Impact of Antenna Correlation on the Capacity of Multiantenna Channels

Antonia M. Tulino,[*] Angel Lozano[†] and Sergio Verdú[‡]

March 9, 2005

Abstract

This paper applies random matrix theory to obtain analytical characterizations of the capacity of correlated multiantenna channels. The analysis is not restricted to the popular separable correlation model, but rather it embraces a more general representation that subsumes most of the channel models that have been treated in the literature. For arbitrary signal-to-noise ratios (SNR), the characterization is conducted in the regime of large numbers of antennas. For the low- and high-SNR regions, in turn, we uncover compact capacity expansions that are valid for arbitrary numbers of antennas and that shed insight on how antenna correlation impacts the tradeoffs between power, bandwidth and rate.

Keywords:

Channel capacity; Multiantenna arrays; Random matrices; Fading channels; Antenna correlation;

I Introduction

A The UIU Model

$\mathcal{C}$ONSIDER n_{T} transmit and n_{R} receive antennas with the corresponding multiantenna channel described by the complex model

$$\mathbf{H}_i = \mathbf{U}_{\mathrm{R}}\tilde{\mathbf{H}}_i\mathbf{U}_{\mathrm{T}}^{\dagger} \tag{1}$$

where $\mathbf{U}_{\mathrm{R}}$ and $\mathbf{U}_{\mathrm{T}}$ are $n_{\mathrm{R}} \times n_{\mathrm{R}}$ and $n_{\mathrm{T}} \times n_{\mathrm{T}}$ deterministic unitary matrices while $\{\tilde{\mathbf{H}}_i\}$ is a stationary and ergodic sequence of $n_{\mathrm{R}} \times n_{\mathrm{T}}$ random matrices. The entries of $\tilde{\mathbf{H}}_i$ are zero-mean and independent with arbitrary marginal distributions and variances, constrained only to satisfy the normalization[1]

$$E\left[\mathrm{Tr}\left\{\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\dagger}\right\}\right] = n_{\mathrm{R}}n_{\mathrm{T}}. \tag{2}$$

Because of the Unitary-Independent-Unitary structure of (1), we refer to this model as UIU. The goal of this paper is to characterize the capacity of UIU channels, whose defining property is being unitarily equivalent (i.e., same singular values) to a channel with IND (independent nonidentically distributed) entries. It will be also useful to recognize that $\tilde{\mathbf{H}}$ is the KLT (Karhunen-Loève transform) of $\mathbf{H}$.

The UIU framework encompasses most of the zero-mean channel models that have been treated in the multiantenna literature:

- If $\tilde{\mathbf{H}}$ has IID (independent identically distributed) Rayleigh-faded entries, then $\mathbf{H}$ is also IID Rayleigh-faded and we are faced with the canonical channel considered in the early pioneering works in multiantenna communication [1, 2]. While very simple, this canonical model enabled a number of fundamental observations on the benefits that accrue as function of the numbers of antennas. Subsequent contributions built around the canonical model include [3]–[5].

- If the entries of $\tilde{\mathbf{H}}$ are Rayleigh-faded with variances that conform to a certain pattern (cf. Section III), the UIU channel reverts to the *separable* correlation model [6]–[8]. Equivalently, this model can be expressed as

$$\mathbf{H} = \mathbf{\Theta}_{\mathrm{R}}^{1/2}\mathbf{H}_w\mathbf{\Theta}_{\mathrm{T}}^{1/2} \tag{3}$$

where $\mathbf{\Theta}_{\mathrm{R}}$ and $\mathbf{\Theta}_{\mathrm{T}}$ are deterministic receive and transmit correlation matrices while $\mathbf{H}_w$ is IID Rayleigh-faded. If the entries of $\mathbf{H}$ are identically distributed, then the diagonal

---

[1]Henceforth, we drop the temporal index i.

entries of $\boldsymbol{\Theta}_{\mathrm{R}}$ and $\boldsymbol{\Theta}_{\mathrm{T}}$ equal 1. The most salient feature of the model in (3) is that the correlation between two transmit antennas is restricted to be the same irrespective of the receive antenna where it is observed and, conversely, the correlation between two receive antennas is restricted to be the same irrespective of the transmit antenna where it is observed. In channels spanned by *space diversity* arrays,[2] this constraint is often satisfied since the immediate surroundings to each array determine the spatial correlation between its antennas but have no impact on the spatial correlation at the other end of the link [9]. The separable model is often adequate and analytically friendly. It has been—and it continues to be—the cornerstone of a large number of analyses [10]–[16].

- If $\mathbf{U}_{\mathrm{R}} = \mathbf{I}$ and $\mathbf{U}_{\mathrm{T}} = \mathbf{I}$, then the UIU channel reduces to an IND channel. This channel can describe the use of either *polarization diversity*[3] or *pattern diversity*,[4] which are becoming increasingly popular complements—and even alternatives—to space diversity [19]. These forms of diversity, especially alluring in small handheld terminals, often result in minimal correlation between the channel coefficients but possibly large disparities in their strength and distribution.

- If $\mathbf{U}_{\mathrm{R}}$ and $\mathbf{U}_{\mathrm{T}}$ are Fourier matrices while $\tilde{\mathbf{H}}$ is IND Rayleigh-faded, the UIU model delivers the *virtual* representation pioneered in [20]. At the expense of restricting its scope to space diversity ULAs (uniform linear arrays), this representation enables the insightful physical interpretation of the columns of $\mathbf{U}_{\mathrm{R}}$ and $\mathbf{U}_{\mathrm{T}}$ as steering vectors receiving and launching energy on specific spatial directions.

- If $\mathbf{U}_{\mathrm{R}}$ and $\mathbf{U}_{\mathrm{T}}$ arbitrary unitary matrices while $\tilde{\mathbf{H}}$ is IND Rayleigh-faded, we obtain the generalization of the virtual model proposed in [21] and [22] and experimentally validated in the latter. This representation encompasses Rayleigh-faded channels with correlation structures that cannot generally be represented through the separable model [23].

The above channels give rise to distinct correlation structures whose impact on the capacity we seek to determine. Models that fall outside the UIU framework include:

- The *keyhole* channel, where $\mathbf{H}$ is given by the outer product of two random vectors with zero-mean independent entries [24, 25]. The entries of $\mathbf{H}$ are thus uncorrelated (but not independent). The capacity of this channel is studied in [5].

---

[2]By *space diversity* arrays we mean those composed of homogeneous antennas. Diversity is obtained exclusively through antenna spacing.

[3]Antennas with orthogonal polarizations may provide low levels of correlation with minimum or no antenna spacing [17] while making the communication link robust to polarization rotations in the channel [18].

[4]Antennas with different radiation patterns or with rotated versions of the same pattern (in which case the term *angle diversity* is also common) may discriminate different multipath components and reduce correlation.

- The channel with diagonal correlations considered in [26, 27]. In this model, defined explicitly for $n_{\text{T}} = n_{\text{R}} = 2$, the so-called diagonal correlations (between $h_{1,1}$ and $h_{2,2}$ on one hand and between $h_{1,2}$ and $h_{2,1}$ on the other hand) are free parameters rather than related to the other correlations as in the UIU model.

Note also that the UIU model only accommodates zero-mean channels. As it turns out, nonetheless, for large numbers of antennas the addition to the model of an unfaded (i.e, deterministic) matrix becomes immaterial in terms of the capacity as long as such matrix does not have full rank. For arbitrary numbers of antennas, capacity characterizations for nonzero-mean channels are available in the specific case of the IID Ricean distribution [28, 29, 30].

B Scope

The capacity of a UIU channel becomes analytically tractable asymptotically, either in the large-dimensional regime or as the signal-to-noise ratio (SNR) approaches zero or infinity. Fortunately, the usually rapid convergence to these limits renders the expressions derived therein relevant to a wide range of nonasymptotic scenarios [9, 31, 32]. Accordingly, we seek to characterize the capacity (*i*) asymptotically in the number of antennas for arbitrary SNR, and (*ii*) for arbitrary numbers of antennas at low and high SNR. Along the way, we exemplify the applicability of these analytical characterizations to several channels of interest. The specific contributions of the paper are:

- We find the asymptotic ($n_{\text{R}}, n_{\text{T}} \to \infty$) capacity of IND channels. Although, in its full generality, such capacity is the solution of a fixed-point equation, for a class of IND matrices that embodies several channels of interest an explicit expression is found.

- We show how, exploiting their unitary equivalence, the asymptotic capacity of the various UIU structures described in the previous section can be mapped from that of an IND channel. Particular emphasis is given to the separable correlation model.

- Several key low- and high-SNR performance measures are characterized for arbitrary numbers of antennas. In addition, we analyze the high-SNR *power offset* introduced in [33] (see also [34]).

- Besides the capacity, the spectral efficiency achievable with isotropic inputs is also characterized.

In addition to single-user multiantenna communication, IND channels surface in other contexts such as multiaccess, broadcast, multicell, etc. Some of our results in a conference version of this paper [35] have already been applied to these related problems [36, 37].

4

The manuscript is organized as follows. Section II puts forth notation and some useful definitions. Section III details the various classes of correlation structures that are accommodated by the UIU model. In Section IV, the capacity-achieving input is identified. Section V is devoted to the large-dimensional analysis for arbitrary SNR. The analysis in Section VI, in contrast, is for arbitrary numbers of antennas. Finally, Section VII asserts the impact of antenna correlation by drawing insight from the analysis in the preceding sections.

In the large-dimensional analysis of the capacity, the main tool is the theory of large random matrices [38]. Specifically, the mathematical framework provided by Girko [39] proves instrumental in the asymptotic analysis of IND channels presented in Section V. Girko's theorem, reproduced in Appendix B, has been applied to diverse problems such as multicarrier CDMA [40] and CDMA networks with multiuser receivers [41]. It has also been used to examine the asymptotic capacity scaling with the number of antennas of certain multiantenna channels [10, 42]. A very fertile result, Girko's theorem allows us not only to characterize the asymptotic multiantenna capacity but to express it as sole function of the MMSE (minimum mean-square error) incurred in the linear estimation of the corresponding input. This dependence imbues the solutions with direct engineering significance. For certain classes of IND matrices, Girko's framework reverts to the more restrictive Silverstein-Bai framework [43], which has also seen extensive use in the information theory literature [33, 44, 45]. Furthermore, for a specific class of IND matrices, both the Girko and Silverstein-Bai frameworks are shown in Section V to be equivalent to the classical Marčenko-Pastur law that governs the asymptotic behavior of matrices with IID entries [46].

II   Preliminaries

A   Definitions

Our complex baseband model is

$$\mathbf{y} = \sqrt{g}\,\mathbf{H}\,\mathbf{x} + \mathbf{n} \tag{4}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the input and output vectors while $\mathbf{n}$ is white Gaussian noise with one-sided spectral density

$$N_0 = \frac{E[\|\mathbf{n}\|^2]}{n_{\mathrm{R}}}. \tag{5}$$

The channel is represented by the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ zero-mean random matrix $\sqrt{g}\,\mathbf{H}$ where $\mathbf{H} = \{h_{i,j}\}$ and $\mathbf{h}_j$ is the jth column of $\mathbf{H}$. When the fading process is frequency selective, the

channel can be decomposed into a number of sufficiently narrow parallel subchannels, each undergoing frequency-flat fading and having the same ergodic capacity as the aggregate channel. Thus, it suffices to consider flat fading which results in the entries $h_{i,j}$ being complex scalars. These entries are normalized in accordance with (2), which puts different channels on an equal footing in the sense that any differences in average gain are factored out of $\mathbf{H}$ and absorbed into

$$\mathsf{SNR} \triangleq g \frac{E[\|\mathbf{x}\|^2]}{N_0}. \tag{6}$$

The covariance of the input, normalized by its energy per dimension, is denoted by

$$\mathbf{\Phi} \triangleq \frac{E[\mathbf{xx}^\dagger]}{\frac{1}{n_\mathrm{T}} E[\|\mathbf{x}\|^2]} \tag{7}$$

where the normalization ensures that $\mathrm{Tr}\{\mathbf{\Phi}\} = n_\mathrm{T}$. If the input is isotropic, then $\mathbf{\Phi} = \mathbf{I}$. An eigenvalue decomposition of $\mathbf{\Phi}$ yields

$$\mathbf{\Phi} = \mathbf{VPV}^\dagger \tag{8}$$

with $\mathbf{V}$ unitary and $\mathbf{P}$ diagonal. The columns of $\mathbf{V}$ represent the signaling directions while the diagonal entries of $\mathbf{P}$ represent the normalized powers allocated to those directions, constrained such that $\mathrm{Tr}\{\mathbf{P}\} = n_\mathrm{T}$.

Our focus is on channels known instantaneously at the receiver but known only in distribution at the transmitter. The mutual information in bits/s/Hz achieved by a zero-mean Gaussian input is

$$I(\mathsf{SNR}, \mathbf{\Phi}) = \log_2 \det \left(\mathbf{I} + \frac{\mathsf{SNR}}{n_\mathrm{T}} \mathbf{H\Phi H}^\dagger \right) \tag{9}$$

from which the ergodic spectral efficiency is $E[I(\mathsf{SNR}, \mathbf{\Phi})]$ and the ergodic capacity is

$$C(\mathsf{SNR}) = \max_{\mathbf{\Phi}:\mathrm{Tr}\{\mathbf{\Phi}\}\leq n_\mathrm{T}} E\left[I(\mathsf{SNR}, \mathbf{\Phi})\right] \tag{10}$$

Additional definitions that will be invoked are:

Definition 1 *A $n_\mathrm{R} \times n_\mathrm{T}$ matrix $\mathbf{A}$ taking values in $\mathcal{A} \subset \Re^+$ is mean column-regular if the mean of the entries of every column is the same, i.e., if* [5]

$$\frac{1}{n_\mathrm{R}} \sum_{i=1}^{n_\mathrm{R}} (\mathbf{A})_{i,j}$$

does not depend on j.

---

[5]Throughout the paper, $(\cdot)_{i,j}$ indicates the (i,j)th entry of a matrix.

Definition 2 *A matrix* $\mathbf{A}$ *is mean row-regular if* $\mathbf{A}^T$ *is mean column-regular.*

Definition 3 *A matrix* $\mathbf{A}$ *is mean doubly-regular if it is both mean column-regular and mean row-regular.*

Some stricter regularity conditions can also be defined [38]:

Definition 4 *A* $n_{\mathrm{R}} \times n_{\mathrm{T}}$ *matrix* $\mathbf{A}$ *taking values in* $\mathcal{A} \subset \Re^+$ *is column-regular if the entries of every column exhibit the same empirical distribution, i.e., if*

$$\frac{1}{n_{\mathrm{R}}} \sum_{i=1}^{n_{\mathrm{R}}} 1\{(\mathbf{A})_{i,j} < \xi\}$$

does not depend on j, *with* $1\{\cdot\}$ *the indicator function.*

Definition 5 *A matrix* $\mathbf{A}$ *is row-regular if* $\mathbf{A}^T$ *is column-regular.*

Clearly, a column-regular matrix is always mean column-regular and a row-regular matrix is always mean row-regular.

When any of the regularity conditions in Definitions 1–5 takes hold only as the size of the matrix goes to infinity, we say that the corresponding matrix is asymptotically regular.

B   Performance Measures

In order to evaluate the tradeoffs between power and bandwidth [13], it is often convenient to express the capacity, not only as function of the SNR, but also as function of the transmitted energy per bit relative to the noise level

$$\frac{E_b}{N_0} \quad = \quad \frac{E[\|\mathbf{x}\|^2]}{N_0\, C(\mathsf{SNR})} \tag{11}$$

$$= \quad \frac{\mathsf{SNR}}{g\, C(\mathsf{SNR})}. \tag{12}$$

For many purposes, expansions of $C(\frac{E_b}{N_0})$ in the low- and high-SNR regions suffice. The value of these expansions is reinforced by the rapid transition between the low- and the high-SNR regimes. We shall evaluate these expressions for $\frac{E_b}{N_0}$ in dB.

C   Low SNR

Denoting by $\frac{E_b}{N_0}_{\min}$ the minimum energy per bit required for reliable communication and by S_0 the slope therein in bits/s/Hz/(3 dB), the low-SNR capacity can be posed as [47]

$$C(\tfrac{E_b}{N_0}) = S_0 \frac{\frac{E_b}{N_0}\big|_{\mathrm{dB}} - \frac{E_b}{N_0}{}_{\min}\big|_{\mathrm{dB}}}{3\,\mathrm{dB}} + \epsilon \tag{13}$$

with $3\,\mathrm{dB} \triangleq 10\log_{10} 2$ and

$$\epsilon = o\left(\frac{E_b}{N_0}\big|_{\mathrm{dB}} - \frac{E_b}{N_0}{}_{\min}\big|_{\mathrm{dB}} \right). \tag{14}$$

Using (11), we can unravel (13) and relate the capacity to SNR via

$$\mathsf{SNR}\big|_{\mathrm{dB}} = \frac{C}{S_0}\,3\,\mathrm{dB} + 10\log_{10} C + \left(g\,\frac{E_b}{N_0}{}_{\min} \right)\big|_{\mathrm{dB}} + o(C). \tag{15}$$

D   High SNR

At high SNR, the capacity behaves as [33]

$$C(\mathsf{SNR}) = S_\infty \left(\frac{\mathsf{SNR}\big|_{\mathrm{dB}}}{3\,\mathrm{dB}} - \mathcal{L}_\infty \right) + o(1)$$

where S_∞ denotes the high-SNR slope in bits/s/Hz/(3 dB) while $\mathcal{L}_\infty$ represents the high-SNR power offset, in 3-dB units, with respect to a reference channel whose dimensions are unfaded and orthogonal, i.e., such that $\frac{1}{n_{\mathrm{T}}}\mathbf{HH}^\dagger = \mathbf{I}$. In multiantenna channels, S_∞ exhibits little dependence on the structure of $\mathbf{H}$ and thus, in most cases, only $\mathcal{L}_\infty$ reflects the distinctiveness of each channel at high SNR [34].

To relate the high-SNR capacity with $\frac{E_b}{N_0}$, we can again use (11) to obtain [33]

$$\frac{E_b}{N_0}\big|_{\mathrm{dB}} = \left(\frac{C}{S_\infty} + \mathcal{L}_\infty \right) 3\,\mathrm{dB} - 10\log_{10} C - g\big|_{\mathrm{dB}} + o(1).$$

We note that S_∞ represents the high-SNR slope in terms of either SNR or $\frac{E_b}{N_0}$, both measured in 3-dB units.

III   The UIU Model Revisited

Since our focus is on characterizing the impact of antenna correlation, a detailed description of the correlation structures accommodated by the UIU model is in order. To this

end, let us denote the correlation between the (i,j)th and (i',j')th entries of $\mathbf{H}$ by

$$R_{\mathbf{H}}(i,j;i',j') \triangleq E\left[h_{i,j}h_{i',j'}^*\right]. \tag{16}$$

A tool that facilitates the connection between correlated channels and the UIU model is the KLT, widely used in signal and image processing.

