MIMO Capacity with Channel State Information at the Transmitter

Antonia M. Tulino
Universitá di Napoli "Federico II"
Naples 80125, Italy
Email: atulino@ee.princeton.edu

Angel lozano
Bell Labs (Lucent Technologies)
Holmdel, NJ 07733, USA
Email: aloz@lucent.com

Sergio Verdú
Princeton University
Princeton, NJ 08544, USA
Email: verdu@princeton.edu

Abstract— **This paper presents analytical expressions for the capacity of single-user multi-antenna channels, known instantaneously by both transmitter and receiver, at moderate and high signal-to-noise ratios. Fading channels with both uncorrelated and correlated antennas are encompassed. The characterization is conducted primarily in the limit of large numbers of antennas, with accompanying examples that illustrate the validity of the results for even small numbers thereof. In the absence of correlation, the capacity is also tightly bounded for fixed numbers of antennas with compact closed-form expressions.**

I. Introduction

The capacity of a MIMO (multi-input multi-output) channel is influenced by the degree of CSI (channel-state information) available to both transmitter and receiver. In most instances of multi-antenna communication, the receiver can accurately track the instantaneous state of the channel from pilot signals that are typically embedded within the transmissions. In terms of CSI at the transmitter, on the other hand, several scenarios are possible:

- In frequency-duplexed systems, where uplink and downlink are apart in frequency, the link fading is not reciprocal and thus the CSI must be conveyed through feedback, which may incur round-trip delays that are nonnegligible with respect to the coherence time of the CSI being reported. Consequently, the transmitter is usually deprived of instantaneous CSI.
- In time-duplexed systems, in contrast, the links are reciprocal as long as the coherence time of the fading process exceeds the duplex time.[1] Thus, the transmitter may have access to reliable CSI at low and moderate levels of mobility.
- At high levels of mobility, even in time-duplexed systems the CSI becomes rapidly outdated.

In terms of the characterization of the single-user capacity, these various scenarios are usually mapped onto distinct operational regimes:

(a) The transmitter has instantaneous CSI.
(b) The transmitter has only statistical CSI.
(c) The transmitter has no CSI.

For regimes (b) and (c), closed-form expressions are available for Rayleigh-faded channels whose entries are IID (independent identically distributed), both for arbitrary numbers of antennas [1], [2] and asymptotically in the number of antennas [3], [4]. Also for correlated Rayleigh-faded channels, analytical characterizations in these regimes abound, for arbitrary numbers of antennas [5]–[7] and asymptotically therein [8]–[11]. These analyses reveal how the corresponding capacities are impacted by antenna correlation. The asymptotic expressions, specifically, turn out to be particularly insightful and very accurate even for small numbers of antennas [12].

For regime (a), on the other hand, the capacity had been analyzed only for channels with IID entries, for arbitrary numbers of antennas in [13] and asymptotically in [14].[2] In this paper, we present a more extensive analysis that encompasses correlated channels. Specifically:

- For arbitrary numbers of antennas, we tightly bound the capacity of uncorrelated Rayleigh-faded channels. The bound, in compact closed-form, becomes exact at high signal-to-noise ratio.
- Asymptotically in the number of antennas, we find expressions for the capacity of Rayleigh-faded channels with either correlated or uncorrelated entries.

II. Definitions and Models

With frequency-flat fading, the baseband complex model we consider is

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

where $\mathbf{x}$ and $\mathbf{y}$ are the input and output vectors while $\mathbf{n}$ is white Gaussian noise. The channel is represented by the zero-mean random matrix $\sqrt{g}\,\mathbf{H}$ where the scalar g is such that

$$E[\mathrm{Tr}\{\mathbf{H}\mathbf{H}^\dagger\}] = n_\mathrm{R} n_\mathrm{T}. \tag{1}$$

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

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

[1] The reciprocity applies to the radio channel but not necessarily to the filters and amplifiers at transmitter and receiver. Differences therein may require careful calibration.

