5

Multiantenna capacity: myths and realities

Angel Lozano

Bell Laboratories

Antonia M. Tulino

Universitá degli Studi di Napoli

Sergio Verdú

Princeton University

Over the last decade, the increases in capacity promised by multiantenna communication techniques have spurred many information-theoretic analyses. Furthermore, information theory has been used as a design tool to optimize the signals fed to the transmit array and to motivate signal processing strategies at the receiver. In this chapter, we catalog a number of misconceptions that have arisen in the multiantenna literature. The focus is on information-theoretic results and their interpretations, rather than on the validity of the various modeling assumptions. After some introductory material, we address several misconceptions about optimum signaling and about the impact of model features, such as antenna correlation, line-of-sight components, and intercell interference. Particular attention is given to the low- and high-power regions. We also briefly touch upon the relationship of the capacity results with some practical transmit and receive architectures. The chapter deals mostly, but not exclusively, with coherent communication.

5.1 Definitions

Denoting by n_{T} and n_{R} the number of transmit and receive antennas, respectively, we shall abide by the frequency-flat complex vector model[†]

$$\mathbf{y} = \sqrt{g}\,\mathbf{H}\mathbf{x} + \mathbf{n}, \tag{5.1}$$

where g is a deterministic scalar that represents the average channel gain, while the random matrix $\mathbf{H}$ is normalized to satisfy

$$E[\mathrm{Tr}\{\mathbf{H}\mathbf{H}^{\dagger}\}] = n_{\mathsf{T}} n_{\mathsf{R}}. \tag{5.2}$$

[†]If the channel is frequency-selective, it can be decomposed into parallel noninteracting channels, each of which conforms to (5.1).

While the distribution of $\mathbf{H}$ is known to both transmitter and receiver, we shall specify when its realization is known to either.

The additive noise is modeled as Gaussian with i.i.d. (independent identically distributed) entries, and†

$$N_0 = \frac{E\left[\|\mathbf{n}\|^2\right]}{n_\mathsf{R}}.$$

The input is constrained such that

$$E\left[\|\mathbf{x}\|^2\right] = P, \tag{5.3}$$

with P the power available for transmission. The normalized input covariance is denoted by

$$\mathbf{\Phi}_\mathbf{x} = \frac{E[\mathbf{x}\mathbf{x}^\dagger]}{P/n_\mathsf{T}},$$

where (5.3) translates into $\mathrm{Tr}\{\mathbf{\Phi}_\mathbf{x}\} = n_\mathsf{T}$. If the input is isotropic, $\mathbf{\Phi}_\mathbf{x} = \mathbf{I}$.

5.2 Multiantenna capacity: the coherent realm

With transmit power P and bandwidth B, the rate R [bits/s] achievable with arbitrary reliability must obey the fundamental limit $R/B \leq C(\mathsf{SNR})$, where $C(\mathsf{SNR})$ is the capacity [bits/s/Hz], and

$$\mathsf{SNR} = g\frac{P}{N_0}.$$

Most existing results on multiantenna capacity apply to the coherent regime, where the channel realization is available at the receiver. A glimpse of several issues concerning the capacity of noncoherent multiantenna channels is given in Section 5.11.

With coherent reception, the input-output mutual information is maximized if the input $\mathbf{x}$ in (5.1) is circularly symmetric complex Gaussian. Conditioned on $\mathbf{H}$, the mutual information [bits/s/Hz] is

$$\mathcal{I}(\mathsf{SNR}, \mathbf{\Phi}_\mathbf{x}) = \log_2 \det\left(\mathbf{I} + \frac{\mathsf{SNR}}{n_\mathsf{T}}\mathbf{H}\mathbf{\Phi}_\mathbf{x}\mathbf{H}^\dagger\right). \tag{5.4}$$

Formula (5.4) was given in the mid 1990s by Foschini (1996) and Telatar (1999). Earlier embodiments of this formula and its generalizations can be found by Pinsker (1964), Tsybakov (1965), Root and Varaiya (1968), Brandenburg and Wyner (1974), Verdú (1986), and Cover and Thomas (1990).

†Any nonsingular noise covariance can be absorbed into the correlation of $\mathbf{H}$. Section 5.9 deals specifically with noise that is colored and subject to fading.

Without knowledge of the realizations of the channel matrices at the transmitter, the capacity is

$$C(\mathsf{SNR}) = \max_{\mathbf{\Phi_x}:\mathrm{Tr}\{\mathbf{\Phi_x}\}=n_\mathsf{T}} E[\mathcal{I}(\mathsf{SNR}, \mathbf{\Phi_x})], \tag{5.5}$$

as long as the sequence of random matrices $\{\mathbf{H}_k\}$ is stationary and ergodic.