A   The Karhunen-Loève Transform

Applied to any channel $\mathbf{H}$ with zero-mean correlated entries, the KLT delivers a transformed channel whose entries are uncorrelated and given by

$$\tilde{h}_{k,\ell} = \sum_{i=1}^{n_{\mathrm{R}}} \sum_{j=1}^{n_{\mathrm{T}}} h_{i,j}\,\psi_{k,\ell}^*(i,j) \tag{17}$$

where the KLT kernel $\{\psi_{k,\ell}(i,j)\}$ is a set of complete orthonormal discrete basis functions formed by the eigenfunctions of $R_{\mathbf{H}}$, i.e., satisfying

$$\sum_{i'=1}^{n_{\mathrm{R}}} \sum_{j'=1}^{n_{\mathrm{T}}} R_{\mathbf{H}}(i,j;i',j')\,\psi_{k,\ell}(i',j') = \lambda_{k,\ell}(R_{\mathbf{H}})\,\psi_{k,\ell}(i,j) \tag{18}$$

where $\lambda_{i,j}(\cdot)$ denotes the (i,j)th eigenvalue.

A.1   Factorable KLT Kernel

If the KLT kernel can be factored as

$$\psi_{k,\ell}(i,j) = u_k(i)\,v_\ell(j), \tag{19}$$

then

$$\mathbf{H} = \mathbf{U}_{\mathrm{R}}\tilde{\mathbf{H}}\mathbf{U}_{\mathrm{T}}^\dagger \tag{20}$$

with

$$(\mathbf{U}_{\mathrm{R}})_{i,k} = u_k(i) \tag{21}$$
$$(\mathbf{U}_{\mathrm{T}})_{\ell,j} = v_j^*(\ell) \tag{22}$$

which renders $\mathbf{U}_{\mathrm{R}}$ and $\mathbf{U}_{\mathrm{T}}$ unitary. Condition (19) is tantamount to the requirement that the covariance and cross-covariance matrices of the rows of $\mathbf{H}$ share the same eigenvectors, $\mathbf{U}_{\mathrm{R}}$, while the covariance and cross-covariance matrices of the columns of $\mathbf{H}$ share

the same eigenvectors, $\mathbf{U}_\mathrm{T}$ [20, 22]. Whenever (19) is upheld and the entries of $\tilde{\mathbf{H}}$ are independent (e.g., if $\mathbf{H}$ is Rayleigh-faded), then the channel falls within the UIU framework. The columns of $\mathbf{U}_\mathrm{R}$ and $\mathbf{U}_\mathrm{T}$ correspond, respectively, to the eigenvectors of $E[\mathbf{H}\mathbf{H}^\dagger]$ and $E[\mathbf{H}^\dagger\mathbf{H}]$ whereas the variances of the entries of $\tilde{\mathbf{H}}$ are

$$E\left[|\tilde{h}_{i,j}|^2\right] = \lambda_{i,j}(R_\mathbf{H}) \tag{23}$$

evidencing that the $n_\mathrm{R} n_\mathrm{T}$ eigenvalues of $R_\mathbf{H}$, which determine the distribution of the mutual information, are embedded in the variances of the entries of $\tilde{\mathbf{H}}$.

If (19) is not upheld, then the KLT cannot be expressed as in (20) and the corresponding correlation falls outside the UIU framework.

A.2 The Virtual Representation

If $\mathbf{H}$ is Rayleigh-faded and $R_\mathbf{H}(i, j; i', j')$ is a function of only $i - i'$ and $j - j'$, as may be the case when both transmitter and receiver are equipped with space diversity ULAs, then the channel falls within the UIU framework and, as the number of antennas grows large, $\mathbf{U}_\mathrm{R}$ and $\mathbf{U}_\mathrm{T}$ in (22) become Fourier matrices. This gives way to the intuitive virtual representation propounded in [20]. (For arbitrary numbers of antennas, the independence between the entries of $\tilde{\mathbf{H}}$ in the virtual representation is only approximate.) Furthermore, the variances of the entries of $\tilde{\mathbf{H}}$ can in this case be physically interpreted as the power coupling for specific transmit-receive angle pairs [42].

A.3 Separability

The separable correlation model is defined by (3), whereby $R_\mathbf{H}(i, j; i', j')$ can be expressed as the product of two marginal correlations that are functions respectively of (i,i') and (j,j'), i.e.,

$$R_\mathbf{H}(i, j; i', j') = (\mathbf{\Theta}_\mathrm{R})_{i,i'}(\mathbf{\Theta}_\mathrm{T})_{j,j'}. \tag{24}$$

As a result,

$$\lambda_{i,j}(R_\mathbf{H}) = \lambda_i(\mathbf{\Theta}_\mathrm{R})\lambda_j(\mathbf{\Theta}_\mathrm{T}) \tag{25}$$

evidencing that the $n_\mathrm{R} n_\mathrm{T}$ eigenvalues of $R_\mathbf{H}$ are determined by the eigenvalues of $\mathbf{\Theta}_\mathrm{R}$ and $\mathbf{\Theta}_\mathrm{T}$, that is, by only $n_\mathrm{R} + n_\mathrm{T} - 1$ quantities. This restriction on the number of degrees of freedom in the eigenvalues of $R_\mathbf{H}$ evidences the narrower scope of the separable correlation model [20, 23, 48].

The separable model in (3) admits a UIU expression in which $\mathbf{U}_\mathrm{R}$ and $\mathbf{U}_\mathrm{T}$ coincide with the eigenvector matrices of $\mathbf{\Theta}_\mathrm{R}$ and $\mathbf{\Theta}_\mathrm{T}$, respectively, while $\tilde{\mathbf{H}}$ is an IND Rayleigh-faded channel whose (i, j)th entry has variance $\lambda_i(\mathbf{\Theta}_\mathrm{R})\lambda_j(\mathbf{\Theta}_\mathrm{T})$.

B IND Channels

Since the core of the UIU model is an IND channel, the approach taken in the sequel is the following. Initially, we analyze IND channels $\tilde{\mathbf{H}}$ and, subsequently, we show how the corresponding capacities can be applied to other UIU channels by virtue of their unitary equivalence. In addition to serving as a stepping stone towards more general results, the IND channel is interesting on its own for it models situations where correlation is negligible but the arrays contain antennas with distinct polarizations or radiation patterns. It is also of interest in multiuser and multicell contexts [36, 37].

In order to manipulate IND channels, it is convenient to assemble the variances of the entries of $\tilde{\mathbf{H}}$ in a gain matrix $\mathbf{G}$ such that

$$(\mathbf{G})_{i,j} \triangleq E\left[|\tilde{h}_{i,j}|^2\right] \tag{26}$$

where (2) implies that

$$\frac{1}{n_{\mathrm{R}} n_{\mathrm{R}}} \sum_{i=1}^{n_{\mathrm{R}}} \sum_{j=1}^{n_{\mathrm{T}}} (\mathbf{G})_{i,j} = 1. \tag{27}$$

Also useful are the column- and row-wise averages of $\mathbf{G}$, which yield the vectors

$$\mathbf{g}_{\mathrm{R}} = \frac{1}{n_{\mathrm{T}}} \sum_{j=1}^{n_{\mathrm{T}}} (\mathbf{G})_j \tag{28}$$

$$\mathbf{g}_{\mathrm{T}} = \frac{1}{n_{\mathrm{R}}} \sum_{i=1}^{n_{\mathrm{R}}} (\mathbf{G}^T)_i. \tag{29}$$

where we have used $(\cdot)_j$ to denote the jth column of a matrix. The ith entry of $\mathbf{g}_{\mathrm{R}}$ represents the average of the gains between the various transmit antennas and the ith receive antenna while the jth entry of $\mathbf{g}_{\mathrm{T}}$ represents the average of the gains between the jth transmit antenna and the various receive antennas.

Definition 6 *We denote by n'_{R} and n'_{T} the number of nonzero entries in $\mathbf{g}_{\mathrm{R}}$ and $\mathbf{g}_{\mathrm{T}}$, i.e., respectively the number of rows and columns in $\tilde{\mathbf{H}}$ that are not identically zero.*

In an IND channel, n'_{R} and n'_{T} signify the number of receive antennas that capture power and the number of transmit antennas from which power can be received. In a corresponding UIU channel, these map respectively onto the number of nonzero receive and transmit correlation eigenvalues.

IV   Input Optimization

A necessary step in the characterization of the capacity is the optimization of the normalized input covariance, $\boldsymbol{\Phi} = \mathbf{V}\mathbf{P}\mathbf{V}^{\dagger}$, which for channels that are known only statistically at the transmitter cannot depend on $\mathbf{H}$. Henceforth, we shall use the superscript * to distinguish the capacity-achieving value of any quantity that relates to the input covariance.

Theorem 1 *[49, 42] Consider the UIU channel*

$$\mathbf{H} = \mathbf{U}_{\mathrm{R}} \tilde{\mathbf{H}} \mathbf{U}_{\mathrm{T}}^{\dagger} \tag{30}$$

with $\mathbf{U}_{\mathrm{R}}$ and $\mathbf{U}_{\mathrm{T}}$ unitary while $\tilde{\mathbf{H}}$ has independent entries whose marginal distributions are symmetric with respect to zero. The eigenvector matrix of the capacity-achieving input covariance is $\mathbf{V}^{\star} = \mathbf{U}_{\mathrm{T}}$.

A set of necessary and sufficient conditions that characterize the power allocation $\mathbf{P}^{\star} = \mathrm{diag}\{p_j^{\star}\}$, for any arbitrary channel, are given in [49]. A central role in the formulation of these conditions is played by the MMSE on the linear estimation of the signals transmitted along each of the directions determined by the columns of $\mathbf{V}^{\star}$. Deconstructing the input as

$$\mathbf{x} = \mathbf{V}^{\star} \mathbf{P}^{1/2} \mathbf{s} \tag{31}$$

where the information-bearing vector $\mathbf{s} = [s_1, \dots, s_{n_{\mathrm{T}}}]^T$ has IID unit-variance Gaussian entries and defining

$$\mathbf{B} \triangleq \left(\mathbf{I} + \tfrac{\mathsf{SNR}}{n_{\mathrm{T}}} \tilde{\mathbf{H}} \mathbf{P} \tilde{\mathbf{H}}^{\dagger} \right)^{-1}, \tag{32}$$

the MMSE of a linear estimate of s_j from $\mathbf{y}$, conditioned on the channel realization, is

$$\mathsf{MMSE}_j = 1 - p_j \tfrac{\mathsf{SNR}}{n_{\mathrm{T}}} \tilde{\mathbf{h}}_j^{\dagger} \mathbf{B} \tilde{\mathbf{h}}_j. \tag{33}$$

The signal power recovered from the jth signalling direction is given by $1 - \mathsf{MMSE}_j$ and thus the corresponding signal-to-interference-and-noise ratio is

$$\mathsf{SINR}_j = \frac{1 - \mathsf{MMSE}_j}{\mathsf{MMSE}_j} \tag{34}$$

$$= p_j \gamma_j \tag{35}$$

with

$$\gamma_j \triangleq \frac{\tfrac{\mathsf{SNR}}{n_{\mathrm{T}}} \tilde{\mathbf{h}}_j^{\dagger} \mathbf{B} \tilde{\mathbf{h}}_j}{1 - p_j \tfrac{\mathsf{SNR}}{n_{\mathrm{T}}} \tilde{\mathbf{h}}_j^{\dagger} \mathbf{B} \tilde{\mathbf{h}}_j}. \tag{36}$$

The expectations of (33) and (36) with respect to $\tilde{\mathbf{H}}$, in turn, are denoted by

$$\overline{\mathsf{MMSE}}_j \triangleq E[\mathsf{MMSE}_j] \tag{37}$$

$$\overline{\gamma}_j \triangleq E[\gamma_j]. \tag{38}$$

Using the foregoing definitions, the normalized transmit powers that achieve capacity satisfy [49]

$$p_j^\star = 0 \qquad\qquad \overline{\gamma_j} \leq \frac{1}{n_\mathrm{T}} \sum_{\ell=1}^{n_\mathrm{T}} (1 - \overline{\mathrm{MMSE}_\ell}) \qquad (39)$$

$$p_j^\star = \frac{1 - \overline{\mathrm{MMSE}_j}}{\frac{1}{n_\mathrm{T}} \sum_{\ell=1}^{n_\mathrm{T}} (1 - \overline{\mathrm{MMSE}_\ell})} \qquad \overline{\gamma_j} > \frac{1}{n_\mathrm{T}} \sum_{\ell=1}^{n_\mathrm{T}} (1 - \overline{\mathrm{MMSE}_\ell}) \qquad (40)$$

An iterative algorithm to find the power allocation that satisfies these conditions is given in [49]. This allocation simplifies drastically at both low and high SNR:

- For SNR $\to 0$, the entire power budget should be allocated to the maximal-eigenvalue eigenspace of $E[\mathbf{H}^\dagger \mathbf{H}]$ for first- and second-order optimality [47, Theorem 12]. For a UIU channel, this implies that the only nonzero diagonal entry of $\mathbf{P}^\star$ is the one corresponding to the maximal diagonal value of $E[\tilde{\mathbf{H}}^\dagger \tilde{\mathbf{H}}]$. If the multiplicity of such maximal value is plural, the corresponding diagonal entries of $\mathbf{P}^\star$ are equal [47].

- For SNR $\to \infty$, only the diagonal entries of $\mathbf{P}^\star$ that correspond to nonzero diagonal entries of $E[\tilde{\mathbf{H}}^\dagger \tilde{\mathbf{H}}]$ are nonzero. While this suffices to maximize the slope S_∞, minimizing $\mathcal{L}_\infty$ further demands that these nonzero diagonal entries of $\mathbf{P}^\star$ be appropriately adjusted. For $n_\mathrm{T}' \leq n_\mathrm{R}'$, the correct adjustment is to make them equal. For $n_\mathrm{T}' > n_\mathrm{R}'$, this need not be the case [49, 50].

In the special case of separable correlations:

- $\mathbf{V}^\star$ is defined by the eigenspace of the transmit correlation, $\mathbf{\Theta}_\mathrm{T}$, as claimed in [12, 15, 51, 52]. (See also [53, 54] for multi-transmit single-receive channels.)

- The receive correlation, $\mathbf{\Theta}_\mathrm{R}$, enters only the computation of the powers within $\mathbf{P}^\star$. Moreover, at low SNR it plays no role.

- If the correlation takes place only at the receiver, then the capacity-achieving input is isotropic.

In an IND channel, $\mathbf{V}^\star = \mathbf{I}$ and thus each diagonal entry of $\mathbf{P}$ maps directly onto a transmit antenna. Consequently, p_j represents the normalized power allocated to the jth transmit antenna while $(\mathbf{GP})_{i,j}$ represents the normalized power received by the ith receive antenna from the jth transmit antenna.

The capacity-achieving power allocation is uniform for IND Rayleigh-faded channels whose gain matrix $\mathbf{G}$ in (26) is column-regular [49]. As we shall see in the next section (cf. Theorem 2), asymptotically in the numbers of antennas this result extends to any IND channel whose gain matrix is column-regular.

V Asymptotic Analysis

The mutual information in (9) can be rewritten as function of the eigenvalues of $\frac{1}{n_{\mathrm{T}}}\mathbf{H}\mathbf{\Phi}\mathbf{H}^{\dagger}$ via

$$I(\mathsf{SNR}, \mathbf{\Phi}) = \sum_{i=1}^{n_{\mathrm{R}}} \log_2 \left(1 + \mathsf{SNR}\, \lambda_i\big(\tfrac{\mathbf{H}\mathbf{\Phi}\mathbf{H}^{\dagger}}{n_{\mathrm{T}}}\big) \right). \tag{41}$$

In the regime of large numbers of antennas, it is more convenient to operate with the mutual information per receive antenna, sure to remain finite, which converges asymptotically a.s. (almost surely) to

$$\frac{1}{n_{\mathrm{R}}} I(\mathsf{SNR}) \overset{a.s.}{\to} \int \log_2(1 + \mathsf{SNR}\,\xi)\, dF(\xi) \tag{42}$$

where $F(\xi)$ returns the fraction of the eigenvalues $\{\lambda_i\}$ that fall below ξ. Since (42) depends on the channel only through $F(\cdot)$, its study continues to profit from advances in the characterization of the eigenvalues of random matrices (see [38] and references therein).

Although the asymptotic analysis in the remainder of this section deals exclusively with zero-mean channels $\mathbf{H}$, the asymptotic expressions apply also to any nonzero-mean channel $\bar{\mathbf{H}} + \mathbf{H}$ where the deterministic mean matrix $\bar{\mathbf{H}}$ has rank κ satisfying [38, Lemma 2.22]

$$\lim_{n_{\mathrm{R}} \to \infty} \frac{\kappa}{n_{\mathrm{R}}} = 0. \tag{43}$$

In this case, however, the convergence of the capacity to its asymptotic value may be significantly slower than in their zero-mean counterparts [55].

A Asymptotic Analysis: IND Channels

Define the ratio

$$\beta \triangleq \frac{n_{\mathrm{T}}}{n_{\mathrm{R}}} \tag{44}$$

and consider an IND channel $\tilde{\mathbf{H}}$ whose entries are arbitrarily distributed[6] with uniformly bounded variances. Associated with the gain matrix $\mathbf{G}$ in (26), we define a *gain profile* that maps the entries of $\mathbf{G}$ onto a two-dimensional piecewise constant function

$$\mathcal{G}^{(n_{\mathrm{R}})}(r, t) \triangleq (\mathbf{G})_{i,j} \qquad \tfrac{i}{n_{\mathrm{R}}} \le r < \tfrac{i+1}{n_{\mathrm{R}}}, \qquad \tfrac{j}{n_{\mathrm{R}}} \le t < \tfrac{j+1}{n_{\mathrm{R}}} \tag{45}$$

---

[6]In the solutions for the capacity, which involve input optimization, the distribution of each entry must be symmetric with respect to zero. For all other results, this need not be the case.

supported on $r \in (0,1)$ and $t \in (0, \beta)$. We can interpret r and t as normalized receive and transmit antenna indices. Similarly, we define marginal transmit and receive gain profiles by mapping the entries of the vectors $\mathbf{g}_\mathrm{T}$ and $\mathbf{g}_\mathrm{R}$ onto the functions

$$\mathcal{G}_\mathrm{R}^{(n_\mathrm{R})}(r) \triangleq \left(\mathbf{g}_\mathrm{R}^T\right)_i \qquad \frac{i}{n_\mathrm{R}} \leq r < \frac{i+1}{n_\mathrm{R}} \tag{46}$$

$$\mathcal{G}_\mathrm{T}^{(n_\mathrm{R})}(t) \triangleq \left(\mathbf{g}_\mathrm{T}^T\right)_j \qquad \frac{j}{n_\mathrm{R}} \leq t < \frac{j+1}{n_\mathrm{R}} \tag{47}$$

Invoking Theorem 1, we can restrict the input covariance to be diagonal (i.e., $\mathbf{V}^\star = \mathbf{I}$) with the entries along the diagonal of $\mathbf{P}$ uniformly bounded. Associated with $\mathbf{P}$ we define an input *power profile* given by

$$\mathcal{P}^{(n_\mathrm{R})}(t) \triangleq p_j \qquad \frac{j}{n_\mathrm{R}} \leq t < \frac{j+1}{n_\mathrm{R}} \tag{48}$$

and supported on $t \in (0, \beta)$. This profile specifies the power allocation across the transmit array.