[2] The explicit expression given in [13] for arbitrary numbers of antennas is function of a parameter that must be solved for numerically.

where the normalization ensures that $E[\mathrm{Tr}\{\mathbf{\Phi}\}]{=}n_{\mathrm{T}}$. From (1) and (2), we can define the signal-to-noise ratio

$$\mathsf{SNR} = g\frac{E[\|\mathbf{x}\|^2]}{\frac{1}{n_{\mathrm{R}}}E[\|\mathbf{n}\|^2]}.$$

In terms of the correlation between the entries of $\mathbf{H}$, we adhere to the widely used *separable* model whereby the correlation between the (i,j) and (i',j') entries is expressed as [15]

$$E\left[(\mathbf{H})_{i,j}(\mathbf{H})^*_{i',j'}\right] = (\mathbf{\Theta}_{\mathrm{R}})_{i,i'}(\mathbf{\Theta}_{\mathrm{T}})_{j,j'}$$

where $\mathbf{\Theta}_{\mathrm{R}}$ and $\mathbf{\Theta}_{\mathrm{T}}$ are $(n_{\mathrm{R}}{\times}n_{\mathrm{R}})$ and $(n_{\mathrm{T}}{\times}n_{\mathrm{T}})$ correlation matrices whose entries indicate the correlation between receive antennas and between transmit antennas, respectively, while $(\cdot)_{i,j}$ denotes the (i,j)-th entry of a matrix.

III. Capacity with CSI at the Transmitter

The unique capacity-achieving input is zero-mean Gaussian. It is useful to decompose its covariance as $\mathbf{\Phi}{=}\mathbf{VPV}^\dagger$ where $\mathbf{V}$ contains the eigenvectors while $\mathbf{P}{=}\mathrm{diag}\{p_1, p_2, \ldots, p_{n_{\mathrm{T}}}\}$ holds the eigenvalues, each signifying the normalized power allocated to the corresponding eigenvector.

With instantaneous CSI, $\mathbf{\Phi}$ can be made a function of $\mathbf{H}$. Capacity is achieved by signalling with $\mathbf{V}{=}\mathbf{H}^\dagger\mathbf{H}$ thus creating a set of orthogonal parallel channels [16], [1]. The corresponding power allocation, $\mathbf{P}$, is obtained via waterfill over the eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ [17]. Hence,

$$p_j = \left[\nu - \frac{n_{\mathrm{T}}}{\mathsf{SNR}\,\lambda_j(\mathbf{H}^\dagger\mathbf{H})}\right]^+ \tag{3}$$

where ν is such that $\mathrm{Tr}\{\mathbf{P}\}{=}n_{\mathrm{T}}$ whereas $\lambda_j(\cdot)$ denotes the j-th eigenvalue of a matrix and $[z]^+{=}\max(0,z)$. The ergodic capacity that results is [1]

$$\begin{aligned} C(\mathsf{SNR}) &= E\left[\log_2\det\left(\mathbf{I} + \tfrac{\mathsf{SNR}}{n_{\mathrm{T}}}\mathbf{\Phi}\mathbf{H}^\dagger\mathbf{H}\right)\right] \\ &= E\left[\sum_{j=1}^{n_{\mathrm{T}}}\log_2\left(1+\tfrac{\mathsf{SNR}}{n_{\mathrm{T}}}p_j\lambda_j(\mathbf{H}^\dagger\mathbf{H})\right)\right] \\ &= E\left[\sum_{j=1}^{n_{\mathrm{T}}}\left[\log_2\left(\mathsf{SNR}\,\nu\lambda_j\left(\tfrac{\mathbf{H}^\dagger\mathbf{H}}{n_{\mathrm{T}}}\right)\right)\right]^+\right] \end{aligned} \tag{4}$$

with expectation over the distribution of $\mathbf{H}$. Since the nonzero eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ coincide with those of $\mathbf{H}\mathbf{H}^\dagger$, the capacity is unaffected by a reversal of roles between transmitter and receiver [1]. Also noteworthy is that, for $\mathsf{SNR}{\to}\infty$, the waterfill process results in a uniform power allocation over the eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ that are not identically zero and hence capacity analyses conducted with such an input become applicable.