When the variation of $\{\mathbf{H}_k\}$ within the horizon of each codeword is not fast enough for the time-averaged mutual information to approach its statistical average, the ergodic capacity (5.5) loses its operational meaning, and the distribution (not just the average) of the mutual information (the so-called outage capacity) has to be considered. In addition to fast-fading channels, the distribution of the mutual information concentrates around its mean in wideband frequency-selective fading channels, and in channels with hybrid ARQ. The ergodic capacity is further operationally meaningful, even if the channel is invariant over the codeword, with achievable-rate feedback.

A particularly attractive ergodic channel, from which plenty of insight can be drawn, is the *canonical* multiantenna channel, where the matrix has i.i.d. Rayleigh-fading entries. The evaluation of the capacity of the canonical channel is facilitated by several facts:

- The isotropy of the capacity-achieving input distribution (Telatar, 1999).
- The availability of an explicit expression for the marginal distribution of the squared singular values of random matrices with Gaussian i.i.d. entries(Bronk, 1965; Telatar, 1999).
- The availability of an explicit asymptotic expression for the squared singular value distribution of matrices with i.i.d. entries whose size goes to infinity with a constant aspect ratio (Marčenko and Pastur, 1967). This asymptotic expression, furthermore, does not depend on the marginal distribution of the matrix entries.

Using these expressions, closed-form formulas for the capacity of the canonical channel were derived, first asymptotically in the numbers of antennas (Verdú and Shamai, 1999; Rapajic and Popescu, 2000) and then for arbitrary numbers thereof (Shin and Lee, 2003).

5.3 Input optimization

5.3.1 Transmitter side information

If the channel realization is known at the transmitter, capacity is achieved with an input covariance whose eigenvectors coincide at time k with those of $\mathbf{H}_k^\dagger \mathbf{H}_k$, and whose eigenvalues are obtained via waterfilling on those of $\mathbf{H}_k^\dagger \mathbf{H}_k$. This solution was first derived by Tsybakov (1965) for deterministic matrix channels, inspired by Shannon's original frequency-domain waterfilling approach (Shannon, 1949). It was rederived by several authors in the 1990s, specifically for multiantenna communication, once the interest on the topic was sparked (Raleigh and Cioffi, 1998; Telatar, 1999).

Depending on the time horizon over which the power is allowed to be averaged and the temporal dynamics of the problem, several solutions have been obtained for the power assigned to each channel realization:

- Temporal waterfilling (Goldsmith and Varaiya, 1997; Biglieri *et al.*, 2001).
- Dynamic programming (Negi and Cioffi, 2002).

The same techniques can be used if the variability takes place along the frequency (rather than time) axis.

Note that, in practice, discrete constellations such as PSK or QAM are used in lieu of the capacity-achieving Gaussian inputs. Not only does this reduce the mutual information, but power allocation via waterfilling on the channel singular values is no longer optimal (Lozano *et al.*, 2005b).

5.3.2 No transmitter side information

The optimization of the input becomes more problematic, and in fact it has not been fully solved, when the channel realizations are unknown to the transmitter. For the canonical channel (i.i.d. Rayleigh-fading entries), the intuitively appealing isotropy of the optimal input was proved by Telatar (1999). It is, however, frequently forgotten that isotropic inputs are not necessarily optimal for noncanonical channels. A popular nonisotropic signaling, which is a natural counterpart to the solution with transmit side information, is *eigenbeamforming*: signaling on the eigenvectors of $E[\mathbf{H}^\dagger \mathbf{H}]$.

Myth 1 *Regardless of the distribution of* $\mathbf{H}$*, eigenbeamforming achieves capacity.*

The optimality of inputs whose eigenstructure is that of $E[\mathbf{H}^\dagger \mathbf{H}]$ holds true for several important classes of channels such as Rayleigh fading with certain correlation structures (Jafar *et al.*, 2004; Jorswieck and Boche, 2004a) and Ricean fading (Venkatesan *et al.*, 2003; Hoesli and Lapidoth, 2004a). However, this principle fails to hold in general, as the following counterexample shows.