We shall assume that, as the number of antennas grows large, the gain profile converges uniformly to a bounded function

$$\mathcal{G}(r,t) \triangleq \lim_{n_\mathrm{R} \to \infty} \mathcal{G}^{(n_\mathrm{R})}(r,t) \tag{49}$$

referred to as asymptotic gain profile. At the same time, the marginal transmit and receive gain profiles in (47) converge to corresponding asymptotic marginal gain profiles, $\mathcal{G}_\mathrm{R}(r)$ and $\mathcal{G}_\mathrm{T}(t)$, which are immediately related to $\mathcal{G}(r,t)$ via

$$\mathcal{G}_\mathrm{R}(r) = E\left[\mathcal{G}(r, \textsf{T})\right] \tag{50}$$

$$\mathcal{G}_\mathrm{T}(t) = E\left[\mathcal{G}(\textsf{R}, t)\right] \tag{51}$$

where the expectations are over the auxiliary random variables $\textsf{R}$ and $\textsf{T}$, independent and uniform on $[0,1]$ and $[0, \beta]$ respectively, which serve as scanning normalized antenna indices. The relationships in (50) and (51) are simply the asymptotic counterparts of (28) and (29).

The power profile is also assumed to converge uniformly to an asymptotic power profile

$$\mathcal{P}(t) \triangleq \lim_{n_\mathrm{R} \to \infty} \mathcal{P}^{(n_\mathrm{R})}(t) \tag{52}$$

The asymptotic capacity-achieving power profile is denoted by $\mathcal{P}^\star(t, \textsf{SNR})$, which in general depends on $\textsf{SNR}$.

Note that (7) and (27) imply that $E[\mathcal{P}(\textsf{T})] = 1$ and $E[\mathcal{G}(\textsf{R}, \textsf{T})] = 1$.

In order to find asymptotic counterparts to n'_T and n'_R, we define the subsets

$$\mathcal{S}_\text{R} \triangleq \{t \in [0,1] : \mathcal{G}_\text{R}(r) \neq 0\} \tag{53}$$

$$\mathcal{S}_\text{T} \triangleq \{r \in [0,\beta] : \mathcal{G}_\text{T}(t) \neq 0 \tag{54}$$

such that $\Pr\{\text{R} \in \mathcal{S}_\text{R}\}$ and $\Pr\{\text{T} \in \mathcal{S}_\text{T}\}$ give, respectively, the fraction of rows and the fraction of columns of $\tilde{\mathbf{H}}$ that have a nonnegligible fraction of nonzero entries. All of our asymptotic results (i.e., whenever we write $n_\text{R} \to \infty$) require that these two probabilities be nonzero. We can then further define

$$\beta' = \beta \frac{\Pr\{\text{T} \in \mathcal{S}_\text{T}\}}{\Pr\{\text{R} \in \mathcal{S}_\text{R}\}} \tag{55}$$

satisfying $0 < \beta' < \infty$.

A.1   MMSE, SINR and Input Optimization

Since, as shown in Section IV, the MMSE plays a central role in defining the capacity-achieving power allocation, such is the starting point of our asymptotic analysis of IND channels. More precisely, we begin with an asymptotic characterization of the SINR, from which the MMSE follows. Recall, from (35) and (36), the SINR exhibited by the signal radiated from the jth transmit antenna at the output of a linear MMSE receiver. Invoking the matrix inversion lemma, such SINR is equivalent to

$$\mathsf{SINR}_j = \tfrac{\mathsf{SNR}}{n_\text{T}} p_j \tilde{\mathbf{h}}_j^\dagger \left(\mathbf{I} + \tfrac{\mathsf{SNR}}{n_\text{R}} \sum_{\ell \neq j} p_\ell \tilde{\mathbf{h}}_\ell \tilde{\mathbf{h}}_\ell^\dagger \right)^{-1} \tilde{\mathbf{h}}_j. \tag{56}$$

Associated with SINR_j we can define the sequence, indexed by n_R, of functions

$$\Gamma^{(n_\text{R})}(t, \mathsf{SNR}) \triangleq \frac{\mathsf{SINR}_j}{\mathsf{SNR}}, \qquad \tfrac{j}{n_\text{R}} \leq t < \tfrac{j+1}{n_\text{R}} \tag{57}$$

Lemma 1 *Let $\tilde{\mathbf{H}}$ be IND with asymptotic gain profile $\mathcal{G}(r,t)$ and let the asymptotic power profile of the input be $\mathcal{P}(t)$. As $n_\text{R} \to \infty$, $\Gamma^{(n_\text{R})}(t, \mathsf{SNR})$ converges a.s. to $\Gamma(t, \mathsf{SNR})$ satisfying*

$$\Gamma(t, \mathsf{SNR}) = \frac{\mathcal{P}(t)}{\beta} E\left[\frac{\mathcal{G}(\text{R}, t)}{1 + \mathsf{SNR}\, E\left[\frac{\mathcal{G}(\text{R},\text{T})\mathcal{P}(\text{T})}{1+\mathsf{SNR}\,\Gamma(\text{T},\mathsf{SNR})} \,\middle|\, \text{R} \right]} \right] \tag{58}$$

where the expectations are over R and T, independent and respectively uniform on $[0,1]$ and $[0,\beta]$.

Proof: *See Appendix B.*

In the case that $\mathbf{G}$ is row-regular, (58) reverts to the well-known Tse-Hanly formula [44] and hence Lemma 1 provides a generalization thereof.

Analogous to $\Gamma^{(n_\mathrm{R})}(\cdot, \mathsf{SNR})$, we can define a sequence of functions associated with the MMSE as

$$\Upsilon^{(n_\mathrm{R})}(t, \mathsf{SNR}) \triangleq \mathsf{SNR}\ \mathsf{MMSE}_j, \qquad \tfrac{j}{n_\mathrm{R}} \le t < \tfrac{j+1}{n_\mathrm{R}} \tag{59}$$

which converges to the function

$$\Upsilon(t, \mathsf{SNR}) = \frac{\mathsf{SNR}}{1 + \mathsf{SNR}\,\Gamma(t, \mathsf{SNR})}. \tag{60}$$

A related quantity whose asymptotic behavior is of interest is the matrix $\mathbf{B}$ defined in (32), which appears in the expressions of the SINR and the MMSE. Specifically, the significance lies on the diagonal entries of $\mathbf{B}$ whose average equals

$$\frac{1}{n_\mathrm{T}} \sum_{i=1}^{n_\mathrm{R}} (\mathbf{B})_{i,i} = \frac{1}{n_\mathrm{T}} \sum_{j=1}^{n_\mathrm{T}} \mathsf{MMSE}_j + \frac{n_\mathrm{R}}{n_\mathrm{T}} - 1. \tag{61}$$

The sequence of functions

$$\mathcal{B}^{(n_\mathrm{R})}(r, \mathsf{SNR}) \triangleq (\mathbf{B})_{i,i} \qquad \tfrac{i}{n_\mathrm{R}} \le r < \tfrac{i+1}{n_\mathrm{R}} \tag{62}$$

converges a.s. to $\mathcal{B}(r, \mathsf{SNR})$, which is the solution to (cf. Theorem 6)

$$\mathcal{B}(r, \mathsf{SNR}) = \frac{1}{1 + \mathsf{SNR}\beta\,E\left[\dfrac{\mathcal{G}(r, \mathsf{T})\mathcal{P}(\mathsf{T})}{\beta + \mathsf{SNR}\,E[\mathcal{B}(\mathsf{R}, \mathsf{SNR})\mathcal{G}(\mathsf{R}, \mathsf{T})\mathcal{P}(\mathsf{T})|\mathsf{T}]}\right]} \tag{63}$$

Comparing (58) and (63), respectively satisfied by $\Gamma(\cdot, \mathsf{SNR})$ and $\mathcal{B}(\cdot, \mathsf{SNR})$, it is inferred that these functions relate through

$$\Gamma(t, \mathsf{SNR}) = \frac{\mathcal{P}(t)}{\beta}\,E[\mathcal{G}(\mathsf{R}, t)\mathcal{B}(\mathsf{R}, \mathsf{SNR})]. \tag{64}$$

Using $\Gamma(\cdot, \mathsf{SNR})$, $\Upsilon(\cdot, \mathsf{SNR})$ and $\mathcal{B}(\cdot, \mathsf{SNR})$, we can now formulate the asymptotic counterparts of (39) and (40) and obtain the conditions that the asymptotic power profile must satisfy to achieve capacity. As detailed in Appendix D, the corresponding $\mathcal{P}^\star(\cdot, \mathsf{SNR})$ satisfies

$$\mathcal{P}^\star(t, \mathsf{SNR}) = 0 \qquad\qquad t \notin \mathcal{T}_0 \tag{65}$$

$$\mathcal{P}^\star(t, \mathsf{SNR}) = \frac{\Gamma(t, \mathsf{SNR})\Upsilon(t, \mathsf{SNR})}{1 - \frac{1}{\mathsf{SNR}}E[\Upsilon(\mathsf{T}, \mathsf{SNR})]} \qquad t \in \mathcal{T}_0 \tag{66}$$

where

$$\mathcal{T}_0 \triangleq \left\{ t \in [0, \beta] : \frac{\mathsf{SNR}}{\beta}E[\mathcal{G}(\mathsf{R}, t)\mathcal{B}(\mathsf{R}, \mathsf{SNR})] > 1 - \frac{1}{\mathsf{SNR}}E[\Upsilon(\mathsf{T}, \mathsf{SNR})] \right\} \tag{67}$$

The measure of this subset does not vanish at any SNR > 0 if $\beta' > 0$, i.e., $\Pr\{\textsc{t} \in \mathcal{T}_0\} \neq 0$ if $\beta' > 0$.

For a certain class of channels of interest, the asymptotic capacity-achieving power profile simplifies drastically. Specifically, the isotropy of the capacity-achieving input for IND Rayleigh-faded channels whose gain matrix is column regular [49] extends asymptotically to any IND channel.

Theorem 2 *Consider an IND channel $\tilde{\mathbf{H}}$ whose gain matrix $\mathbf{G}$ is column-regular. The asymptotic capacity-achieving input is isotropic, i.e., $\mathcal{P}^\star(t, \textsc{snr}) = 1$ at every SNR.*

Proof: *See Appendix 2.*

A.2   Capacity

The following asymptotic characterization of the mutual information, which capitalizes on Girko's theorem (cf. Appendix B), is one of the pivotal results in the paper.

Theorem 3 *Let $\tilde{\mathbf{H}}$ be IND with asymptotic gain profile $\mathcal{G}(r, t)$ and let the asymptotic power profile be $\mathcal{P}(t)$. As $n_{\mathrm{R}} \to \infty$,*

$$\frac{1}{n_{\mathrm{R}}} I(\textsc{snr}) \overset{a.s.}{\to} \beta \, E\left[\log_2(1 + \textsc{snr}\,\Gamma(\textsc{t}, \textsc{snr}))\right] + E\left[\log_2(1 + E[\mathcal{G}(\textsc{r}, \textsc{t})\mathcal{P}(\textsc{t})\Upsilon(\textsc{t}, \textsc{snr})|\textsc{r}])\right]$$
$$-\beta \, E\left[\Gamma(\textsc{t}, \textsc{snr})\Upsilon(\textsc{t}, \textsc{snr})\right] \log_2 e \tag{68}$$

with expectation over $\textsc{r}$ and $\textsc{t}$, independent and respectively uniform on $[0, 1]$ and $[0, \beta]$, and with $\Gamma(\cdot, \textsc{snr})$ and $\Upsilon(\cdot, \textsc{snr})$ defined in (58) and (60).

Proof: *See Appendix C.*

The asymptotic capacity is given by the right-hand side of (68) if the arbitrary $\mathcal{P}(t)$ is replaced by $\mathcal{P}^\star(t, \textsc{snr})$. The asymptotic spectral efficiency achievable with a uniform power allocation can also be obtained from (68) simply by setting $\mathcal{P}(t) = 1$.

Proposition 1 *For any channel with IND entries satisfying the normalization in (27), the asymptotic spectral efficiency achieved by a uniform power allocation is upper-bounded by the value it takes when the channel is IID.*

Proof: *See Appendix E.*

As an alternative to (68), the asymptotic mutual information can be posed as function of $\mathcal{B}(\cdot, \textsc{snr})$. Plugging (64) into (68),

$$\frac{1}{n_{\mathrm{R}}} I(\textsc{snr}) \overset{a.s.}{\to} \beta \, E\left[\log_2\left(1 + \frac{\textsc{snr}}{\beta} E[\mathcal{G}(\textsc{r}, \textsc{t})\mathcal{P}(\textsc{t})\mathcal{B}(\textsc{r}, \textsc{snr})|\textsc{t}]\right)\right] - E\left[\log_2\left(\mathcal{B}(\textsc{r}, \textsc{snr})\right)\right]$$
$$+\left(E\left[\mathcal{B}(\textsc{r}, \textsc{snr})\right] - 1\right) \log_2 e. \tag{69}$$

Although (68) and (69) are equivalent, they provide different interpretations. The former, in particular, appears as functional of $\Gamma(\cdot, \textsc{snr})$ and $\Upsilon(\cdot, \textsc{snr})$, which are not only functions with direct engineering significance (SINR and MMSE respectively) but also sufficient performance measures for a linear MMSE receiver. In fact, the first term on the right-hand side of (68) is nothing but the asymptotic mutual information achievable with a linear MMSE receiver and thus the remaining terms quantify the additional mutual information achievable with an optimum receiver. Remarkably, this additional mutual information is itself a functional of the SINR and MMSE and hence the asymptotic mutual information of an optimum nonlinear receiver is completely determined by the performance measures of a linear MMSE receiver. Additionally, as we saw in Section IV, the power allocation that achieves capacity can also be formulated solely on the basis of the MMSE. Far from coincidental, these various dependences are manifestations of a fundamental relationship between the mutual information and the MMSE [56].

Theorem 3 adopts relevant special forms whenever $\mathbf{G}$ exhibits certain structures. A common such structure occurs when $\mathbf{G}$ is given by the outer product of two vectors, in which case $\mathcal{G}(r,t) = \mathcal{G}_{\textsc{R}}(r)\mathcal{G}_{\textsc{T}}(t)$.

Corollary 1 *Given an IND channel with $\mathcal{G}(r,t) = \mathcal{G}_{\textsc{R}}(r)\mathcal{G}_{\textsc{T}}(t)$,*

$$\frac{1}{n_{\textsc{R}}}I(\textsc{snr}) \overset{a.s.}{\to} \beta\, E\left[\log\left(1 + \textsc{snr}\, \mathcal{G}_{\textsc{T}}(\textsc{t})\mathcal{P}(\textsc{t})\hat{\Gamma}(\textsc{snr})\right)\right] + E\left[\log\left(1 + \textsc{snr}\, \mathcal{G}_{\textsc{R}}(\textsc{r})\hat{\Upsilon}(\textsc{snr})\right)\right]$$
$$-\beta\, \textsc{snr}\, \hat{\Gamma}(\textsc{snr})\hat{\Upsilon}(\textsc{snr})\log_2 e \tag{70}$$

where

$$\hat{\Gamma}(\textsc{snr}) \;=\; \frac{1}{\beta}E\left[\frac{\mathcal{G}_{\textsc{R}}(\textsc{r})}{1 + \textsc{snr}\, \mathcal{G}_{\textsc{R}}(\textsc{r})\hat{\Upsilon}(\textsc{snr})}\right] \tag{71}$$

$$\hat{\Upsilon}(\textsc{snr}) \;=\; E\left[\frac{\mathcal{G}_{\textsc{T}}(\textsc{t})\mathcal{P}(\textsc{t})}{1 + \textsc{snr}\, \mathcal{G}_{\textsc{T}}(\textsc{t})\mathcal{P}(\textsc{t})\hat{\Gamma}(\textsc{snr})}\right] \tag{72}$$

Proof: *See Appendix H.*

In this case, the SINR at the output of a linear MMSE receiver experienced by the signal transmitted from the jth transmit antenna corresponds with the function

$$\textsc{snr}\, \hat{\Gamma}(\textsc{snr})\mathcal{G}_{\textsc{T}}(t)\mathcal{P}(t) \qquad\qquad \frac{j}{n_{\textsc{R}}} \leq t < \frac{j+1}{n_{\textsc{R}}} \tag{73}$$

whereas the MMSE corresponds with

$$\frac{1}{1 + \textsc{snr}\, \hat{\Gamma}(\textsc{snr})\mathcal{G}_{\textsc{T}}(t)\mathcal{P}(t)} \qquad\qquad \frac{j}{n_{\textsc{R}}} \leq t < \frac{j+1}{n_{\textsc{R}}} \tag{74}$$

Besides the outer product of two vectors, other relevant structures for $\mathbf{G}$ that are created by many arrays of practical interest are those of regular matrices (cf. Definitions 4–5).

If $\mathbf{G}$ is row-regular, then Theorem 3 reduces to [33, Theorem 4.1] which can be obtained directly through the Silverstein-Bai framework.

As stated by Theorem 2, for any column-regular $\mathbf{G}$ the asymptotic capacity-achieving power allocation is uniform and the capacity emerges in a familiar explicit form.

Theorem 4 *Consider an IND channel whose gain matrix $\mathbf{G}$ is column-regular and asymptotically mean row-regular. The asymptotic capacity coincides with that of a channel with unit-variance independent entries, given by [3] [7]*

$$\frac{1}{n_{\mathrm{R}}} C \overset{a.s.}{\to} \log_2 \left(1 + \mathsf{SNR} - \tfrac{1}{4}\mathcal{F}(\beta, \tfrac{\mathsf{SNR}}{\beta}) \right)$$
$$+ \beta \log_2 \left(1 + \tfrac{\mathsf{SNR}}{\beta} - \tfrac{1}{4}\mathcal{F}(\beta, \tfrac{\mathsf{SNR}}{\beta}) \right) - \beta \frac{\log_2 e}{4\,\mathsf{SNR}} \mathcal{F}(\beta, \tfrac{\mathsf{SNR}}{\beta}) \tag{75}$$

with the auxiliary function

$$\mathcal{F}(x, y) \triangleq \left(\sqrt{1 + y(1 + \sqrt{x})^2} - \sqrt{1 + y(1 - \sqrt{x})^2} \right)^2. \tag{76}$$

Proof: *See Appendix F.*

Theorem 4 states that the asymptotic capacity expressions derived for IID channels [3, 4] through the Marčenko-Pastur law are, in fact, valid for any IND channel whose gain matrix $\mathbf{G}$ is column-regular and asymptotically mean row-regular. Moreover, the stricter column-regular condition is required solely in order to ensure that the capacity-achieving input of the IND channel is isotropic. By specifying a uniform power allocation, we can relax this condition as follows.