In order to evaluate (4) asymptotically, we will need to introduce the empirical cumulative distribution of the eigenvalues of $\frac{1}{n_{\mathrm{T}}}\mathbf{H}^\dagger\mathbf{H}$, defined as

$$\mathsf{F}^{n_{\mathrm{T}}}_{\mathbf{H}^\dagger\mathbf{H}}(z) = \frac{1}{n_{\mathrm{T}}}\sum_{j=1}^{n_{\mathrm{T}}} 1\left\{\lambda_j\left(\tfrac{\mathbf{H}^\dagger\mathbf{H}}{n_{\mathrm{T}}}\right)\le z\right\} \tag{5}$$

where $1\{\cdot\}$ is the indicator function. For every channel that we shall consider, as n_{T} and n_{R} are driven to infinity with some constant ratio $\beta{=}\frac{n_{\mathrm{T}}}{n_{\mathrm{R}}}$, $\mathsf{F}^{n_{\mathrm{T}}}_{\mathbf{H}^\dagger\mathbf{H}}(\cdot)$ converges almost surely to a nonrandom limit that we note by $\mathsf{F}_{\mathbf{H}^\dagger\mathbf{H}}(\cdot)$. Accordingly, for $n_{\mathrm{T}}, n_{\mathrm{R}}{\to}\infty$ the capacity per receive antenna converges to

$$\tfrac{1}{n_{\mathrm{R}}}C(\mathsf{SNR}) \to \beta\int[\log_2(\mathsf{SNR}\,\nu\lambda)]^+ d\mathsf{F}_{\mathbf{H}^\dagger\mathbf{H}}(\lambda) \tag{6}$$

IV. Analytical Characterization

The analysis of (4) and (6) is greatly facilitated whenever $p_j{>}0\ \forall j$, i.e., when the waterfill process allocates power to every signalling eigenvector. This condition is satisfied above a certain SNR threshold. Nonetheless, the expressions derived under this premise usually cover—as will be illustrated—most of the SNR range of interest to wireless communication applications. The exception lies in channels with $n_{\mathrm{T}}{=}n_{\mathrm{R}}$, for which this condition is only satisfied at high SNR.

A. Uncorrelated Channels

Let us begin by evaluating the capacity of channels with IID entries. For fixed n_{T} and n_{R}, we present the following result.

Proposition 1: Consider a Rayleigh-faded channel with IID entries. The capacity (bits/s/Hz) with instantaneous CSI at the transmitter satisfies, for $n_{\mathrm{T}}{<}n_{\mathrm{R}}$,

$$C \le n_{\mathrm{T}}\log_2\left(\tfrac{\mathsf{SNR}}{n_{\mathrm{T}}} + \tfrac{1}{n_{\mathrm{R}}-n_{\mathrm{T}}}\right) + \left(\sum_{k=1}^{n_{\mathrm{T}}}\sum_{\ell=1}^{n_{\mathrm{R}}-k}\tfrac{1}{\ell} - n_{\mathrm{T}}\gamma\right)\log_2 e$$

and, for $n_{\mathrm{T}}{>}n_{\mathrm{R}}$,

$$C \le n_{\mathrm{R}}\log_2\left(\tfrac{\mathsf{SNR}}{n_{\mathrm{R}}} + \tfrac{1}{n_{\mathrm{T}}-n_{\mathrm{R}}}\right) + \left(\sum_{k=1}^{n_{\mathrm{R}}}\sum_{\ell=1}^{n_{\mathrm{T}}-k}\tfrac{1}{\ell} - n_{\mathrm{R}}\gamma\right)\log_2 e$$

where γ is the Euler-Mascheroni constant, $\gamma{\approx}0.5772$. For $n_{\mathrm{T}}{=}n_{\mathrm{R}}{=}n$ and high SNR, in turn,

$$C = n\log_2\tfrac{\mathsf{SNR}}{n} + n\left(\sum_{\ell=2}^{n}\tfrac{1}{\ell} - \gamma\right)\log_2 e + O(\tfrac{1}{\mathsf{SNR}})$$

Proof: See Appendix A.

Before proceeding with the analysis, we illustrate the tightness of this bound.

Example 1: Depicted in Fig. 1 is $C(\mathsf{SNR})$ for several numbers of transmit and receive antennas, both from Proposition 1 and from Montecarlo simulations. Because of reciprocity, results for $(n_{\mathrm{T}}{\times}n_{\mathrm{R}})$ apply also to $(n_{\mathrm{R}}{\times}n_{\mathrm{T}})$. Note the tight correspondence between the closed-form expressions and the simulations, for a wide range of SNR and numbers of antennas.

