrm{R}}, n_{\mathrm{T}})$ when $n_{\mathrm{T}} > n_{\mathrm{R}}$. If transmit antenna correlation is considered the per-user capacity is going to zero. Note that the capacity increases for low $(n_{\mathrm{R}}, n_{\mathrm{T}})$ until a turning point where the transmit antenna correlation leads to an ill-conditioning of the overall channel matrix. The maximum capacity is reached for $d < \lambda/2$. This is observed in figure 5. We also provide, in figure 6, the total capacity of ZF-beamforming for $n_{\mathrm{T}} = \frac{3}{2}n_{\mathrm{R}}$. It is observed that even the total capacity cannot sustain the poor conditionning of large channel matrices and the capacity decreases down to zero. In the simulation results of figure 7, MMSE-beamforming is applied. Since no closed-form solution for α under the constraint (53) is available, the algorithm carries out an exhaustive search for this parameter. We observe again that the per-user capacity is going asymptotically zero, which is in accordance with equations (60) and (61). The same observation can be made for the overall system capacity in figure 8. In both cases the curves are decreasing less rapidly with increasing $(n_{\mathrm{R}}, n_{\mathrm{T}})$ than in the ZF precoding scheme.

5. DISCUSSION

A few limitations are worth mentioning about our previous conclusions. In the MIMO case we stated that, under perfect CSIT, the channel capacity grows unbounded even with a strong antenna correlation at the transmitter side. This might indicate that densifying the array of transmit antennas is the preferred option to increase the capacity (rather than increasing the transmitted power or the channel bandwidth). However, perfect CSIT implies that the receiver has to feed back channel information to the transmitter. For a dense MIMO system, this introduces an immense feedback overhead, thus reducing the achievable throughput.

The same conclusion stands for channel state information at the receiver (CSIR). As Tse demonstrated [17], the capacity with perfect CSIR is limited by the coherence time of the channel. If the number of antennas grows, one needs to estimate more and more degrees of freedom with less power. An optimal trade-off must then be found between increasing the number of antennas (and thus the capacity) and decreasing the amount of channel state information required for reliable transmission.

However, if the channel coherence time is infinite and a long synchronization stage prior to data transmission is allowed, then the channel capacity can effectively go unbounded. The only limitation that would then appear lies in the physical ability to design a dense array of antennas on a limited surface. Also, for the model to be realistic, an increasingly large number of scatterers in the medium is mandatory.

6. CONCLUSION

In this work we analyzed the asymptotic capacity of different dense antenna MIMO configurations. We have shown in particular that in the absence of CSIT, the capacity is bounded and related to the ratio between the size of the antenna array and the wavelength. However, we found that the capacity is unbounded if perfect CSIT is available. This rate can only be achieved with a great amount of feedback which might be impossible for time varying channels.

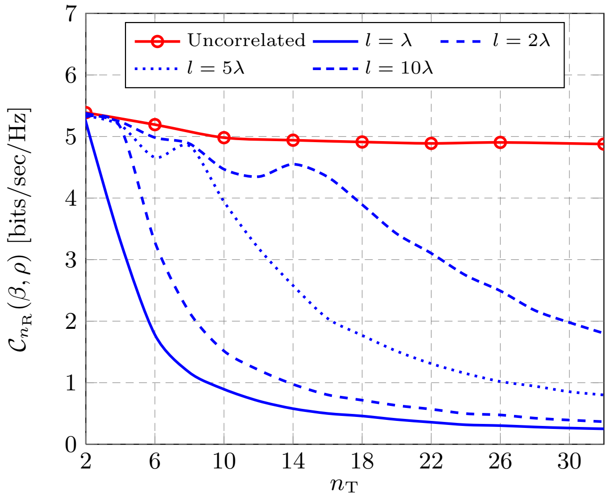

Figure 7: MIMO-GBC-MMSE per user capacity $\mathcal{C}_{n_\mathrm{R}}(\beta, \rho)$ for different l/λ - $n_\mathrm{T} = \frac{3}{2}n_\mathrm{R}$

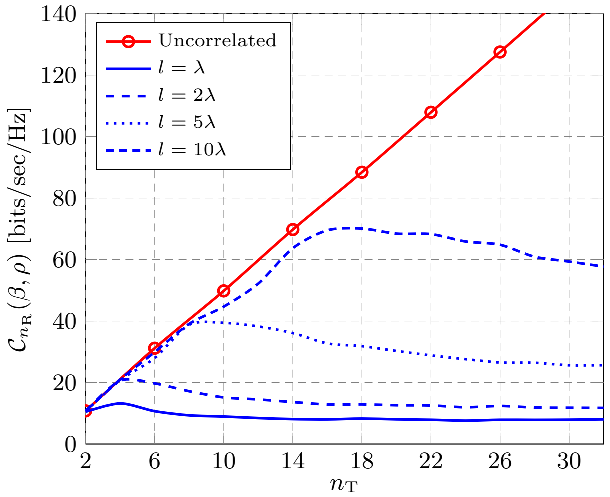

Figure 8: MIMO-GBC-MMSE capacity $n_\mathrm{R}\mathcal{C}_{n_\mathrm{R}}(\beta, \rho)$ for different l/λ - $n_\mathrm{T} = \frac{3}{2}n_\mathrm{R}$