Theorem 5 *Consider a channel with IND entries whose gain matrix $\mathbf{G}$ is asymptotically mean doubly-regular. The asymptotic spectral efficiency achieved by an isotropic input coincides with that of a channel with unit-variance independent entries given in (75).*

Proof: *See Appendix F.*

In order to compare regular channels to which Theorems 4 and 5 apply, it is important to remember that, while the function that relates their capacity/spectral efficiency with

---

[7]Eq. (75) was originally derived for IID matrices in the context of CDMA [3]. As presented here, it has been modified to conform with the SNR defined in this paper.

the receive signal-to-noise ratio may be common, the corresponding *transmit* signal-to-noise ratios may differ if the average channel gain g in (4) is not the same. Any difference in average channel gain translates directly into a dB-shift in the curve relating capacity/spectral efficiency and transmit power, even while $I(\text{SNR})$ may remain unchanged. A more detailed discussion of this issue can be found in Section VII.

Example 1 *Consider a transmitter and a receiver each with the antennas evenly split between two orthogonal polarizations. Denoting by $\mathcal{X}$ the cross-polar discrimination (gain between crosspolar antennas relative to gain between copolar antennas), $\mathbf{G}$ equals*

$$\mathbf{G} = \frac{2}{1+\mathcal{X}} \begin{bmatrix} 1 & \mathcal{X} & 1 & \mathcal{X} & \cdots & \mathcal{X} \\ \mathcal{X} & 1 & \mathcal{X} & 1 & \cdots & 1 \\ 1 & \mathcal{X} & 1 & \mathcal{X} & \cdots & \mathcal{X} \\ \mathcal{X} & 1 & \mathcal{X} & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{X} & 1 & \mathcal{X} & 1 & \cdots & 1 \end{bmatrix} \tag{77}$$

which is column- and row-regular. From Theorem 4, in the absence of antenna correlation its asymptotic $C(\text{SNR})$ function equals that of an IID channel where all the antennas are copolar. However, the amount of transmit power required the attain a certain capacity differs by $\frac{2}{1+\mathcal{X}}\big|_{dB}$ from the amount it would take in an IID channel.

A.3 Performance Measures: $\mathcal{L}_\infty$

The study of $\frac{E_b}{N_0}{}_{\min}$, S_0 and S_∞ for IND channels is deferred to Section VI, where exact expressions for arbitrary numbers of antennas are derived. In contrast, the characterization of $\mathcal{L}_\infty$ is asymptotic in the number of antennas and thus we elaborate it here. Recall from Section IV that, for $\text{SNR} \to \infty$, $\mathcal{P}^\star(t, \text{SNR})$ becomes strictly positive for t on the subset $\mathcal{S}_\text{T}$ defined in (54). Furthermore, for $\beta' \le 1$ it becomes uniform on that subset. For $\beta' > 1$, this may not be the case and we use $\mathcal{P}_\infty^\star(t)$ to denote the limiting capacity-achieving power allocation, i.e.,

$$\mathcal{P}_\infty^\star(t) \triangleq \lim_{\text{SNR}\to\infty} \mathcal{P}^\star(t, \text{SNR}) \tag{78}$$

Proposition 2 *Let $\tilde{\mathbf{H}}$ be an IND channel with asymptotic gain profile $\mathcal{G}(r,t)$ whose marginals are $\mathcal{G}_\text{R}(r)$ and $\mathcal{G}_\text{T}(t)$. As $n_\text{R} \to \infty$,*

$$\mathcal{L}_\infty \overset{a.s.}{\to} \begin{cases} -E\left[\log_2\left(\frac{\Pr\{\text{T}\in\mathcal{S}_\text{T}\}}{e} E\left[\frac{\mathcal{G}(\text{R}',\text{T}')\mathcal{P}_\infty^\star(\text{T}')}{1+\alpha(\text{T}')}\big|\text{R}'\right]\right)\right] - \beta' E\left[\log_2\left(1+\alpha(\text{T}')\right)\right] & \beta' > 1 \\[2ex] -E\left[\log_2\frac{\mathcal{G}(\text{R}',\text{T}')}{e}\right] & \beta' = 1 \\[2ex] -E\left[\log_2\frac{\Gamma_\infty(\text{T}')}{e}\right] - \frac{1}{\beta'}E\left[\log_2\left(1+E\left[\frac{\mathcal{G}(\text{R}',\text{T}')}{\Gamma_\infty(\text{T}')}\big|\text{R}'\right]\right)\right] & \beta' < 1 \end{cases}$$

with R′ and T′ the restrictions of R and T to the subsets $\mathcal{S}_{\mathrm{R}}$ and $\mathcal{S}_{\mathrm{T}}$, respectively. For $\beta' > 1$, the function $\alpha(t) \triangleq \lim\limits_{\mathrm{SNR}\to\infty} \mathrm{SNR}\,\Gamma(t, \mathrm{SNR})$ is the solution of

$$\alpha(t) = \frac{1}{\beta'} E\left[\frac{\mathcal{G}(\mathrm{R}',t)\mathcal{P}_\infty^\star(t)}{E\left[\frac{\mathcal{G}(\mathrm{R}',\mathrm{T}')\mathcal{P}_\infty^\star(\mathrm{T}')}{1+\alpha(\mathrm{T}')}\Big|\mathrm{R}'\right]}\right] \tag{79}$$

whereas, for $\beta' < 1$, $\Gamma_\infty(t) \triangleq \lim\limits_{\mathrm{SNR}\to\infty} \Gamma(t, \mathrm{SNR})$ is the solution of

$$E\left[\frac{1}{1 + E\left[\frac{\mathcal{G}(\mathrm{R}',\mathrm{T}')}{\Gamma_\infty(\mathrm{T}')}\Big|\mathrm{R}'\right]}\right] = 1 - \beta'. \tag{80}$$

Additionally, for $\beta' > 1$

$$\mathcal{P}_\infty^\star(t) = \frac{\frac{\alpha(t)}{1+\alpha(t)}}{E\left[\frac{\alpha(\mathrm{T})}{1+\alpha(\mathrm{T})}\right]} \tag{81}$$

Proof: *See Appendix I.*

Notice that $\alpha(t)$ indicates, for $\frac{j}{n_{\mathrm{R}}} \leq t < \frac{j+1}{n_{\mathrm{R}}}$, the limiting SINR, approached as the SNR grows large, of the signal radiated from the jth transmit antenna. In turn, $\Gamma_\infty(t)$ indicates the limiting ratio between the SINR and the SNR. Except outside $\mathcal{S}_{\mathrm{T}}(t)$, where both $\alpha(t)$ and $\Gamma_\infty(t)$ are zero, the behavior of these two quantities depends very strongly on β'.

- If $\beta' \leq 1$, then $\alpha(t)$ is not finite and thus the SINR of every signal reaching the receiver grows without bound for SNR $\to \infty$. This simply reflects the ability of a linear MMSE receiver to become a perfect decorrelator whenever the number of dimensions exceeds the number of interfering signals and the noise is negligible. At the same time, $\Gamma_\infty(t)$ is nonzero and thus the SINR growth is linear with the SNR.

- If $\beta' > 1$, then $\alpha(t)$ is bounded and thus the SINR saturates for SNR $\to \infty$. Perfect decorrelation is no longer possible once the number of signals overwhelms the dimensionality of the receiver. Naturally then, $\Gamma_\infty(t) = 0$.

With a uniform power allocation over *all* transmit antennas, $\mathcal{L}_\infty$ is further increased, for $\beta' \leq 1$, by $-\log_2\left(\Pr\{\mathrm{T} \in \mathcal{S}_{\mathrm{T}}\}\right)$, reflecting the power wasted on transmit antennas whose signals do not reach the receiver. For $\beta' > 1$, $\mathcal{L}_\infty$ is also bound to increase if the power allocation is uniform over all transmit antennas, in this case by an amount that depends on $\mathcal{P}_\infty(\cdot)$ but that is no smaller than $-\log_2\left(\Pr\{\mathrm{T} \in \mathcal{S}_{\mathrm{T}}\}\right)$.

Corollary 2 *If $\mathcal{G}(r,t) = \mathcal{G}_{\mathrm{R}}(r)\mathcal{G}_{\mathrm{T}}(t)$, Proposition 2 particularizes to*

$$\mathcal{L}_{\infty} \overset{a.s.}{\to} \begin{cases} -E\left[\log_2\left(\frac{\Pr\{\mathsf{T}\in\mathcal{S}_{\mathrm{T}}\}}{e}\frac{\mathcal{G}_{\mathrm{R}}(\mathsf{R}')}{\hat{\alpha}\,\beta'}\right)\right] - \beta' E\left[\log_2\left(1 + \hat{\alpha}\,\mathcal{G}_{\mathrm{T}}(\mathsf{T}')\mathcal{P}_{\infty}^{\star}(\mathsf{T}')\right)\right] & \beta' > 1 \\[2ex] -E\left[\log_2 \frac{\mathcal{G}_{\mathrm{T}}(\mathsf{T}')\mathcal{G}_{\mathrm{R}}(\mathsf{R}')}{e}\right] & \beta' = 1 \\[2ex] -E\left[\log_2 \frac{\hat{\Gamma}_{\infty}\,\mathcal{G}_{\mathrm{T}}(\mathsf{T}')}{e}\right] - \frac{1}{\beta'}E\left[\log_2\left(1 + \frac{\mathcal{G}_{\mathrm{R}}(\mathsf{R}')}{\hat{\Gamma}_{\infty}}\right)\right] & \beta' < 1 \end{cases}$$

with $\hat{\alpha}$ and $\hat{\Gamma}_{\infty}$ solutions, respectively, to

$$E\left[\frac{1}{1 + \hat{\alpha}\,\mathcal{G}_{\mathrm{T}}(\mathsf{T}')\mathcal{P}_{\infty}^{\star}(\mathsf{T}')}\right] = 1 - \frac{1}{\beta'} \tag{82}$$

$$E\left[\frac{1}{1 + \frac{\mathcal{G}_{\mathrm{R}}(\mathsf{R}')}{\hat{\Gamma}_{\infty}}}\right] = 1 - \beta'. \tag{83}$$

For $\beta' = 1$, this result had been derived in [10]. If $\mathcal{G}_{\mathrm{R}}(r) = 1$ and we set $\mathcal{P}_{\infty}(t) = 1$, the expressions for $\mathcal{L}_{\infty}$ further simplify to those reported in [33]. Finally, if $\mathcal{G}(r,t)$ is mean doubly-regular and $\mathcal{P}_{\infty}(t) = 1$, then $\hat{\alpha}$ and $\hat{\Gamma}_{\infty}$ can be solved for and the high-SNR power offset specializes to

$$\mathcal{L}_{\infty} \overset{a.s.}{\to} \begin{cases} (\beta - 1)\log_2\frac{\beta-1}{\beta} + \log_2 e & \beta > 1 \\[2ex] \log_2 e & \beta = 1 \\[2ex] \frac{1-\beta}{\beta}\log_2(1-\beta) + \log_2(\beta e) & \beta < 1 \end{cases} \tag{84}$$

previously derived for IID channels [31] and hereby generalized to mean doubly-regular IND channels. For $\beta' \leq 1$, the expressions in (84) also serve as a lower bound, i.e., for any IND channel with $\beta' \leq 1$

$$\mathcal{L}_{\infty} \geq \mathcal{L}_{\infty} \quad . \tag{85}$$

B Asymptotic Analysis: UIU Channels

Through the unitary equivalence between IND and UIU channels, the asymptotic expressions for mutual information, capacity and $\mathcal{L}_{\infty}$ derived thus far for the former are immediately applicable to the latter. Precisely, any channel with correlation $R_{\mathbf{H}}$ that falls within the UIU framework and has bounded eigenvalues is unitarily equivalent to an IND channel $\tilde{\mathbf{H}}$ with variances $(\mathbf{G})_{i,j} = \lambda_{i,j}(R_{\mathbf{H}})$.

For Rayleigh-faded channels with separable correlations, cf. (3), the unitary relationship can be further elaborated on. Specifically, since for the equivalent IND channel $(\mathbf{G})_{i,j} =$

$\lambda_i(\boldsymbol{\Theta}_{\mathrm{R}})\lambda_j(\boldsymbol{\Theta}_{\mathrm{T}})$, the asymptotic gain profile of $\tilde{\mathbf{H}}$ then factors as $\mathcal{G}(r,t) = \mathcal{G}_{\mathrm{R}}(r)\mathcal{G}_{\mathrm{T}}(t)$ and Corollary 1 yields

$$
\begin{aligned}
\frac{1}{n_{\mathrm{R}}}I(\textsf{SNR}) \overset{a.s.}{\to} \;& \beta\,E\left[\log_2(1 + \textsf{SNR}\,\Lambda_{\mathrm{T}}\hat{\Gamma}(\textsf{SNR}))\right] + E\left[\log_2(1 + \textsf{SNR}\,\Lambda_{\mathrm{R}}\hat{\Upsilon}(\textsf{SNR}))\right] \\
& -\beta\,\textsf{SNR}\,\hat{\Gamma}(\textsf{SNR})\hat{\Upsilon}(\textsf{SNR})\log_2 e
\end{aligned}
\tag{86}
$$

with expectation over Λ_{R} and Λ_{T} distributed according to the asymptotic eigenvalue distributions of $\boldsymbol{\Theta}_{\mathrm{R}}$ and $\boldsymbol{\Theta}_{\mathrm{T}}^{1/2}\boldsymbol{\Phi}\boldsymbol{\Theta}_{\mathrm{T}}^{1/2}$, respectively, and with

$$
\hat{\Gamma}(\textsf{SNR}) = \frac{1}{\beta}E\left[\frac{\Lambda_{\mathrm{R}}}{1 + \Lambda_{\mathrm{R}}\,\textsf{SNR}\,\hat{\Upsilon}(\textsf{SNR})}\right]
\tag{87}
$$

$$
\hat{\Upsilon}(\textsf{SNR}) = E\left[\frac{\Lambda_{\mathrm{T}}}{1 + \Lambda_{\mathrm{T}}\textsf{SNR}\,\hat{\Gamma}(\textsf{SNR})}\right]
\tag{88}
$$

The capacity is given by the right-hand side of (86) when the eigenvectors of $\boldsymbol{\Phi}$ equal those of $\boldsymbol{\Theta}_{\mathrm{T}}$ while the power allocation equals $\mathbf{P}^{\star}$, whose diagonal entries satisfy (39) and (40).

Corollary 2 can be similarly applied to UIU channels taking into account that β' indicates the ratio between the number of nonzero eigenvalues in $\boldsymbol{\Theta}_{\mathrm{T}}$ and $\boldsymbol{\Theta}_{\mathrm{R}}$.

It must be noted that, for Rayleigh-faded channels with separable correlations, some asymptotic characterizations are already available in the literature [10, 11, 14].[8] As it turns out, if $\boldsymbol{\Theta}_{\mathrm{R}}$ and $\boldsymbol{\Theta}_{\mathrm{T}}$ are Toeplitz matrices, then (86) holds for any channel described by (3) even when the entries of $\mathbf{H}_w$ are IID but their fading is not Rayleigh [38, Theorem 2.44].

VI   Arbitrary Numbers of Antennas

Having analyzed the large-dimensional regime in Section V, we now turn our attention to the realm of arbitrary numbers of antennas, for which we derive exact expressions for $\frac{E_b}{N_0}_{\min}$, S_0 and S_∞. In addition, this section also illustrates the power of the asymptotic expressions in approximating their nonasymptotic counterparts even when n_{T} and n_{R} are small.

---

[8]The analysis in [10] is limited to square channels ($\beta' = 1$). Arbitrary β' is considered in [11], which relies on the (not fully rigorized yet) replica method of statistical mechanics. In [14], an explicit expression is given for a specifically crafted form of one-sided correlation (either transmit or receive only).

A Arbitrary Numbers of Antennas: IND Channels

Invoking Theorem 3, the capacity of an IND channel is approximated by

$$C(\mathsf{SNR}) \approx \sum_{j=1}^{n_\mathrm{T}} \log_2\left(\frac{1 + \mathsf{SNR}\,\Gamma_j}{e^{\Gamma_j \Upsilon_j}}\right) + \sum_{i=1}^{n_\mathrm{R}} \log_2\left(1 + \tfrac{1}{n_\mathrm{T}} \sum_{j=1}^{n_\mathrm{T}} p_j^\star (\mathbf{G})_{i,j} \Upsilon_j\right) \tag{89}$$

where

$$\Gamma_j = \sum_{i=1}^{n_\mathrm{R}} \frac{p_j^\star (\mathbf{G})_{i,j}}{n_\mathrm{T} + \sum_{j=1}^{n_\mathrm{T}} p_j^\star (\mathbf{G})_{i,j} \Upsilon_j} \tag{90}$$

$$\Upsilon_j = \frac{\mathsf{SNR}}{1 + \mathsf{SNR}\,\Gamma_j} \tag{91}$$

with the powers $p_j^\star$ satisfying (39) and (40). The spectral efficiency with uniform power allocation is also approximated by (89) by simply setting $p_j = 1$ for every j. The SINR exhibited by the signal radiated from the jth transmit antenna at the output of a linear MMSE receiver is approximated by $\mathsf{SNR}\,\Gamma_j$ while the corresponding MMSE is approximated by Υ_j/SNR.

For the results that follow, we recall that $\mathbf{g}_\mathrm{R}$ and $\mathbf{g}_\mathrm{T}$ are the column- and row-wise averages of the gain matrix $\mathbf{G}$ while n_R' and n_T' are the number of nonzero entries in $\mathbf{g}_\mathrm{R}$ and $\mathbf{g}_\mathrm{T}$, respectively.

Proposition 3 *The low-SNR performance measures of an IND channel are*

$$\frac{E_b}{N_0}_{\min} = \frac{\log_e 2}{g\, n_\mathrm{R} \max(\mathbf{g}_\mathrm{T})} \tag{92}$$

$$S_0 = \frac{2 n_\mathrm{R}^2 m^2 \max^2(\mathbf{g}_\mathrm{T})}{n_\mathrm{R}^2 m \max^2(\mathbf{g}_\mathrm{T}) + \|\sum_{j \in \mathcal{M}} (\mathbf{G})_j\|^2} \tag{93}$$

where $\max(\mathbf{g}_\mathrm{T})$ denotes the largest entry of $\mathbf{g}_\mathrm{T}$, $\mathcal{M}$ is the set of indices of the entries attaining it and $m = |\mathcal{M}|$.

Proof: *Apply [47, Theorems 1 and 13].*

It can be verified that

$$\frac{\log_e 2}{g\, n_\mathrm{T} n_\mathrm{R}} \leq \frac{E_b}{N_0}_{\min} \leq \frac{\log_e 2}{g\, n_\mathrm{R}} \tag{94}$$

with the upper bound corresponding to an IID channel [47]. On the other hand,

$$1 \leq S_0 \leq \frac{2 n_\mathrm{R} m}{n_\mathrm{R} + m} \leq \frac{2 n_\mathrm{T} n_\mathrm{R}}{n_\mathrm{T} + n_\mathrm{R}} \tag{95}$$

where the rightmost expression corresponds to an IID channel [47].