Asymptotically in the number of antennas, on the other hand, the following applies. (The same asymptotic characterization was undertaken in [14]).

Theorem 1: Consider a channel with zero-mean IID entries, arbitrarily distributed. Let $n_{\mathrm{T}}, n_{\mathrm{R}}{\to}\infty$ with $\beta{=}\frac{n_{\mathrm{T}}}{n_{\mathrm{R}}}$ and define

$$\mathsf{SNR}_0 = \frac{2\min(1,\beta^{3/2})}{|1-\sqrt{\beta}||1-\beta|}$$

Fig. 1. $C(\mathsf{SNR})$ for an uncorrelated Rayleigh-faded channel with several numbers of transmit and receive antennas: (2×4), (2×6) and (4×6). Solid lines indicate analytical upper bound, circles indicate simulation.

Fig. 2. $C(\mathsf{SNR})$ for an uncorrelated Rayleigh-faded channel with several numbers of antennas: (2×4), (2×6) and (4×6). Solid lines indicate scaled asymptotic expressions evaluated at $\beta{=}\frac{n_{\rm T}}{n_{\rm R}}$, circles indicate simulation.

For $\mathsf{SNR}{\geq}\mathsf{SNR}_0$, the capacity with instantaneous CSI at the transmitter is, for $\beta{<}1$,

$$\frac{1}{n_{\rm R}} C \rightarrow \beta \log_2\left(\frac{\mathsf{SNR}}{\beta} + \frac{1}{1{-}\beta}\right) + (1{-}\beta)\log_2 \frac{1}{1{-}\beta} - \beta \log_2 e$$

while, for $\beta{>}1$,

$$\frac{1}{n_{\rm R}} C \rightarrow \log_2\left(\beta\,\mathsf{SNR} + \frac{\beta}{\beta{-}1}\right) + (\beta{-}1)\log_2 \frac{\beta}{\beta{-}1} - \log_2 e$$

For $\beta{=}1$ and high SNR,

$$\frac{1}{n_{\rm R}} C \rightarrow \log_2 \frac{\mathsf{SNR}}{e} + O\left(\frac{1}{\mathsf{SNR}}\right)$$

Proof: See Appendix B.

In this case, the fading is not constrained to be Rayleigh. Rather, the expressions are valid if only the entries of the channel matrix are zero-mean with uniformly bounded variances.

Example 2: The applicability of Theorem 1 is exemplified in Fig. 2, for the same antenna numbers and SNR range used in Example 1. The expressions in the theorem are evaluated with the role of β played by $\frac{n_{\rm T}}{n_{\rm R}}$ and contrasted with Montecarlo simulations. On each curve, the threshold SNR_0 is also indicated. Notice how the asymptotic expressions, scaled by the corresponding number of receive antennas, approximate very closely the actual capacities well below the nominal threshold.

B. Correlated Rayleigh-faded Channels

Let us now extend the asymptotic analysis to channels with transmit and receive correlations. In terms of the correlation matrices, $\mathbf{\Theta}_{\rm T}$ and $\mathbf{\Theta}_{\rm R}$, only their eigenvalues are relevant to the capacity. We thus define their empirical eigenvalue distributions as we did in (5), i.e., for a generic $(n{\times}n)$ correlation matrix $\mathbf{\Theta}$,

$$\mathsf{F}^n_{\mathbf{\Theta}}(z) = \frac{1}{n} \sum_{j=1}^{n} 1\left\{\lambda_j\left(\mathbf{\Theta}\right) \leq z\right\} \tag{7}$$

which, as $n{\rightarrow}\infty$, converges to $\mathsf{F}_{\mathbf{\Theta}}(\cdot)$. The following is the main result in the paper.