7. REFERENCES

[1] C. E. Shannon, “A Mathematical Theory of Communication”, The Bell System Technical Journal, Vol. 27, pp. 379-423, 623-656, July, October, 1948.

[2] G. J. Forschini and M. J. Gans, “On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas”, Wireless Personal Communications, vol. 6, no. 3, pp. 311-335, 1998.

[3] I.E. Telatar, “Capacity of multi-antenna Gaussian channels,” Tech. Rep., ATT Bell Labs, 1995.

[4] T.S. Pollock, T.D. Abhayapala, and R.A. Kennedy, “Antenna saturation effects on dense array MIMO capacity,” in IEEE Intl. Conf. on Commun., ICC’03, Anchorage, Alaska, May. 11-15 2003, pp. 2301-2305.

[5] E. T. Jaynes, “Probability Theory: the Logic of Science”, Cambridge University Press, 2003.

[6] M. Guillaud, M. Debbah, A. Moustakas, “Maximum Entropy MIMO Wireless Channel Models”, Submitted to IEEE Trans. on Information Theory dec. 2006, arxiv.org/abs/cs.IT/0612101.

[7] D. V. Voiculescu, “Free Probability Theory”, American Mathematical Society, 1997.

[8] P. Viswanath, D. Tse, “Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality”, vol. 49, no. 8, pp. 1912-1921, 2003.

[9] G. Caire and S. Shamai, “On the achievable throughput of a multiantenna Gaussian broadcast channel”, IEEE Trans. on Information Theory, vol. 49, no. 7, pp. 1691-1706, 2003.

[10] V. L. Girko, “Theory of random determinants. Transl. from the Russian”, Mathematics and Its Applications (Soviet Series), 45, 1988.

[11] A.M. Tulino, A. Lozano, and S. Verdú, “Impact of correlation on the capacity of multi-antenna channels,” IEEE Trans. on Information Theory), vol. 51, no. 7, pp. 2491-2509, 2005.

[12] P. Dent, G. E. Bottomley and T. Croft, “Jakes fading model revisited”, Electonics Letters, vol. 29, no. 13, June 1993.

[13] J. W. Silverstein and Z.D Bai. “On the empirical distribution of eigenvalues of a class of large dimensional random matrices,” Journal of Multivariate Analysis, 54: 175-192, 1995

[14] A. M. Tulino, S. Verdú, “Random Matrix Theory and Wireless Communications”, Now Publishers, vol. 1, Issue 1, 2004.

[15] Z. D. Bai and J. W. Silverstein, “No eigenvalue outside the support of the limiting spectral distribution of large dimensional sample covariance matrices”, Ann. Probab. vol. 26, no. 1, pp. 316-345, 1998.

[16] L.Schumacher K.I. Pedersen and P. E. Mogensen, “From antenna spacings to theoretical capacities - Guidelines for simulating MIMO systems”, PIMRC 2002

[17] L. Zheng and D. Tse, “Communicating on the Grassmann Manifold: A Geo- metric Approach to Noncoherent Multi-antenna Channels”, IEEE Transactions on Information Theory , vol. 48, February 2002, pp. 359-383.

[18] M. Debbah and R. Müller, ”MIMO Channel Modeling and the Principle of Maximum Entropy”, IEEE Transactions on Information Theory, Vol. 51 , May, No. 5, 2005

[19] E. Wigner, ”Random Matrices in Physics”, SIAM Review, Vol. 9, No. 1, January, 1967.

[20] R.M. Gray, ”Toeplitz and Circulant Matrices: A Review”, Foundations and Trends in Communications and Information Theory, 2006, vol. 2, No. 3, pp. 155-239

[21] D. Tse and P. Viswanath, ”Fundamentals of Wireless Communication”, Cambridge University Press, 2005.

[22] C. Peel, B. Hochwald and A. Swindlehurst, “A vector-perturbation technique for near-capacity multiantenna multiuser communication - Part I: Channel inversion and regularization”, IEEE Trans. on Communications, vol. 53, no. 1, pp. 195-202, Jan. 2005.