If, in lieu of the optimum power allocation in Proposition 3, the power allocation is uniform, then

$$\frac{E_b}{N_0}\bigg|_{\min} = \frac{\log_e 2}{g\, n_{\mathrm{R}}} \tag{96}$$

$$S_0 = \frac{2n_{\mathrm{T}}^2 n_{\mathrm{R}}^2}{n_{\mathrm{T}}^2 \|\mathbf{g}_{\mathrm{R}}\|^2 + n_{\mathrm{R}}^2 \|\mathbf{g}_{\mathrm{T}}\|^2}. \tag{97}$$

Although S_0 is lower on an IND channel than on an IID channel, $\frac{E_b}{N_0}\big|_{\min}$ is also lower and thus the low-SNR capacity of an IND channel exceeds that of an IID channel.

Proposition 4 *The high-*SNR *slope of an IND channel is*

$$S_\infty = \min(n_{\mathrm{T}}', n_{\mathrm{R}}'). \tag{98}$$

Proof: *See Appendix G.*

The high-SNR slope in (98) can be attained with a uniform power allocation. Clearly, $1 \leq S_\infty \leq \min(n_{\mathrm{T}}, n_{\mathrm{R}})$ with the upper bound corresponding to an IID channel. Since, in most practical instances, all antennas contribute positive power, this upper bound is usually achieved bolstering the importance of $\mathcal{L}_\infty$ in the assessment of the effects of antenna correlation at high SNR.

Denote by $\mathbf{G}'$ and $\mathbf{P}'$ the submatrices obtained by excluding from $\mathbf{G}$ and $\mathbf{P}$ the rows and columns that are identically zero. In the high-SNR limit, the capacity-achieving power allocation is uniform for $n_{\mathrm{T}}' \leq n_{\mathrm{R}}'$ while, for $n_{\mathrm{T}}' > n_{\mathrm{R}}'$, we denote it by $\mathbf{P}'^\star = \mathrm{diag}\{p_j'^\star\}$ whose entries satisfy (40). With that, $\mathcal{L}_\infty$ in 3-dB units is, from Proposition 2,

$$\mathcal{L}_\infty \approx \begin{cases} -\frac{1}{n_{\mathrm{R}}'} \sum\limits_{i=1}^{n_{\mathrm{R}}'} \log_2 \left(\frac{1}{n_{\mathrm{T}} e} \sum\limits_{j=1}^{n_{\mathrm{T}}'} \frac{p_j'^\star (\mathbf{G}')_{i,j}}{1+\alpha_j} \right) - \frac{1}{n_{\mathrm{R}}'} \sum\limits_{j=1}^{n_{\mathrm{T}}'} \log_2 (1+\alpha_j) & n_{\mathrm{T}}' > n_{\mathrm{R}}' \\[2em] -\frac{1}{n_{\mathrm{T}}' n_{\mathrm{R}}'} \sum\limits_{j=1}^{n_{\mathrm{T}}'} \sum\limits_{i=1}^{n_{\mathrm{R}}'} \log_2 \frac{(\mathbf{G}')_{i,j}}{e} & n_{\mathrm{T}}' = n_{\mathrm{R}}' \\[2em] -\frac{1}{n_{\mathrm{T}}'} \sum\limits_{j=1}^{n_{\mathrm{T}}'} \log_2 \frac{\Gamma_j}{e} - \frac{1}{n_{\mathrm{T}}'} \sum\limits_{i=1}^{n_{\mathrm{R}}'} \log_2 \left(1 + \frac{1}{n_{\mathrm{T}}'} \sum\limits_{j=1}^{n_{\mathrm{T}}'} \frac{(\mathbf{G}')_{i,j}}{\Gamma_j} \right) & n_{\mathrm{T}}' < n_{\mathrm{R}}' \end{cases}$$

with α_j and Γ_j solutions to

$$\alpha_j = \sum_{i=1}^{n_{\mathrm{R}}'} \frac{p_j'^\star (\mathbf{G}')_{i,j}}{\sum_{\ell=1}^{n_{\mathrm{T}}'} p_\ell'^\star (\mathbf{G}')_{i,\ell}/(1+\alpha_\ell)} \tag{99}$$

26

and

$$\sum_{i=1}^{n'_{\text{R}}} \frac{1}{n'_{\text{T}} + \sum_{j=1}^{n'_{\text{T}}} (\mathbf{G}')_{i,j}/\Gamma_j} = \frac{n'_{\text{R}}}{n'_{\text{T}}} - 1 \tag{100}$$

respectively.

It is interesting to note that the low-SNR performance measures depend on n_{R} and n_{T} while the high-SNR measures depend on n'_{R} and n'_{T}. This is a consequence of the fact that the former are connected to the energy in the channel matrix, which from (2) is determined by n_{R} and n_{T}, whereas the latter relate essentially to the number of active degrees of freedom, which from Definition 6 is given by n'_{R} and n'_{T}.

Example 2 *Consider $n_{\text{T}} = 3$ and $n_{\text{R}} = 2$ with*

$$\mathbf{G} = \begin{bmatrix} 0.4 & 3.6 & 0.5 \\ 0.3 & 1 & 0.2 \end{bmatrix} \tag{101}$$

and with each (i,j)th entry of $\tilde{\mathbf{H}}$ having independent real and imaginary parts distributed uniformly on the interval

$$\left[-\sqrt{1.5\,(\mathbf{G})_{i,j}}, \sqrt{1.5\,(\mathbf{G})_{i,j}} \right].$$

The capacity-achieving power allocation is tabled within Fig. 1 along with plots of the capacity and of the spectral efficiency with uniform power allocation, both provided by (89), (90) and (91). The fixed-point equations in (90) and (91) are easily solved by initializing $\Gamma_j = \Upsilon_j = 1$ for $j = 1, 2, 3$ and iterating between them. Also shown in Fig. 1 are the results of corresponding Montecarlo simulations as well as the analytical high-SNR expansion of the spectral efficiency. As function of $\frac{E_b}{N_0}$, the capacity and spectral efficiency with uniform power allocation are portrayed in Fig. 2 along with the analytical low-SNR expansion.

The high-SNR slope in Example 2 is $S_\infty = 2$ bits/s/Hz/(3 dB). From our asymptotic expression, $\mathcal{L}_\infty \approx 3.6$ dB while the exact value obtained numerically is $\mathcal{L}_\infty = 3.5$ dB.

B Arbitrary Numbers of Antennas: UIU Channels

The nonasymptotic expressions for IND channels can be immediately applied to UIU channels.

For Rayleigh-faded channels with separable correlations, we can further specialize the expressions. From (86), the capacity becomes

$$C(\text{SNR}) \approx \sum_{j=1}^{n_{\text{T}}} \log_2 \left(\frac{1 + \text{SNR}\, p_j^\star \lambda_j(\boldsymbol{\Theta}_{\text{T}})\, \hat{\Gamma}}{e^{\text{SNR}\,\Gamma\Upsilon}} \right) + \sum_{i=1}^{n_{\text{R}}} \log_2 \left(1 + \text{SNR}\, \lambda_i(\boldsymbol{\Theta}_{\text{R}})\, \hat{\Upsilon} \right) \tag{102}$$
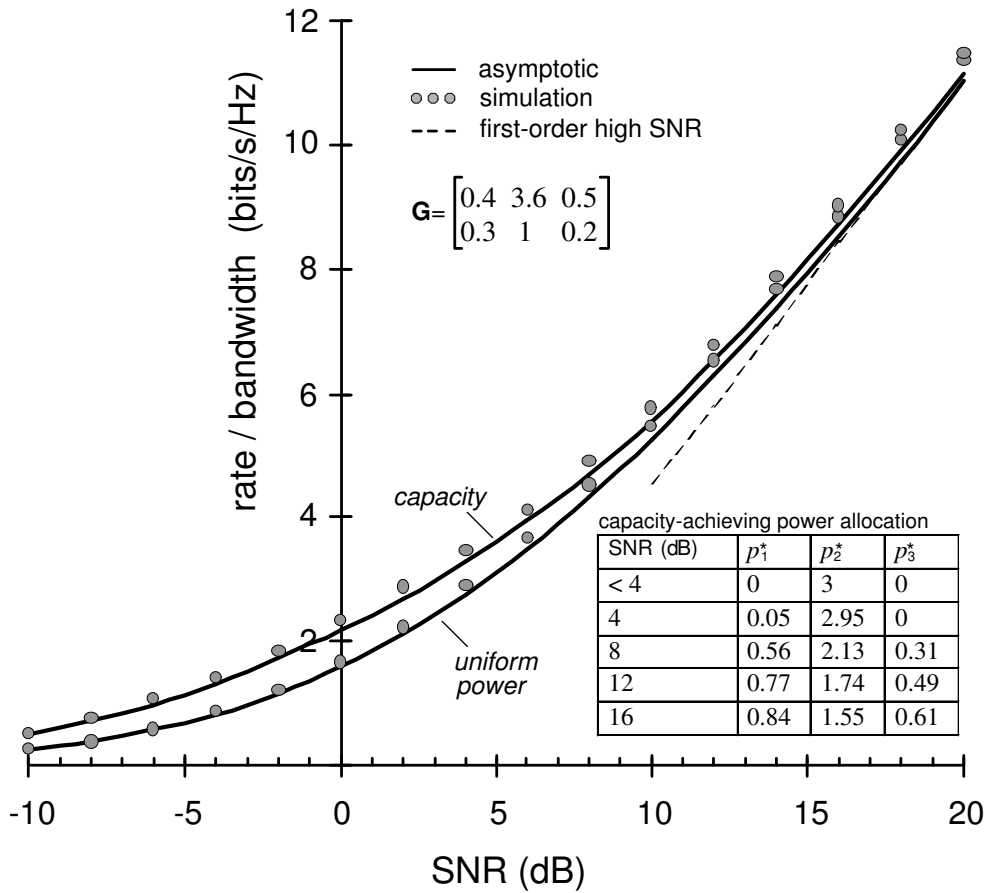
Figure 1: $C(\text{SNR})$ with $n_{\text{R}}{=}2$ and $n_{\text{T}}{=}3$ for the gain matrix **G** on display. Also shown is the spectral efficiency with uniform power allocation. Solid lines indicate the asymptotic approximation in (89), circles indicate simulation (fading having independent real and imaginary parts uniformly distributed around zero), dashed line indicates high-SNR expansion.
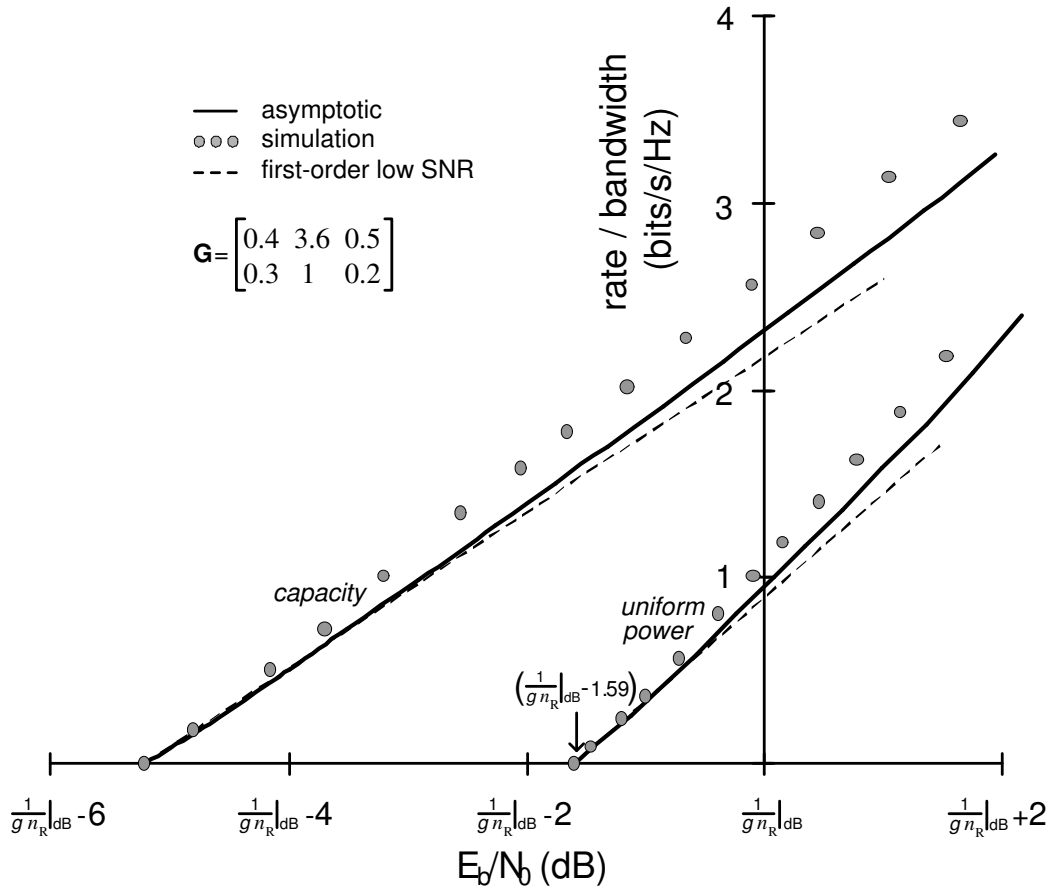
Figure 2: $C(\frac{E_b}{N_0})$ with $n_R=2$ and $n_T=3$ for the gain matrix $\mathbf{G}$ on display. Also shown is the spectral efficiency with uniform power allocation. Solid lines indicate the asymptotic approximation in (89), circles indicate simulation (fading having independent real and imaginary parts uniformly distributed around zero), dashed lines indicate low-SNR expansions.

with the powers $p_j^\star$ satisfying (39) and (40) while

$$\hat{\Gamma} \;=\; \frac{1}{n_\mathrm{T}} \sum_{i=1}^{n_\mathrm{R}} \frac{\lambda_i(\boldsymbol{\Theta}_\mathrm{R})}{1 + \mathsf{SNR}\,\lambda_i(\boldsymbol{\Theta}_\mathrm{R})\,\hat{\Upsilon}} \tag{103}$$

$$\hat{\Upsilon} \;=\; \frac{1}{n_\mathrm{T}} \sum_{j=1}^{n_\mathrm{T}} \frac{p_j^\star \lambda_j(\boldsymbol{\Theta}_\mathrm{T})}{1 + \mathsf{SNR}\,p_j^\star \lambda_j(\boldsymbol{\Theta}_\mathrm{T})\hat{\Gamma}}. \tag{104}$$

The low-SNR performance measures are

$$\frac{E_b}{N_0}_{\min} \;=\; \frac{\log_e 2}{g\,\mathrm{Tr}\{\boldsymbol{\Theta}_\mathrm{R}\}\lambda_{\mathsf{max}}(\boldsymbol{\Theta}_\mathrm{T})} \tag{105}$$

$$S_0 \;=\; \frac{2\,\mathrm{Tr}^2\{\boldsymbol{\Theta}_\mathrm{R}\}}{\mathrm{Tr}^2\{\boldsymbol{\Theta}_\mathrm{R}\}/m + \mathrm{Tr}\{\boldsymbol{\Theta}_\mathrm{R}^2\}} \tag{106}$$

where $\lambda_{\mathsf{max}}(\boldsymbol{\Theta}_\mathrm{T})$ denotes the largest eigenvalue of $\boldsymbol{\Theta}_\mathrm{T}$ and m its multiplicity. Notably, $\frac{E_b}{N_0}_{\min}$ depends mainly on $\boldsymbol{\Theta}_\mathrm{T}$ whereas S_0 depends mainly on $\boldsymbol{\Theta}_\mathrm{R}$.

For a uniform power allocation, the low-SNR performance measures of a Rayleigh-faded channel with separable correlations are given in [13]. Recall that such an input achieves capacity whenever there is correlation only at the receiver.

The high-SNR slope is $S_\infty = \min(n'_\mathrm{T}, n'_\mathrm{R})$ where n'_R and n'_T are the number of nonzero eigenvalues of $\boldsymbol{\Theta}_\mathrm{R}$ and $\boldsymbol{\Theta}_\mathrm{T}$, respectively. Thus, when such matrices are strictly positive definite, the linear scaling with the number of antennas of the first-order term in the high-SNR expansion of the capacity vs $\mathsf{SNR}|_{\mathrm{dB}}$ is retained. This observation, first made in [10] for large numbers of antennas and shown here to hold for arbitrary n_T and n_R, does not preclude a possibly large penalty in the zero-order term embodied by the power offset. Closed-form expressions for the high-SNR power offset of Rayleigh-faded channels with separable correlations and arbitrary numbers of antennas can be found in [34, 57].

In general, computing $C(\mathsf{SNR})$ requires finding $\hat{\Gamma}$ and $\hat{\Upsilon}$ by iterating between their expressions in (103) and (104). In some instances, however, e.g. when either transmitter or receiver are equipped with a single antenna or with uncorrelated antennas, $\hat{\Gamma}$ and $\hat{\Upsilon}$ can be found directly without the need to iterate. For example, if $\lambda_i(\boldsymbol{\Theta}_\mathrm{R})=1 \;\forall i$, then the expressions in (103) and (104) can be rearranged as

$$\hat{\Gamma} \;=\; \frac{n_\mathrm{R}}{n_\mathrm{T}} - \mathsf{SNR}\,\hat{\Gamma}\hat{\Upsilon} \tag{107}$$

$$\mathsf{SNR}\,\hat{\Gamma}\hat{\Upsilon} \;=\; 1 - \frac{1}{n_\mathrm{T}} \sum_{j=1}^{n_\mathrm{T}} \frac{1}{1 + \mathsf{SNR}\,p_j^\star \lambda_j(\boldsymbol{\Theta}_\mathrm{T})\hat{\Gamma}}. \tag{108}$$

Plugging in a value for $\mathsf{SNR}\,\hat{\Gamma}$ on (108) provides the corresponding $\mathsf{SNR}\,\hat{\Gamma}\hat{\Upsilon}$, from which (107) yields $\hat{\Gamma}$. Both SNR and $\hat{\Upsilon}$ can be then obtained and hence $C(\mathsf{SNR})$, as illustrated by the next example.

30

Example 3 *Consider a 4-antenna ULA transmit array and a receiver with 2 uncorrelated antennas. Further consider, at the transmitter, a broadside (truncated) Gaussian power azimuth spectrum with a 2° root-mean-square spread. With these conditions, typical of a mobile system deployed in a suburban area, the correlation between transmit antennas can be approximated by [58]*

$$(\boldsymbol{\Theta}_{\mathrm{T}})_{i,j} \approx e^{-0.05\, d^2 (i-j)^2} \tag{109}$$

where d is the antenna spacing (wavelengths). The approximation of the capacity provided by (102), (107) and (108) is plotted in Fig. 3, for Rayleigh fading and various values of d, alongside corresponding Montecarlo simulations as well as low-SNR expansions. Also shown, besides the capacity, is the spectral efficiency with an isotropic input.