Theorem 2: Consider a Rayleigh-faded channel whose nonsingular transmit and receive correlation matrices are $\mathbf{\Theta}_{\rm T}$ and $\mathbf{\Theta}_{\rm R}$. Define $\mathsf{\Lambda}_{\rm T}$ and $\mathsf{\Lambda}_{\rm R}$ as variables whose respective distributions are $\mathsf{F}_{\mathbf{\Theta}_{\rm T}}(\cdot)$ and $\mathsf{F}_{\mathbf{\Theta}_{\rm R}}(\cdot)$. For $\mathsf{SNR}{\geq}\mathsf{SNR}_0$, if $\beta{<}1$,

$$\frac{1}{n_{\rm R}} C \rightarrow E\left[\log_2\left(1 + \frac{\mathsf{\Lambda}_{\rm R}}{\varphi}\right)\right] + \beta \log_2\left(\mathsf{SNR} + \frac{1}{\varphi} E\left[\frac{1}{\mathsf{\Lambda}_{\rm T}}\right]\right) + \beta E\left[\log_2 \frac{\varphi \mathsf{\Lambda}_{\rm T}}{e}\right]$$

while, if $\beta{>}1$,

$$\frac{1}{n_{\rm R}} C \rightarrow \beta E[\log_2(1 + \alpha \mathsf{\Lambda}_{\rm T})] + \log_2\left(\mathsf{SNR} + \alpha E\left[\frac{1}{\mathsf{\Lambda}_{\rm R}}\right]\right) + E\left[\log_2 \frac{\mathsf{\Lambda}_{\rm R}}{\alpha e}\right]$$

with the expectations taken over $\mathsf{\Lambda}_{\rm T}$ and $\mathsf{\Lambda}_{\rm R}$ while the parameters φ and α are solutions to

$$E\left[\frac{1}{1 + \frac{\mathsf{\Lambda}_{\rm R}}{\varphi}}\right] = 1 - \beta \qquad E\left[\frac{1}{1 + \alpha \mathsf{\Lambda}_{\rm T}}\right] = 1 - \frac{1}{\beta}.$$

For $\beta{=}1$ and high SNR,

$$\frac{1}{n_{\rm R}} C \rightarrow \log_2 \frac{\mathsf{SNR}}{e} + E[\log_2 \mathsf{\Lambda}_{\rm T}] + E[\log_2 \mathsf{\Lambda}_{\rm R}] + O\left(\frac{1}{\mathsf{SNR}}\right)$$

The threshold SNR_0 is, for $\beta{<}1$,

$$\mathsf{SNR}_0 = \frac{1}{\kappa} - \frac{1}{\varphi} E\left[\frac{1}{\mathsf{\Lambda}_{\rm T}}\right] \tag{8}$$

with κ the infimum of the support of $\mathsf{F}_{\mathbf{H}^\dagger \mathbf{H}}(\cdot)$ while,[3] for $\beta{>}1$,

$$\mathsf{SNR}_0 = \frac{1}{\beta\,\kappa} - \alpha E\left[\frac{1}{\mathsf{\Lambda}_{\rm R}}\right] \tag{9}$$

Proof: See Appendix B.

[3]Results on the support of $\mathsf{F}_{\mathbf{H}\mathbf{H}^\dagger}$ can be found in [18] and [19].

Fig. 3. $C(\mathsf{SNR})$ for a correlated Rayleigh-faded channel with $n_{\rm T}{=}2$ and $n_{\rm R}{=}4$. Also shown is the corresponding capacity with no correlation. Solid lines indicate scaled asymptotic expressions evaluated non-asymptotically, circles indicate simulation.

Example 3: Let $n_{\rm T}{=}2$ with correlation

$$(\mathbf{\Theta}_{\rm T})_{i,j} = e^{-0.2\,(i-j)^2}$$

which corresponds to a 2-wavelength antenna separation and a broadside (truncated) Gaussian power azimuth spectrum with $2°$ root-mean-square spread [20]. Further let $n_{\rm R}{=}4$ with

$$(\mathbf{\Theta}_{\rm R})_{i,j} = J_0(\pi|i-j|) \tag{10}$$

where $J_0(\cdot)$ is the zero-order bessel function of the first kind. The correlation structure in (10) is representative of a linear array with 1-wavelength antenna spacing and a uniform power azimuth spectrum. Theorem 2 yields

$$\begin{aligned} C \approx & \sum_{i=1}^{n_{\rm R}} \log_2\!\left(1+\frac{\lambda_i(\mathbf{\Theta}_{\rm R})}{\varphi}\right) + n_{\rm T}\log_2\!\left(\mathsf{SNR}+\frac{1}{\varphi}\sum_{j=1}^{n_{\rm T}}\frac{1}{\lambda(\mathbf{\Theta}_{\rm T})}\right) \\ & + \sum_{j=1}^{n_{\rm T}} \log_2 \frac{\varphi\lambda_j(\mathbf{\Theta}_{\rm T})}{e} \end{aligned} \tag{11}$$

where φ is obtained from

$$\sum_{i=1}^{n_{\rm R}} \frac{1}{1+\frac{\lambda_i(\mathbf{\Theta}_{\rm R})}{\varphi}} = n_{\rm R}-n_{\rm T}.$$

Eq. (11) is plotted in Fig. 3 alongside Montecarlo simulations. The correspondence is excellent above 4-5 dB.