VII   Conclusion: Impact of Correlation

Throughout this paper, we have presented analytical characterizations of the capacity and spectral efficiency (achieved by an isotropic input) of correlated multiantenna channels. Although with some emphasis on the prominent separable correlation model, our analysis has embraced the more general UIU model. The distinctive feature of a UIU channel is an IND matrix, which may surface in contexts other than the single-user multiantenna problem.

In the remainder of this section, we put these characterizations to use in order to derive insight on how multiantenna links are impacted by correlation.

A   Spectral Efficiency with Isotropic Input

When the input is isotropic, antenna correlation can only diminish the spectral efficiency. At low SNR, the $\frac{E_b}{N_0}_{\min}$ is unaffected and thus the reduction in spectral efficiency takes the form of a bandwidth penalty directly quantified by the decline in the slope S_0 [13]. In the high SNR region, on the other hand, the slope S_∞ is only sensitive to the number of nonnegligible eigenvalues and thus it is not altered by most correlation structures of practical interest. The penalty therein is reflected as a power offset increase determined, in 3-dB units, roughly by the average of the logarithm of the eigenvalues of the channel correlation function.
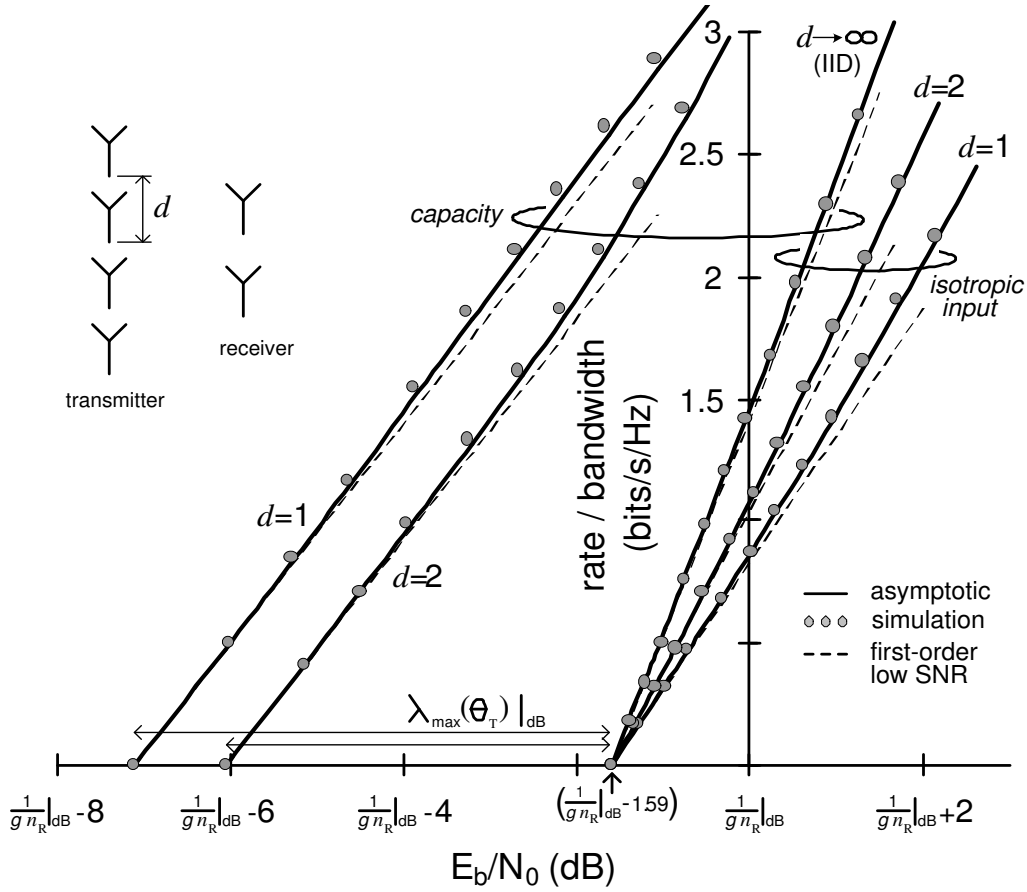
Figure 3: $C(\frac{E_b}{N_0})$ and spectral efficiency with an isotropic input, parameterized by d. The transmitter is a 4-antenna ULA with antenna spacing d (wavelengths) while the receiver has 2 uncorrelated antennas. The power angular spectrum at the transmitter is Gaussian (broadside) with $2°$ spread. Solid lines indicate the asymptotic approximation in (89), circles indicate simulation (Rayleigh fading), dashed lines indicate low-SNR expansion.

B Capacity

For the capacity (i.e., with optimized inputs), the picture is more involved. If we contrast a correlated UIU channel $\mathbf{H}$ against an IID channel normalized to the same average gain, we identify two distinct behaviors:

- If isotropic inputs achieve the capacity of the UIU channel, then it holds that correlation is detrimental at any SNR. This is the case if correlation and variance disparities arise between entries in different rows of $\mathbf{H}$, but there is uniformity across columns.

- If the capacity-achieving input for the UIU channel is nonisotropic (generally the case when correlation and variance disparities exist between entries in different columns of $\mathbf{H}$), the capacity curves may intersect with correlation being deleterious above a critical SNR but beneficial below it. The advantage at low SNR is a mere reflection of the findings in [47], where it is shown that correlation can only decrease $\frac{E_b}{N_0}_{\min}$.

For separable correlations specifically, we can draw the following lessons:

- Receive correlation reduces the effective dimensionality of the receiver without increasing the captured power and thus it is always detrimental.

- Transmit correlation reduces the effective dimensionality of the transmitter, but it also enables focusing power. The net effect is an advantage at low SNR where, as evidenced in Fig. 3, $\frac{E_b}{N_0}_{\min}$ experiences a reduction given precisely by the largest eigenvalue of $\boldsymbol{\Theta}_{\mathrm{T}}$. At high SNR, an advantage can also be realized if $n_{\mathrm{T}} > n_{\mathrm{R}}$.

C Correlation and Channel Gain

The foregoing assessments of the impact of correlation have been made on the basis of a fixed received signal-to-noise ratio, i.e., by contrasting the spectral efficiency and capacity of channel matrices with a common normalization. In doing so, we have intentionally decoupled our study from a second aspect that arises in the comparison of different propagation environments and different antenna arrays: besides the correlation, the average gain g in (4) may also differ. A complete comparison of different channels must necessarily account for both these aspects and should thus be conducted at a fixed transmitted signal-to-noise ratio, i.e., with a fixed transmitted power. The study of the average gain, however, has to do with radio propagation and channel modeling and is thus beyond the scope of this paper. For more on this theme, the interested reader is referred to [59, 60].

Acknowledgements

Appendices

A Proof of Theorem 2

Let $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{H}}_{\mathsf{Ray}}$ be two zero-mean column-regular IND channel having the same gain matrix $\mathbf{G}$. Further let the entries of $\tilde{\mathbf{H}}_{\mathsf{Ray}}$ be Rayleigh-faded while the entries of $\tilde{\mathbf{H}}$ have arbitrary marginal distributions.

Denote by $I_{\mathsf{Ray}}(\mathsf{SNR})$ the spectral efficiency of $\tilde{\mathbf{H}}_{\mathsf{Ray}}$ with an arbitrary input covariance $\mathbf{\Phi}$ and by $C_{\mathsf{R}}(\mathsf{SNR})$ the corresponding capacity, which is achieved by an isotropic input for any value of n_{T} and n_{R} [49]. Clearly,

$$\frac{1}{n_{\mathsf{R}}} C_{\mathsf{Ray}}(\mathsf{SNR}) \geq \frac{1}{n_{\mathsf{R}}} I_{\mathsf{Ray}}(\mathsf{SNR}). \tag{110}$$

Denote by $I(\mathsf{SNR})$ the spectral efficiency of $\tilde{\mathbf{H}}$ with the same input covariance $\mathbf{\Phi}$ above and by $I^{\circ}(\mathsf{SNR})$ the spectral efficiency of $\tilde{\mathbf{H}}$ with an isotropic input. Since, with a given input, the asymptotic mutual information of an IND channel does not depend on the marginal distribution of its entries,

$$\lim_{n_{\mathsf{R}} \to \infty} \tfrac{1}{n_{\mathsf{R}}} I^{\circ}(\mathsf{SNR}) = \lim_{n_{\mathsf{R}} \to \infty} \tfrac{1}{n_{\mathsf{R}}} C_{\mathsf{Ray}}(\mathsf{SNR}) \tag{111}$$

$$\lim_{n_{\mathsf{R}} \to \infty} \tfrac{1}{n_{\mathsf{R}}} I(\mathsf{SNR}) = \lim_{n_{\mathsf{R}} \to \infty} \tfrac{1}{n_{\mathsf{R}}} I_{\mathsf{Ray}}(\mathsf{SNR}) \tag{112}$$

from which, invoking (110) in the limit $n_{\mathsf{R}}' \to \infty$, it follows that

$$\lim_{n_{\mathsf{R}} \to \infty} \tfrac{1}{n_{\mathsf{R}}} I^{\circ}(\mathsf{SNR}) \geq \lim_{n_{\mathsf{R}} \to \infty} \tfrac{1}{n_{\mathsf{R}}} I(\mathsf{SNR}). \tag{113}$$

Thus, an isotropic input is asymptotically capacity-achieving on $\tilde{\mathbf{H}}$.

B Proof of Lemma 1

We shall make extensive use of the following result by Girko [39, Corollary 10.1.2], which we quote couched in the notation of this paper.

Theorem 6 *Let $\tilde{\mathbf{H}}$ be a $n_{\mathrm{R}} \times n_{\mathrm{T}}$ matrix with independent zero-mean arbitrarily distributed complex random entries whose uniformly bounded variances are $(\mathbf{G})_{i,j} = E[|(\tilde{\mathbf{H}})_{i,j}|^2]$. Denote $\beta = n_{\mathrm{T}}/n_{\mathrm{R}}$. Assume that the asymptotic gain profile of $\tilde{\mathbf{H}}$ as defined in (45) and (49) exists and is denoted by $\mathcal{G}(r,t)$ with $(r,t) \in [0,1] \times [0,\beta]$ and that the asymptotic power profile of $\mathbf{P}$ defined in (48) and (52) also exists and is denoted by $\mathcal{P}(t)$ with $t \in [0,\beta]$. Then, the asymptotic empirical eigenvalue distribution of $\frac{1}{n_{\mathrm{T}}} \tilde{\mathbf{H}} \mathbf{P} \tilde{\mathbf{H}}^\dagger$ converges almost surely to a nonrandom limit $F(\cdot)$ whose Stieltjes transform is*

$$\mathcal{S}(z) = \lim_{n_{\mathrm{R}} \to \infty} E\left[\left(\mathbf{I} + \tfrac{z}{n_{\mathrm{T}}} \tilde{\mathbf{H}} \mathbf{P} \tilde{\mathbf{H}}^\dagger \right)^{-1} \right] \tag{114}$$

$$= E[\mathcal{B}(\textsf{R}, z)] \tag{115}$$

with $\mathcal{B}(\cdot, \cdot)$ solution to the fixed-point equation

$$\mathcal{B}(r, z) = \cfrac{1}{1 + z\beta E\left[\cfrac{\mathcal{P}(\textsf{T})\mathcal{G}(r, \textsf{T})}{\beta + z E\left[\mathcal{B}(\textsf{R}, z)\mathcal{P}(\textsf{T})\mathcal{G}(\textsf{R}, \textsf{T}) | \textsf{T} \right]} \right]} \tag{116}$$

where the expectations are over $\textsf{R}$ and $\textsf{T}$, independent and respectively uniform in $(0,1]$ and $(0,\beta]$. The solution to (116) exists and is unique in the class of functions $\mathcal{B}(r, z)$ that are analytical on z and continuous on $r \in [0,1]$.

Define

$$\mathbf{W}_j \triangleq \tfrac{\textsf{SNR}}{n_{\mathrm{T}}} \sum_{\ell \neq j} p_\ell \tilde{\mathbf{h}}_\ell \tilde{\mathbf{h}}_\ell^\dagger. \tag{117}$$

Recalling $\frac{j}{n_{\mathrm{R}}} \leq t < \frac{j+1}{n_{\mathrm{R}}}$ and using [43, Proposition 2.7],

$$E\left[\left| \Gamma^{(n_{\mathrm{R}})}(t, \textsf{SNR}) - \frac{p_j}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} (\mathbf{G})_{i,j} \left(\mathbf{I} + \mathbf{W}_j \right)_{i,i}^{-1} \right|^2 \right] \leq \frac{K}{n_{\mathrm{R}}^2} \tag{118}$$

with K a constant.

Since $(\mathbf{G})_{i,j} \leq b, \forall i, j$ and for some bound $b < \infty$, then

$$\frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} (\mathbf{G})_{i,j} \left(\mathbf{I} + \mathbf{W}_j \right)_{i,i}^{-1} = \frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} \frac{(\mathbf{G})_{i,j}}{1 - \mathbf{w}_{j,i}^\dagger \left(\mathbf{I} + \mathbf{W}_{j,i} \right)^{-1} \mathbf{w}_{j,i} + \frac{\textsf{SNR}}{n_{\mathrm{T}}} \mathbf{r}_i \mathbf{P} \mathbf{r}_i^\dagger} \tag{119}$$

where $\mathbf{r}_i$ is the ith row of $\tilde{\mathbf{H}}$ while $\mathbf{w}_{j,i}$ is the ith column of $\mathbf{W}_j$ excluding the ith entry and $\mathbf{W}_{j,i}$ is the $(n_{\mathrm{R}} - 1) \times (n_{\mathrm{R}} - 1)$ submatrix obtained by eliminating from $\mathbf{W}_j$ the ith column and the ith row. Further defining

$$\mathcal{D}_i^{(n_{\mathrm{R}})}(\textsf{SNR}) \triangleq \left(\mathbf{I} + \tfrac{\textsf{SNR}}{n_{\mathrm{T}}} \sum_{\ell \neq j} p_\ell \tilde{\mathbf{h}}_\ell \tilde{\mathbf{h}}_\ell^\dagger \right)^{-1}_{i,i} \tag{120}$$

and

$$\varpi_i^{(n_{\mathrm{R}})} \triangleq \frac{\mathsf{SNR}}{n_{\mathrm{T}}} \sum_{j=1}^{n_{\mathrm{T}}} \frac{p_j (\mathbf{G})_{i,j}}{1 + \frac{\mathsf{SNR}}{n_{\mathrm{T}}} p_j \sum_{\ell=1}^{n_{\mathrm{R}}} (\mathbf{G})_{\ell,j} \mathcal{D}_\ell^{(n_{\mathrm{R}})}(\mathsf{SNR})} \tag{121}$$

we can write

$$\frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} (\mathbf{G})_{i,j} \mathcal{D}_i^{(n_{\mathrm{R}})}(\mathsf{SNR}) = \frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} \frac{(\mathbf{G})_{i,j}}{1 + \varpi_i^{(n_{\mathrm{R}})} + \epsilon_i^{(n_{\mathrm{R}})}} \tag{122}$$

$$= \frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} \frac{(\mathbf{G})_{i,j}}{1 + \varpi_i^{(n_{\mathrm{R}})}} - \delta_i^{(n_{\mathrm{R}})} \tag{123}$$

with

$$\epsilon_i^{(n_{\mathrm{R}})} \triangleq -\mathbf{w}_{j,i}^\dagger (\mathbf{I} + \mathbf{W}_{j,i})^{-1} \mathbf{w}_{j,i} + \frac{\mathsf{SNR}}{n_{\mathrm{T}}} \mathbf{r}_i \mathbf{P} \mathbf{r}_i^\dagger - \varpi_i^{(n_{\mathrm{R}})} \tag{124}$$

and

$$\delta_i^{(n_{\mathrm{R}})} \triangleq \frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} \frac{\epsilon_i^{(n_{\mathrm{R}})} (\mathbf{G})_{i,j}}{(1 + \varpi_i^{(n_{\mathrm{R}})})(1 + \varpi_i^{(n_{\mathrm{R}})} + \epsilon_i^{(n_{\mathrm{R}})})}. \tag{125}$$

Since the entries of $\mathbf{G}$ are uniformly bounded by b with respect to i, j and n_{T}, invoking Theorem 6 we have that a.s.

$$|\delta_i^{(n_{\mathrm{R}})}| \leq \left| \frac{b}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} \frac{\epsilon_i^{(n_{\mathrm{R}})}}{(1 + \varpi_i^{(n_{\mathrm{R}})})(1 + \varpi_i^{(n_{\mathrm{R}})} + \epsilon_i^{(n_{\mathrm{R}})})} \right| \tag{126}$$

$$= o(1) \tag{127}$$

Consequently,

$$\lim_{n_{\mathrm{R}} \to \infty} \frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} (\mathbf{G})_{i,j} \mathcal{D}_i^{(n_{\mathrm{R}})}(\mathsf{SNR}) = \lim_{n_{\mathrm{R}} \to \infty} \frac{1}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} \frac{(\mathbf{G})_{i,j}}{1 + \frac{\mathsf{SNR}}{n_{\mathrm{T}}} \sum_{j=1}^{n_{\mathrm{T}}} \frac{p_j(\mathbf{G})_{i,j}}{1 + \frac{\mathsf{SNR}}{n_{\mathrm{T}}} p_j \sum_{\ell=1}^{n_{\mathrm{R}}} (\mathbf{G})_{\ell,j} \mathcal{D}_\ell^{(n_{\mathrm{R}})}(\mathsf{SNR})}} \tag{128}$$

Denoting

$$\Gamma(t, \mathsf{SNR}) = \lim_{n_{\mathrm{R}} \to \infty} \frac{p_j}{n_{\mathrm{T}}} \sum_{i=1}^{n_{\mathrm{R}}} (\mathbf{G})_{i,j} \mathcal{D}_i^{(n_{\mathrm{R}})}(\mathsf{SNR}) \tag{129}$$

we obtain (58) from (128). Furthermore, using the first Borel-Cantelli Lemma [62], the sequence of functions $\Gamma^{(n_{\mathrm{R}})}(\cdot, \mathsf{SNR})$ converges to

$$\lim_{n_{\mathrm{R}} \to \infty} \Gamma^{(n_{\mathrm{R}})}(t, \mathsf{SNR}) \overset{a.s.}{=} \Gamma(t, \mathsf{SNR}) \tag{130}$$

with $\Gamma(\cdot, \mathsf{SNR})$ satisfying (58).

C   Proof of Theorem 3

Defining

$$\mathcal{I}(\textsc{snr}) \triangleq \lim_{n_{\textrm{R}} \to \infty} \frac{1}{n_{\textrm{R}}} I(\textsc{snr}), \tag{131}$$

its derivative with respect to $\textsc{snr}$ in (42) is

$$\dot{\mathcal{I}}(\textsc{snr}) = \frac{\log_2 e}{\textsc{snr}} \left(1 - \int \frac{1}{1 + \textsc{snr}\, \xi} dF(\xi) \right) \tag{132}$$

$$= \frac{\log_2 e}{\textsc{snr}} \left(1 - E\left[\mathcal{B}(\textsc{r}, \textsc{snr})\right] \right) \tag{133}$$

where (133) follows from Theorem 6 and $\mathcal{B}(r, \textsc{snr})$ satisfies (63). From (64),

$$\textsc{snr}\, \Gamma(\textsc{t}, \textsc{snr}) = \mathcal{P}(\textsc{t}) \frac{\textsc{snr}}{\beta} E[\mathcal{G}(\textsc{r}, \textsc{t}) \mathcal{B}(\textsc{r}, \textsc{snr}) | \textsc{t}] \tag{134}$$

is the SINR at the output of a linear MMSE receiver and thus $\beta\, E\left[\log_2(1 + \textsc{snr}\, \Gamma(\textsc{t}, \textsc{snr}))\right]$ is the mutual information achieved by a linear MMSE receiver. Notice also that

$$
\begin{aligned}
E\left[\frac{d}{d\,\textsc{snr}} \log_e(1 + \textsc{snr}\, \Gamma(\textsc{t}, \textsc{snr}))\right] &= E\left[\frac{\textsc{snr}\, E[\mathcal{P}(\textsc{t})\mathcal{G}(\textsc{r}, \textsc{t})\dot{\mathcal{B}}(\textsc{r}, \textsc{snr}) | \textsc{t}]}{\beta + \textsc{snr}\, E[\mathcal{P}(\textsc{t})\mathcal{G}(\textsc{r}, \textsc{t})\mathcal{B}(\textsc{r}, \textsc{snr}) | \textsc{t}]}\right] \\
&\quad + E\left[\frac{E[\mathcal{P}(\textsc{t})\mathcal{G}(\textsc{r}, \textsc{t})\mathcal{B}(\textsc{r}, \textsc{snr}) | \textsc{t}]}{\beta + \textsc{snr}\, E[\mathcal{P}(\textsc{t})\mathcal{G}(\textsc{r}, \textsc{t})\mathcal{B}(\textsc{r}, \textsc{snr}) | \textsc{t}]}\right]
\end{aligned}
\tag{135}
$$

which, in conjunction with (133) and (134), yields

$$
\begin{aligned}
\dot{\mathcal{I}}(\textsc{snr}) - \beta\, E\left[\tfrac{d}{d\textsc{snr}} \log_2(1 + \textsc{snr}\, \Gamma(\textsc{t}, \textsc{snr}))\right] &= \frac{-\beta}{\log_e 2} E\left[\frac{\textsc{snr}\, E[\mathcal{P}(\textsc{t})\mathcal{G}(\textsc{r}, \textsc{t})\dot{\mathcal{B}}(\textsc{r}, \textsc{snr}) | \textsc{t}]}{\beta + \textsc{snr}\, E[\mathcal{P}(\textsc{t})\mathcal{G}(\textsc{r}, \textsc{t})\mathcal{B}(\textsc{r}, \textsc{snr}) | \textsc{t}]}\right] \\
&= -\beta\, E\left[\dot{\Gamma}(\textsc{t}, \textsc{snr})\, \Upsilon(\textsc{t}, \textsc{snr})\right] \log_2 e \tag{136}
\end{aligned}
$$

where we have further used

$$\Upsilon(t, \textsc{snr}) = \frac{\beta\, \textsc{snr}}{\beta + \textsc{snr}\, E[\mathcal{B}(\textsc{r}, \textsc{snr})\mathcal{P}(t)\mathcal{G}(\textsc{r}, t)]}. \tag{137}$$

Note that

$$\frac{d}{d\textsf{SNR}}E\left[\Gamma(\textsf{T},\textsf{SNR})\Upsilon(\textsf{T},\textsf{SNR})\right] = E\left[\Gamma(\textsf{T},\textsf{SNR})\dot{\Upsilon}(\textsf{T},\textsf{SNR})\right] + E\left[\dot{\Gamma}(\textsf{T},\textsf{SNR})\Upsilon(\textsf{T},\textsf{SNR})\right]$$

$$= \frac{1}{\beta}E\left[\frac{\dot{\Upsilon}(\textsf{T},\textsf{SNR})\mathcal{P}(\textsf{T})\mathcal{G}(\textsf{R},\textsf{T})}{1 + E[\mathcal{P}(\textsf{T})\mathcal{G}(\textsf{R},\textsf{T})\Upsilon(\textsf{T},\textsf{SNR})|\textsf{R}]}\right] + E\left[\dot{\Gamma}(\textsf{T},\textsf{SNR})\Upsilon(\textsf{T},\textsf{SNR})\right]$$

$$= \frac{1}{\beta}E\left[\frac{E\left[\dot{\Upsilon}(\textsf{T},\textsf{SNR})\mathcal{P}(\textsf{T})\mathcal{G}(\textsf{R},\textsf{T})|\textsf{R}\right]}{1 + E[\mathcal{P}(\textsf{T})\mathcal{G}(\textsf{R},\textsf{T})\Upsilon(\textsf{T},\textsf{SNR})|\textsf{R}]}\right] + E\left[\dot{\Gamma}(\textsf{T},\textsf{SNR})\Upsilon(\textsf{T},\textsf{SNR})\right]$$

$$= \frac{1}{\beta}\frac{d}{d\textsf{SNR}}E\left[\log_e(1 + E[\mathcal{P}(\textsf{T})\mathcal{G}(\textsf{R},\textsf{T})\Upsilon(\textsf{T},\textsf{SNR})|\textsf{R}])\right]$$

$$+ E\left[\dot{\Gamma}(\textsf{T},\textsf{SNR})\Upsilon(\textsf{T},\textsf{SNR})\right] \tag{138}$$

from which, integrating over SNR and using $I(0) = 0$, the claimed result is found.

In addition,

$$-\beta E\left[\Gamma(\textsf{T},\textsf{SNR})\,\Upsilon(\textsf{T},\textsf{SNR})\right] = E[\mathcal{B}(\textsf{R},\textsf{SNR})] - 1 \tag{139}$$

and also, using (63) and (137),

$$E\left[\log_2(1 + E[\mathcal{G}(\textsf{R},\textsf{T})\mathcal{P}(\textsf{T})\Upsilon(\textsf{T},\textsf{SNR})|\textsf{R}])\right] = -E\left[\log_2 \mathcal{B}(\textsf{R},\textsf{SNR})\right]. \tag{140}$$

Using (137), (139) and (140), the alternative expression in (69) is easily proved.

D Proof of Conditions (65) and (66)

Although (39) and (40) are formulated on the basis of expected mean-square errors, the expectations become asymptotically immaterial as the mean-square errors converge a.s. to nonrandom limits. Equations (66) and (65) are immediately obtained simply by replacing the corresponding quantities in (39) and (40) with their asymptotic limits. The subset in (67) is elaborated in the remainder of this appendix.

Recalling that $\frac{j}{n_\textsf{R}}\leq t<\frac{j+1}{n_\textsf{R}}$ and using [43, Proposition 2.7],

$$E\left[\left|\frac{\textsf{SNR}}{n_\textsf{T}}\tilde{\mathbf{h}}_j^\dagger\left(\mathbf{I}+\mathbf{W}_j\right)^{-1}\tilde{\mathbf{h}}_j - \frac{\textsf{SNR}}{n_\textsf{T}}\sum_{i=1}^{n_\textsf{R}}(\mathbf{G})_{i,j}\left(\mathbf{I}+\mathbf{W}_j\right)_{i,i}^{-1}\right|^2\right] \leq \frac{K}{n_\textsf{R}^2} \tag{141}$$

with $\mathbf{W}_j$ defined in (117) and with K a constant. $(\mathbf{I}+\mathbf{W}_j)^{-1}$. Following the same steps of Appendix B and using (141) and (128), we conclude that

$$\lim_{n_\textsf{R}\to\infty}\gamma_j \stackrel{a.s.}{=} \frac{\textsf{SNR}}{\beta}E\left[\mathcal{G}(\textsf{R},t)\mathcal{B}(\textsf{R},\textsf{SNR})\right]. \tag{142}$$

E Proof of Proposition 1

Letting $\mathcal{P}(t)=1$ and averaging both sides of (63), we obtain

$$\frac{1 - E[\mathcal{B}(\textsc{r}, \textsc{snr})]}{\textsc{snr}} = \beta\, E\left[\frac{E[\mathcal{B}(\textsc{r}, \textsc{snr})\, \mathcal{G}(\textsc{r}, \textsc{t})|\textsc{t}]}{\beta + \textsc{snr}\, E[\mathcal{B}(\textsc{r}, \textsc{snr})\, \mathcal{G}(\textsc{r}, \textsc{t})|\textsc{t}]}\right] \tag{143}$$

which, recalling (64), can be plugged into (133) to yield

$$\dot{\mathcal{I}}(\textsc{snr}) = \beta\, E\left[\frac{\Gamma(\textsc{t}, \textsc{snr})}{1 + \textsc{snr}\, \Gamma(\textsc{t}, \textsc{snr})}\right] \log_2 e. \tag{144}$$

Defining $\eta(\textsc{snr}) \triangleq E[\Gamma(\textsc{t}, \textsc{snr})]$ and using Jensen's inequality,

$$\dot{\mathcal{I}}(\textsc{snr}) \leq \beta\, \frac{\eta(\textsc{snr})}{1 + \textsc{snr}\, \eta(\textsc{snr})} \log_2 e. \tag{145}$$

If $\eta(\textsc{snr})$ satisfies

$$\beta\, \eta(\textsc{snr}) = \frac{1}{1 + \frac{\textsc{snr}}{1 + \textsc{snr}\, \eta(\textsc{snr})}} \tag{146}$$

then the right-hand side of (145) represents the derivative of the asymptotic capacity per antenna of an IID channel. Thus, $\dot{\mathcal{I}}(\textsc{snr})$ is upper-bounded at every $\textsc{snr}$ by the value it takes on an IID channel and, consequently, so is $\mathcal{I}(\textsc{snr})$.

This upper bound is proved in [63] for arbitrary numbers of antennas, albeit only for Rayleigh-faded channels with separable correlations.

F Proof of Theorems 4 and 5

We will start by proving Theorem 5, from which Theorem 4 then follows.

In order to prove Theorem 5, we need to show that, when $\mathbf{G}$ is asymptotically mean doubly-regular, the asymptotic empirical eigenvalue distribution of $\frac{1}{n_{\textsc{t}}}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\dagger}$ coincides with that of a matrix having unit-variance independent entries. This equivalence has already been proved for Hermitian band matrices in [61, Corollary 3.4]. Inspired by this result, and making once again use of Theorem 6, we extend this equivalence to non-Hermitian matrices.

When $\tilde{\mathbf{H}}$ has independent arbitrarily distributed unit-variance entries, the asymptotic empirical eigenvalue distribution of $\frac{1}{n_{\textsc{t}}}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\dagger}$ converges a.s. to the Marčenko-Pastur law with parameter $1/\beta$ whose density function is [46, 38]

$$\mathsf{f}_{1/\beta}(\xi) = (1 - \beta)^+ \delta(\xi) + \beta\, \frac{\sqrt{[\xi - a]^+ [b - \xi]^+}}{2\pi\xi} \tag{147}$$

39

where $[z]^+ = \max(0, z)$ and

$$a = (1 - \frac{1}{\sqrt{\beta}})^2 \tag{148}$$

$$b = (1 + \frac{1}{\sqrt{\beta}})^2. \tag{149}$$

The Stieltjes transform of (147), denoted $\mathcal{B}_w(\cdot)$, is solution to the fixed-point equation

$$\mathcal{B}_w(z) = \frac{1}{1 + \dfrac{z\beta}{\beta + z\mathcal{B}_w(z)}}. \tag{150}$$

For $\tilde{\mathbf{H}}$ having IND arbitrarily distributed entries whose variances are normalized as per (2), the asymptotic empirical eigenvalue distribution of $\frac{1}{n_{\mathsf{T}}}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger$ converges a.s. to a non-random limit whose Stieltjes transform satisfies (116). The solution to (116) exists and is unique in the class of functions $\mathcal{B}(r, z)$ that are analytical on z and continuous on $r \in [0, 1]$. In order to prove that, for $\mathbf{G}$ asymptotically mean doubly-regular, the asymptotic empirical eigenvalue distribution of $\frac{1}{n_{\mathsf{T}}}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger$ converges a.s. to (147), we simply need to verify that under such conditions (116) reduces to (150) and is thus satisfied by $\mathcal{B}_w(z)$.

To this end, recall that if $\mathbf{G}$ is asymptotically mean doubly-regular and normalized as per (2),

$$\lim_{n_{\mathsf{R}} \to \infty} \frac{1}{n_{\mathsf{R}}} \sum_{i=1}^{n_{\mathsf{R}}} (\mathbf{G})_{i,j} = \lim_{n_{\mathsf{T}} \to \infty} \frac{1}{n_{\mathsf{T}}} \sum_{j=1}^{n_{\mathsf{T}}} (\mathbf{G})_{i,j} \tag{151}$$

$$= E\left[\mathcal{G}(\mathsf{R}, t)\right] \tag{152}$$

$$= E\left[\mathcal{G}(r, \mathsf{T})\right] \tag{153}$$

$$= 1. \tag{154}$$

Plugging $\mathcal{B}_w(z)$ in (116),

$$\mathcal{B}(r, z) = \frac{1}{1 + z\beta E\left[\dfrac{\mathcal{G}(r, \mathsf{T})}{\beta + z\mathcal{B}_w(z) E\left[\mathcal{G}(\mathsf{R}, \mathsf{T})|\mathsf{T}\right]}\right]} \tag{155}$$

$$= \frac{1}{1 + z\beta E\left[\dfrac{\mathcal{G}(r, \mathsf{T})}{\beta + z\mathcal{B}_w(z)}\right]} \tag{156}$$

$$= \frac{1}{1 + z\beta \dfrac{E\left[\mathcal{G}(r, \mathsf{T})\right]}{\beta + z\mathcal{B}_w(z)}} \tag{157}$$

$$= \frac{1}{1 + \dfrac{z\beta}{\beta + z\mathcal{B}_w(z)}} \tag{158}$$

40

which is exactly (150). Thus, the asymptotic empirical eigenvalue distribution of $\frac{1}{n_{\mathrm{T}}}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\dagger}$ converges a.s. to (147). Integration over this distribution via (42) with a uniform power allocation yields the expression in (75) [3, 4].

For those channels for which the capacity-achieving input is isotropic, (75) gives the asymptotic capacity. Since a sufficient condition for the optimality of isotropic inputs (cf. Theorem 2) is that $\mathbf{G}$ be column-regular, Theorem 4 is readily obtained by incorporating this stricter condition to Theorem 5.

G Proof of Proposition 4

The high-SNR slope is [33]

$$S_{\infty} = \lim_{\mathrm{SNR}\to\infty} \frac{\mathrm{SNR}\,\dot{C}(\mathrm{SNR})}{\log_2 e} \tag{159}$$

where

$$\dot{C}(\mathrm{SNR}) = \sum_{i=1}^{n_{\mathrm{R}}} E\left[\frac{\lambda_i(\tilde{\mathbf{H}}\mathbf{P}^{\star}\tilde{\mathbf{H}}^{\dagger})}{n_{\mathrm{T}} + \mathrm{SNR}\,\lambda_i(\tilde{\mathbf{H}}\mathbf{P}^{\star}\tilde{\mathbf{H}}^{\dagger})}\right]\log_2 e \tag{160}$$

with the diagonal entries of $\mathbf{P}^{\star}$ satisfying (39) and (40). Then,

$$\lim_{\mathrm{SNR}\to\infty} \mathrm{SNR}\,\dot{C}(\mathrm{SNR}) = E[\mathrm{rank}\{\tilde{\mathbf{H}}\mathbf{P}^{\star}\tilde{\mathbf{H}}^{\dagger}\}]\log_2 e. \tag{161}$$

Since the entries of $\tilde{\mathbf{H}}$ are independent, with probability 1 the rows and columns of $\tilde{\mathbf{H}}$ are linearly independent except for those that are identically zero. Also, recall that for SNR$\to\infty$ nonzero power is allocated to every column of $\tilde{\mathbf{H}}$ that is not identically zero (cf. Section IV) and thus $\mathbf{P}^{\star}$ does not reduce the number of linearly independent columns. Altogether then,

$$E[\mathrm{rank}\{\tilde{\mathbf{H}}\mathbf{P}^{\star}\tilde{\mathbf{H}}^{\dagger}\}] = \min(n'_{\mathrm{T}}, n'_{\mathrm{R}}) \tag{162}$$

from which S_{∞} follows.

H Proof of Corollary 1

When $\mathcal{G}(r,t)=\mathcal{G}_{\mathrm{R}}(r)\mathcal{G}_{\mathrm{T}}(t)$, the expression for $\Gamma(t,\mathrm{SNR})$ in Theorem 3 becomes

$$\Gamma(t,\mathrm{SNR}) = \frac{\mathcal{P}(t)\mathcal{G}_{\mathrm{T}}(t)}{\beta} E\left[\frac{\mathcal{G}_{\mathrm{R}}(\mathrm{R})}{1+\mathrm{SNR}\,\mathcal{G}_{\mathrm{R}}(\mathrm{R})\,\hat{\Upsilon}(\mathrm{SNR})}\right] \tag{163}$$

$$= \mathcal{P}(t)\mathcal{G}_{\mathrm{T}}(t)\,\hat{\Gamma}(\mathrm{SNR}) \tag{164}$$

with

$$\hat{\Gamma}(\text{SNR}) \triangleq \frac{1}{\beta} E\left[\frac{\mathcal{G}_{\text{R}}(\text{R})}{1 + \text{SNR}\,\mathcal{G}_{\text{R}}(\text{R})\hat{\Upsilon}(\text{SNR})}\right] \tag{165}$$

and

$$\hat{\Upsilon}(\text{SNR}) \triangleq \frac{1}{\text{SNR}} E[\mathcal{P}(\text{T})\mathcal{G}_{\text{T}}(\text{T})\Upsilon(\text{T}, \text{SNR})] \tag{166}$$

It also follows from these expressions that

$$E[\mathcal{P}(\text{T})\mathcal{G}(\text{R}, \text{T})\Upsilon(\text{T}, \text{SNR})|\text{R}] = \text{SNR}\,\mathcal{G}_{\text{R}}(\text{R})\,\hat{\Upsilon}(\text{SNR}) \tag{167}$$

Likewise, using the definition of $\Upsilon(t, \text{SNR})$ in Theorem 3 we obtain

$$\hat{\Upsilon}(\text{SNR}) = E\left[\frac{\mathcal{P}(\text{T})\mathcal{G}_{\text{T}}(\text{T})}{1 + \text{SNR}\,\mathcal{P}(\text{T})\mathcal{G}_{\text{T}}(\text{T})\hat{\Gamma}(\text{SNR})}\right] \tag{168}$$

as well as

$$
\begin{aligned}
E[\Gamma(\text{T}, \text{SNR})\,\Upsilon(\text{T}, \text{SNR})] &= \hat{\Gamma}(\text{SNR})E[\mathcal{P}(\text{T})\mathcal{G}_{\text{T}}(\text{T})\Upsilon(\text{T}, \text{SNR})] & (169)\\
&= \text{SNR}\,\hat{\Gamma}(\text{SNR})\,\hat{\Upsilon}(\text{SNR}) & (170)
\end{aligned}
$$

Plugging (164), (167) and (170) into (68), Corollary 1 is obtained.