Although, in their full generality, the expressions in Theorem 2 are in the form of fixed-point solutions, in many cases of interest they become explicit. Specifically, this is the case whenever correlation takes place only at the end of the link with the fewest antennas. If $\beta{<}1$ and $\mathbf{\Theta}_{\rm R}{=}\mathbf{I}$,

$$\begin{aligned} \frac{1}{n_{\rm R}}C \to & \beta E\!\left[\log_2 \frac{\Lambda_{\rm T}}{e}\right] + \beta\log_2\!\left(\mathsf{SNR}\,\frac{1-\beta}{\beta} + E\!\left[\frac{1}{\Lambda_{\rm T}}\right]\right) \\ & -\log_2(1-\beta) \end{aligned}$$

whereas, if $\beta{>}1$ and $\mathbf{\Theta}_{\rm T}{=}\mathbf{I}$,

$$\begin{aligned} \frac{1}{n_{\rm R}}C \to & E\!\left[\log_2 \frac{\Lambda_{\rm R}}{e}\right] + \log_2\!\left(\mathsf{SNR}(\beta-1) + E\!\left[\frac{1}{\Lambda_{\rm R}}\right]\right) \\ & -\beta\log_2(1-\tfrac{1}{\beta}) \end{aligned}$$

V. Conclusions

Various analytical expressions for the capacity of multi-antenna channels with instantaneous CSI at both transmitter and receiver have been derived, with and without antenna correlation. These expressions are valid for SNR levels above a certain threshold that depends on the numbers of antennas and their correlation. (The most constrained instance is that of an equal number of transmitters and receivers, for which the validity is restricted to high SNR.) Specific observations that can be made are:

- For $n_{\rm T}{=}n_{\rm R}{=}n$ at high SNR,

  $$C \approx n\log_2\frac{\mathsf{SNR}}{e} + \sum_{j=1}^{n}\log_2\lambda_j(\mathbf{\Theta}_{\rm T}) + \sum_{i=1}^{n}\log_2\lambda_i(\mathbf{\Theta}_{\rm R})$$

  which, via Jensen's inequality, indicates that correlation can only diminish the high-SNR capacity, as already observed in [8]. The corresponding expressions for $n_{\rm T}{\neq}n_{\rm R}$ further reveal that this is also the case if correlation takes place only at the end of the link with the fewest antennas.
- With no correlation and $n_{\rm T}{\ll}n_{\rm R}$,

  $$C \approx n_{\rm T}\log_2\left(\frac{n_{\rm R}}{n_{\rm T}}\mathsf{SNR}\right) + O(\tfrac{n_{\rm T}}{n_{\rm R}})$$

  which coincides with the corresponding behavior if the transmitter radiates an isotropic signal [21]. In this case, therefore, instantaneous CSI at the transmitter yields no first-order advantage.
- Conversely, with no correlation and $n_{\rm T}{\gg}n_{\rm R}$,

  $$C \approx n_{\rm R}\log_2\left(\frac{n_{\rm T}}{n_{\rm R}}\mathsf{SNR}\right) + O(\tfrac{n_{\rm R}}{n_{\rm T}})$$

  whereas, with an isotropic input [21],

  $$C \approx n_{\rm R}\log_2(1+\mathsf{SNR}) + O(\tfrac{n_{\rm R}}{n_{\rm T}})$$

  from which the first-order value of CSI can be quantified.