I Proof of Proposition 2

Recall the definition of the subsets $\mathcal{S}_{\text{T}}$ and $\mathcal{S}_{\text{R}}$ in (53) and (54). If $t \notin \mathcal{S}_{\text{T}}$, then $\mathcal{G}_{\text{T}}(t)=0$ and hence, $\forall r \in [0, 1]$, $\mathcal{G}(r, t)=0$. If $r \notin \mathcal{S}_{\text{R}}$ then, $\forall t \in [0, \beta]$, $\mathcal{G}(r, t)=0$.

Let us start by studying the behavior of $\Gamma(t, \text{SNR})$ for $\text{SNR} \to \infty$, to which end we define $\Gamma_{\infty}(t) \triangleq \lim_{\text{SNR} \to \infty} \Gamma(t, \text{SNR})$. If $t \notin \mathcal{S}_{\text{T}}$, then $\Gamma_{\infty}(t)=0$. If $t \in \mathcal{S}_{\text{T}}$, there are two scenarios:

a) $\Gamma_{\infty}(t)=0$.

b) $\Gamma_{\infty}(t)>0$. In this case, we apply the law of total probability to (58) and take the limit on SNR to obtain

$$\Gamma_{\infty}(t) = \frac{1}{\beta'} E\left[\frac{\mathcal{G}(\text{R}, t)\mathcal{P}_{\infty}^{\star}(t)\Pr\{\text{T} \in \mathcal{S}_{\text{T}}\}}{1 + E\left[\frac{\mathcal{G}(\text{R}, \text{T})\mathcal{P}_{\infty}^{\star}(\text{T})\Pr\{\text{T} \in \mathcal{S}_{\text{T}}\}}{\Gamma_{\infty}(\text{T})}\Big|\text{R}, \text{T} \in \mathcal{S}_{\text{T}}\right]}\Big|\text{R} \in \mathcal{S}_{\text{R}}\right] \tag{171}$$

where we have exploited the strict positiveness of $\mathcal{P}_\infty^\star(t)$ on $\mathcal{S}_\mathrm{T}$ and we have also used (55). We can rearrange (171) into

$$E\left[\frac{\frac{\mathcal{G}(\mathrm{R},t)\mathcal{P}_\infty^\star(t)\mathrm{Pr}\{\mathrm{T}\in\mathcal{S}_\mathrm{T}\}}{\Gamma_\infty(t)}}{1+E\left[\frac{\mathcal{G}(\mathrm{R},\mathrm{T})\mathcal{P}_\infty^\star(\mathrm{T})\mathrm{Pr}\{\mathrm{T}\in\mathcal{S}_\mathrm{T}\}}{\Gamma_\infty(\mathrm{T})}\Big|\mathrm{R},\mathrm{T}\in\mathcal{S}_\mathrm{T}\right]}\Big|\mathrm{R}\in\mathcal{S}_\mathrm{R}\right]=\beta'. \tag{172}$$

and, averaging both sides over T conditioned on $\mathrm{T}\in\mathcal{S}_\mathrm{T}$, (172) can be further manipulated into

$$E\left[\frac{1}{1+E\left[\frac{\mathcal{G}(\mathrm{R},\mathrm{T})\mathcal{P}_\infty^\star(\mathrm{T})\mathrm{Pr}\{\mathrm{T}\in\mathcal{S}_\mathrm{T}\}}{\Gamma_\infty(\mathrm{T})}\Big|\mathrm{R},\mathrm{T}\in\mathcal{S}_\mathrm{T}\right]}\Big|\mathrm{R}\in\mathcal{S}_\mathrm{R}\right]=1-\beta' \tag{173}$$

Let us evaluate the conditions under which $\Gamma_\infty(t)$ assumes zero or nonzero values. Since $\Gamma_\infty(t)$ is nonnegative:

- If $\beta'\geq 1$, (173) does not admit a nonzero solution and thus, necessarily, $\Gamma_\infty(t)=0$.

- If $\beta'<1$, (173) does admit a nonzero[9] solution for $\Gamma_\infty(t)$. Furthermore, in that case we know that $\mathcal{P}_\infty^\star(t)$ is uniform over $\mathcal{S}_\mathrm{T}$ and zero outside. Thus, as function of R' and T', restrictions of R and T to $\mathcal{S}_\mathrm{R}$ and $\mathcal{S}_\mathrm{T}$ respectively, (173) becomes (80) as claimed.

Since, for $\beta'\geq 1$, $\Gamma(t,\mathrm{SNR})$ vanishes, we need to evaluate the limiting value of $\mathrm{SNR}\,\Gamma(t,\mathrm{SNR})$, to which end we define

$$\alpha(t)\triangleq\lim_{\mathrm{SNR}\to\infty}\mathrm{SNR}\,\Gamma(t,\mathrm{SNR}). \tag{174}$$

For $t\notin\mathcal{S}_\mathrm{T}$, $\Gamma(t,\mathrm{SNR})=0\ \forall\mathrm{SNR}$ and, by continuity, $\alpha(t)=0$. For $t\in\mathcal{S}_\mathrm{T}$ and $\beta'<1$, $\Gamma_\infty(t)$ is a nonzero solution of (173) and thus $\alpha(t)$ diverges. For $t\in\mathcal{S}_\mathrm{T}$ and $\beta'\geq 1$, and under the hypothesis that $0<\alpha(t)<\infty$, we can again use (58) to obtain

$$\alpha(t)=\frac{1}{\beta'}E\left[\frac{\mathcal{G}(\mathrm{R},t)\mathcal{P}_\infty^\star(t)}{E\left[\frac{\mathcal{G}(\mathrm{R},\mathrm{T})\mathcal{P}_\infty^\star(\mathrm{T})}{1+\alpha(\mathrm{T})}\Big|\mathrm{R},\mathrm{T}\in\mathcal{S}_\mathrm{T}\right]}\Big|\mathrm{R}\in\mathcal{S}_\mathrm{R}\right] \tag{175}$$

For $\beta'>1$, (175) admits a unique nonzero solution. Moreover, as function of T' and R', it can be written as in (79).

Having characterized the limiting behavior of $\Gamma(t,\mathrm{SNR})$ for $\mathrm{SNR}\to\infty$, we can now utilize it to calculate $\mathcal{L}_\infty$, which by definition is

$$\mathcal{L}_\infty=\lim_{\mathrm{SNR}\to\infty}\log_2\mathrm{SNR}-\frac{1}{S_\infty}C(\mathrm{SNR}) \tag{176}$$

---

[9]For $\beta'<1$, the solution to (173) can be zero only outside $\mathcal{S}_\mathrm{T}$, where $\Gamma(t,\mathrm{SNR})$ is itself zero as in scenario a).

with $C(\text{SNR})/n_{\text{R}}$ given by (68) and $S_\infty/n_{\text{R}}=\min(\Pr\{\text{R} \in \mathcal{S}_{\text{R}}\}, \beta\Pr\{\text{T} \in \mathcal{S}_{\text{T}}\})$. For $\beta'>1$, the law of total probability and the fact that $\alpha(t)=0$ for $t \notin \mathcal{S}_{\text{T}}$ yield

$$
\begin{aligned}
\mathcal{L}_\infty &= -\beta'E\left[\log_2(1+\alpha(\text{T}))|\text{T} \in \mathcal{S}_{\text{T}}\right] - E\left[\log_2\left(E\left[\frac{\mathcal{G}(\text{R},\text{T})\mathcal{P}_\infty^\star(\text{T})\Pr\{\text{T} \in \mathcal{S}_{\text{T}}\}}{1+\alpha(\text{T})}\Big|\text{R},\text{T} \in \mathcal{S}_{\text{T}}\right]\right)\Big|\text{R} \in \mathcal{S}_{\text{R}}\right] \\
&\quad + \beta'E\left[\frac{\alpha(\text{T})}{1+\alpha(\text{T})}\Big|\text{T} \in \mathcal{S}_{\text{T}}\right]\log_2 e
\end{aligned}
\tag{177}
$$

which, invoking (175) with expectation conditioned on $\text{T} \in \mathcal{S}_{\text{T}}$ on both sides, becomes

$$
\mathcal{L}_\infty = -\beta'E\left[\log_2(1+\alpha(\text{T}))|\text{T} \in \mathcal{S}_{\text{T}}\right] - E\left[\log_2\left(\frac{\Pr\{\text{T} \in \mathcal{S}_{\text{T}}\}}{e}E\left[\frac{\mathcal{G}(\text{R},\text{T})\mathcal{P}_\infty^\star(\text{T})}{1+\alpha(\text{T})}\Big|\text{R},\text{T} \in \mathcal{S}_{\text{T}}\right]\right)\Big|\text{R} \in \mathcal{S}_{\text{R}}\right]
\tag{178}
$$

from which, replacing the conditioned expectations over R and T with expectations over R' and T' respectively, the desired expression for $\beta'>1$ is found. The corresponding expressions for $\beta'\leq 1$ are found via similar expansions from (176) using $\Gamma_\infty(t)=0$ for $t \notin \mathcal{S}_{\text{T}}$.

Finally, condition (81) is easily found from (66) and the definition of $\alpha(t)$.

References

[1] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 315–335, Mar. 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom*, vol. 10, pp. 585–595, Nov. 1999.

[3] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. on Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.

[4] P. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output channel," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1245–1248, Aug. 2000.

[5] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering and keyhole," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 2636–2647, Oct. 2003.

[6] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effects on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[7] D. Chizhik, F. R. Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Commun. Letters*, vol. 4, no. 11, pp. 337–339, Nov. 2000.

[8] K. I. Pedersen, J. B. Andersen, J. P. Kermoal, and P. E. Mogensen, "A stochastic multiple-input multiple-output radio channel model for evaluations of space-time coding algorithms," *Proc. IEEE Vehic. Tech. Conf. (VTC'2000 Fall)*, pp. 893–897, 2000.

[9] A. Moustakas, H. Baranger, L. Balents, A. Sengupta, and S. Simon, "Communication through a diffusive medium: Coherence and capacity," *Science*, vol. 287, pp. 287–290, Jan. 2000.

[10] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in dual-antenna-array wireless systems," *IEEE Trans. on Inform. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[11] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 2545–2561, Oct. 2003.

[12] S. A. Jafar, S. Vishwanath, and A. J. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback," *IEEE Trans. on Wireless Communications*, vol. 3, pp. 1165–1175, July 2004.

[13] A. Lozano, A. M. Tulino, and S. Verdú, "Multi-antenna capacity in the low-power regime," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 2527–2544, Oct. 2003.

[14] X. Mestre, J. R. Fonollosa, and A. Pagès-Zamora, "Capacity of MIMO channels: asymptotic evaluation under correlated fading," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 829–838, June 2003.

[15] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 406–417, Apr. 2003.

[16] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semi-correlated flat fading," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 2781–2788, Oct. 2003.

[17] W. C. Y. Lee and Y. S. Yeh, "Polarization diversity system for mobile radio," *IEEE Trans. on Communic.*, vol. 20, no. 5, pp. 912–923, Oct. 1972.

[18] S. A. Bergmann and H. W. Arnold, "Polarization diversity in portable communications environment," *IEE Electronic Letters*, vol. 22, no. 11, pp. 609–610, May 1986.

[19] C. B. Dietrich Jr., K. Dietze, J. R. Nealy, and W. L. Stutzman, "Spatial, polarization, and pattern diversity for wireless handheld terminals," *IEEE Trans. on Antennas and Propag.*, vol. 49, pp. 1271–1281, Sept. 2001.

[20] A. M. Sayeed, "Deconstructing multi-antenna channels," *IEEE Trans. on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[21] J. H. Kotecha and A. Sayeed, "Optimal signal design for estimation of correlated MIMO channels," *IEEE Intern. Conf. in Communic. (ICC'03)*, pp. 3170–3174, May 2003.

[22] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications (to appear)*, 2005.

[23] H. Özcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of the 'Kronecker' MIMO radio channel model," *IEE Electronics Letters*, vol. 39, no. 16, pp. 1209–1210, Aug. 2003.

[24] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations and capacities of multi-element transmit and receive antennas," *IEEE Trans. on Wireless Commun.*, vol. 1, no. 2, pp. 361–368, Apr. 2002.

[25] D. Gesbert, H. Bolcskei, D. Gore, and A. J. Paulraj, "MIMO wireless channels: Capacity and performance prediction," *Proc. IEEE GLOBECOM'00*, vol. 2, pp. 1083–1088, Dec. 2000.

[26] C. Oestges and A. J. Paulraj, "Beneficial impact of channel correlations on MIMO capacity," *IEE Electronics Letters*, vol. 40, no. 10, pp. 606–608, May 2004.

[27] C. Oestges, B. Clerckx, and D. Vanhoenacker-Janvier A. J. Paulraj, "Impact of diagonal correlations on MIMO capacity: Application to geometrical scattering models," *Proc. of Vehic. Techn. Conf. (VTC'03)*, vol. 1, pp. 394–398, Oct. 2003.

[28] M. Kang and M. S. Alouini, "On the capacity of MIMO Rician channels," in *Proc. 40th Allerton Conf. on Communication, Control, and Computing*, Oct. 2002.

[29] G. Lebrun, M. Faulkner, M. Shafi, and P. J. Smith, "MIMO Ricean channel capacity," *Proc. of IEEE Int. Conf. In Communications (ICC'04)*, pp. 2939–2943, June 2004.

[30] G. Alfano, A. Lozano, A. M. Tulino, and S. Verdú, "Mutual information and eigenvalue distribution of MIMO Ricean channels," in *Proc. of Int. Symp. on Inform. Theory & Applications (ISITA'04)*, Oct. 2004.

[31] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. on Inform. Theory*, vol. 48, pp. 3117–3128, Dec. 2002.

[32] E. Biglieri, A. M. Tulino, and G. Taricco, "How far is infinity ? Using asymptotic analyses in multi-antenna systems," *Intern. Symp. on Spread Spectrum Techn. & Applic. (ISSSTA'02)*, vol. 1, pp. 1–6, Sept. 2002.

[33] S. Shamai and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. on Inform. Theory*, vol. 47, no. 5, pp. 1302–1327, May 2001.

[34] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communication," *Proc. of Intern. Symp. on Inform. Theory (ISIT'04)*, p. 288, July 2004.

[35] A. M. Tulino, S. Verdú, and A. Lozano, "Capacity of antenna arrays with space, polarization and pattern diversity," *Proc. IEEE Inform. Theory Workshop (ITW'03)*, pp. 324–327, Mar. 2003.

[36] O. Somekh, B. M. Zaidel, and S. Shamai, "Spectral efficiency of joint multiple cell-site processors for randomly spread DS-CDMA systems," in *Proc. of IEEE Intern. Symp. on Information Theory (ISIT'04)*, June 2004, p. 278.

[37] D. Aktas, M. N. Bacha, J. Evans, and S. Hanly, "On the sum capacity of multiuser MIMO channels," in *Proc. of IEEE Intern. Symp. on Information Theory & Applications (ISITA'04)*, Oct. 2004.

[38] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[39] V. L. Girko, *Theory of Random Determinants*, Kluwer Academic Publishers, 1990.

[40] A. M. Tulino, L. Li, and S. Verdú, "Spectral efficiency of multicarrier CDMA," *IEEE Trans. on Inform. Theory*, vol. 51, no. 2, pp. 479–505, Feb. 2005.

[41] S. V. Hanly and D. N. C. Tse, "Resource pooling and effective bandwidths in CDMA networks with multiuser receivers and spatial diversity," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1328–1351, May 2004.

[42] V. Veeravalli, Y. Liang, and A. Sayeed, "Correlated MIMO Rayleigh fading channels: Capacity, optimal signaling and scaling laws," *IEEE Trans. on Information Theory (to appear)*, 2005.

[43] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. of Multivariate Analysis*, vol. 54, pp. 175–192, 1995.

[44] D. N. C. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. on Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.

[45] R. R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels," *IEEE Trans. on Information Theory*, vol. 48, no. 7, pp. 2086–2091, July 2002.

[46] V. A. Marcenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

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

[48] K. Liu, V. Raghavan, and A. M. Sayeed, "Capacity scaling and spectral efficiency in wideband correlated MIMO channels," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 2504–2526, Oct. 2003.

[49] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. on Wireless Communications (to appear)*, 2005.

[50] D. Hoesli and A. Lapidoth, "How good is an isotropic input on a MIMO Ricean channel?," *Proc. of IEEE Int. Symp. on Inform. Theory (ISIT'04)*, July 2004.

[51] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Communications*, vol. 3, no. 5, pp. 1543–1553, Sept. 2004.

[52] J. Kotecha and A. Sayeed, "On the capacity of correlated MIMO channels," *Proc. of Intern. Symp. on Inform. Theory (ISIT'03)*, July 2003.

[53] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[54] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 2632–2639, Sept. 2001.

[55] A. Lozano, A. M. Tulino, and S. Verdú, "Multiantenna capacity: myths and realities," *Space-time Wireless Systems: From array processing to MIMO Communications (*Chapter 8), Cambridge University Press, 2005.

[56] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. on Inform. Theory*, Apr. 2005.

[57] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communication," *Int. Conf. in Communications (ICC'05)*, May 2005.

[58] T.-S. Chu and L. J. Greenstein, "A semiempirical representation of antenna diversity gain at cellular and PCS base stations," *IEEE Trans. Commun.*, vol. 45–46, pp. 644–646, June 1997.

[59] A. Sayeed, V. Raghavan, and J. Kotecha, "Capacity of space-time wireless channels: A physical perspective," *Inform. Theory Workshop (ITW'04)*, Oct. 2004.

[60] D. Gesbert, "Multipath: curse or blessing? a system performance analysis of MIMO wireless systems," *Proc. of 2004 International Zurich Seminar on Communications*, pp. 14–17, Jan. 2004.

[61] A. Nica, D. Shlyakhtenko, and R. Speicher, "Operator-valued distributions. I. characterizations of freeness," *IMRN International Mathematics Research Notices*, vol. 2002, no. 29, pp. 1509–1538.

[62] P. Billingsley, *Probability and measure*, New York, Wiley, 1986.

[63] E. Jorswieck and H. Boche, "Average mutual information in spatial correlated MIMO systems with uninformed transmitter," *Proc. of Conf. on Inform. Sciences and Systems (CISS'04)*, pp. 30–35, Mar. 2004.