Appendix A

Consider first $n_{\rm T}{<}n_{\rm R}$. Under the condition that $p_j{>}0\ \forall j$, (3) leads to

$$\nu = 1 + \frac{1}{\mathsf{SNR}}\mathrm{Tr}\left\{\left(\mathbf{H}^\dagger\mathbf{H}\right)^{-1}\right\} \tag{12}$$

which, plugged into (4), yields

$$\begin{aligned} C \;=\; & n_{\rm T}E\left[\log_2\left(\frac{\mathsf{SNR}}{n_{\rm T}}+\frac{1}{n_{\rm T}}\mathrm{Tr}\left\{\left(\mathbf{H}^\dagger\mathbf{H}\right)^{-1}\right\}\right)\right] \\ & +E\left[\log_2\det(\mathbf{H}^\dagger\mathbf{H})\right] \end{aligned} \tag{13}$$

The second expectation in (13) is given by [22]

$$E\left[\log_2\det\left(\mathbf{H}^\dagger\mathbf{H}\right)\right] = \sum_{\ell=0}^{n_{\rm T}-1}\psi(n_{\rm R}-\ell)\,\log_2 e \tag{14}$$

with $\psi(\cdot)$ the digamma function [23]

$$\psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}.$$

The first expectation in (13), in turn, is upper-bounded by

$$n_{\mathrm{T}} \log_2 \left(\frac{\mathsf{SNR}}{n_{\mathrm{T}}} + \frac{1}{n_{\mathrm{T}}} E\left[\mathrm{Tr}\left\{ \left(\mathbf{H}^\dagger \mathbf{H} \right)^{-1} \right\} \right] \right) \tag{15}$$

where [24, Lemma 6]

$$E\left[\mathrm{Tr}\left\{ \left(\mathbf{H}^\dagger \mathbf{H} \right)^{-1} \right\} \right] = \frac{n_{\mathrm{T}}}{n_{\mathrm{R}} - n_{\mathrm{T}}} \tag{16}$$

Plugging (14), (15) and (16) into (13), the bound is found. For $n_{\mathrm{T}} > n_{\mathrm{R}}$, reciprocity can be exploited by reversing the roles of n_{T} and n_{R}. For $n_{\mathrm{T}} = n_{\mathrm{R}}$, both bounds coincide but they are tight only at high SNR.

APPENDIX B

Let us first sketch the proof of Theorem 2 for $\beta < 1$. For $\mathsf{SNR} > \mathsf{SNR}_0$, all the input powers are nonzero and (13) yields

$$\begin{aligned} \frac{1}{n_{\mathrm{R}}} C &= \beta E\left[\log_2 \left(\mathsf{SNR} + \mathrm{Tr}\left\{ \left(\mathbf{H}^\dagger \mathbf{H} \right)^{-1} \right\} \right) \right] \\ &\quad + \frac{1}{n_{\mathrm{R}}} E\left[\log_2 \det \left(\frac{\mathbf{H}^\dagger \mathbf{H}}{n_{\mathrm{T}}} \right) \right] \end{aligned} \tag{17}$$

With the correlation matrices nonsingular, $\mathrm{Tr}\{(\mathbf{H}^\dagger \mathbf{H})^{-1}\}$ in (17) is equivalent to

$$\frac{1}{n_{\mathrm{T}}} \sum_{j=1}^{n_{\mathrm{T}}} \frac{1}{\lambda_j\left(\frac{\mathbf{H}^\dagger \mathbf{H}}{n_{\mathrm{T}}}\right)} = \lim_{\mathsf{SNR} \to \infty} \frac{1}{n_{\mathrm{T}}} \sum_{j=1}^{n_{\mathrm{T}}} \frac{\mathsf{SNR}}{1 + \mathsf{SNR}\lambda_j\left(\frac{\mathbf{H}^\dagger \mathbf{H}}{n_{\mathrm{T}}}\right)}$$

which, as $n_{\mathrm{T}}, n_{\mathrm{R}} \to \infty$, converges almost surely to [25]

$$\lim_{\mathsf{SNR} \to \infty} \mathsf{SNR}\, E\left[\left(1 + \mathsf{SNR} \beta E\left[\frac{\Lambda_{\mathrm{R}} \Lambda_{\mathrm{T}}}{1 + \frac{\Lambda_{\mathrm{R}}}{\varphi}} | \Lambda_{\mathrm{T}} \right] \right)^{-1} \right] = \frac{1}{\varphi} E\left[\frac{1}{\Lambda_{\mathrm{T}}} \right] \tag{18}$$

with φ satisfying the equation in the claim. The second term in (17), in turn, converges almost surely to [25]

$$\frac{1}{n_{\mathrm{R}}} E\big[\log_2 \det(\mathbf{H}^\dagger \mathbf{H})\big] \to E\left[\log_2\left(1 + \frac{\Lambda_{\mathrm{R}}}{\varphi}\right)\right] + \beta E\left[\log_2 \frac{\varphi \Lambda_{\mathrm{T}}}{e}\right]$$

with α as in the claim.

For $\beta > 1$, the same solution applies with β and $\frac{1}{\varphi}$ replaced by $\frac{1}{\beta}$ and α, respectively, and with Λ_{T} and Λ_{R} interchanged.

For $\beta = 1$, the high-SNR behavior coincides with the one observed with no CSI at the transmitter [8].

To find SNR_0, we combine (3) and (12) obtaining

$$\mathsf{SNR} + \frac{1}{n_{\mathrm{T}}} \mathrm{Tr}\left\{ \left(\frac{\mathbf{H}^\dagger \mathbf{H}}{n_{\mathrm{T}}} \right)^{-1} \right\} = \frac{1}{\lambda_{\min}\left(\frac{\mathbf{H}^\dagger \mathbf{H}}{n_{\mathrm{T}}} \right)}$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue. The second term on the left-hand side converges asymptotically to (18) while the smallest eigenvalue converges to κ. For $\beta > 1$, SNR_0 is found reciprocally using the relationship $\mathsf{F}_{\mathbf{H}\mathbf{H}^\dagger}(z) = \frac{1}{\beta} \mathsf{F}_{\mathbf{H}^\dagger \mathbf{H}}(z)$ for $z > 0$.

Theorem 1 follows from Theorem 2 with $\Lambda_{\mathrm{T}} = 1$ and $\Lambda_{\mathrm{R}} = 1$. The explicit thresholds SNR_0 emerge from those in Theorem 2 with $\kappa = (\frac{1}{\sqrt{\beta}} - 1)^2$ [25].

REFERENCES

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

[2] 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.

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

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

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

[6] M. Kiessling and J. Speidel, "Exact ergodic capacity of MIMO channels in correlated Rayleigh fading environments," in *Proc. Intern. Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, Feb. 2004.

[7] G. Alfano, A. Lozano, A. M. Tulino, and S. Verdú, "Capacity of MIMO channels with one-sided correlation," in *Proc. of IEEE Intern. Symp. on Spread Spectrum Tech. and Applications (ISSSTA'04)*, Aug. 2004.

[8] 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.

[9] A. M. Tulino, S. Verdú, and A. Lozano, "Capacity of antenna arrays with space, polarization and pattern diversity," *Inform. Theory Workshop (ITW'03), Paris, France*, Apr. 2003.

[10] X. Mestre, J. Fonollosa, and A. Pages-Zamora, "Capacity of MIMO channels: asymptotic evaluation under correlated fading," *IEEE J. on Sel. Areas in Communications*, vol. 21, no. 5, pp. 829– 838, June 2003.

[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, no. 10, pp. 2545–2561, Oct 2003.

[12] 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)*, Sept. 2002.

[13] S. K. Jayaweera and H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," *IEEE Trans. on Inform. Theory*, vol. 49, no. 10, pp. 2697–2709, Oct. 2003.

[14] A. Grant, "Rayleigh fading multiple-antenna channels," *EURASIP Journal on Applied Signal Processing, Special Issue on Space-Time Coding (Part I)*, vol. 2002, no. 3, pp. 316–329, Mar. 2002.

[15] 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. on Communications*, vol. 48, no. 3, Mar. 2000.

[16] B. S. Tsybakov, "The capacity of a memoryless Gaussian vector channel," *Problems of Inform. Transmission*, vol. 1, no. 1, pp. 18–29, 1965.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, Wiley, 1990.

[18] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *Annals of Probability*, vol. 26, pp. 316–345, 1998.

[19] ——, "Exact separation of eigenvalues of large dimensional sample covariance matrices," *Annals of Probability*, vol. 27, no. 3, pp. 1536–1555, 1999.

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

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

[22] R. J. Muirhead, *Aspects of multivariate statistical theory*. New York, Wiley, 1982.

[23] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, 1972.

[24] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Transactions on Information Theory*, vol. 49, no. 10, October 2003.

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