Example 5.1 *Given $n_{\mathsf{T}} = n_{\mathsf{R}} = 3$, let $\mathbf{H}$ take three possible realizations with respective probabilities $q_1 = 0.4$, $q_2 = 0.58$ and $q_3 = 0.02$. Specifically,*

$$\mathbf{H}_k = \boldsymbol{\Lambda}_k^{1/2} \mathbf{V}_k^{\dagger},$$

where $\mathbf{V}_3 = \mathbf{I}$ while $\mathbf{V}_1 = \mathbf{V}_2 = \mathbf{V} \neq \mathbf{I}$ is a unitary matrix and

$$\begin{aligned}
\boldsymbol{\Lambda}_1 &= \mathrm{diag}\{2.5, 2.5, 1.75\}, \\
\boldsymbol{\Lambda}_2 &= \mathrm{diag}\{0, 0, 0.5172\}, \\
\boldsymbol{\Lambda}_3 &= \mathrm{diag}\{15.2, 10.3, 8.1\}.
\end{aligned}$$

The eigenvectors of $E[\mathbf{H}^{\dagger}\mathbf{H}] = \mathrm{diag}\{1.304, 1.206, 1.162\}$ are the columns of the identity, but an input covariance $\boldsymbol{\Phi}_{\mathbf{x}}$ whose eigenvectors coincide with those of $\mathbf{V}$ can achieve strictly higher $E[\mathcal{I}(\mathsf{SNR}, \boldsymbol{\Phi}_{\mathbf{x}})]$.

The optimization of the eigenvalues of $\boldsymbol{\Phi}_{\mathbf{x}}$ has only been explicitly solved at low SNR (Verdú, 2002b) and, for $n_{\mathsf{T}} \leq n_{\mathsf{R}}$, at high SNR (Hoesli and Lapidoth, 2004b; Tulino *et al.*, 2004). Beyond these limiting scenarios, the eigenvalues have only been characterized through fixed-point conditions that must be solved for iteratively (Hoesli and Lapidoth, 2004b; Tulino *et al.*, 2006).

A low-complexity approach consists of determining the eigenvalues of the input covariance matrix by *statistical waterfilling*, namely, a waterfilling on the eigenvalues of $E[\mathbf{H}^{\dagger}\mathbf{H}]$ (e.g., Ivrlac *et al.* (2003), Simeone and Spagnolini (2003)). For $n_{\mathsf{R}} = 1$, this approach minimizes the pairwise error probability at high SNR (Zhou and Giannakis, 2003).

Myth 2 *The eigenvalues of the capacity-achieving input covariance are determined by statistical waterfilling.*

The capacity-achieving eigenvalues are in general not given by a waterfilling on the eigenvalues of any channel covariance matrix.

At low SNR, it is inviting to solve the optimization of $\boldsymbol{\Phi}_{\mathbf{x}}$ by simply maximizing the first term in the corresponding series expansion of $\mathcal{I}(\mathsf{SNR}, \boldsymbol{\Phi}_{\mathbf{x}})$, i.e., by solving the much simpler problem

$$C(\mathsf{SNR}) = \max_{\boldsymbol{\Phi}_{\mathbf{x}}:\mathrm{Tr}\{\boldsymbol{\Phi}_{\mathbf{x}}\}=n_{\mathsf{T}}} \frac{\mathsf{SNR}}{n_{\mathsf{T}}} E\left[\mathrm{Tr}\left\{\mathbf{H}\boldsymbol{\Phi}_{\mathbf{x}}\mathbf{H}^{\dagger}\right\}\right] \log_2 e + o(\mathsf{SNR}). \tag{5.6}$$

The cost function in (5.6), however, only reflects the energy at the output of the channel. There is no notion of bandwidth in the optimization, which leads to the following misconception.

Myth 3 *Rank-1 beamforming in the direction of the largest-eigenvalue eigenvector of $E[\mathbf{H}^\dagger \mathbf{H}]$ is always optimal at low* SNR*. If such eigenvalue is multiple, beamforming in the direction of any associated eigenvector is optimal.*

As the following example shows unequivocally, unless bandwidth is a free commodity, beamforming is not optimal (regardless of how low the SNR is) if the largest eigenvalue of $E[\mathbf{H}^\dagger \mathbf{H}]$ has plural multiplicity. For second-order optimality, in particular, the power should be uniformly divided among the associated eigenvectors. An extreme case is that of the canonical channel, where the multiplicity is n_T and the power has to be isotropically radiated.

Example 5.2 *(Verdú, 2002a) Consider n_T transmit and one receive antennas with i.i.d. channel entries. If B_ℓ is the bandwidth required to sustain rate R with power P and rank-ℓ signaling, then for vanishing P (and R)*

$$\frac{B_1}{B_{n_\mathsf{T}}} = \frac{2\, n_\mathsf{T}}{1 + n_\mathsf{T}}.$$

For $n_\mathsf{T} = 2$, beamforming requires $1/3$ more bandwidth than an isotropic input. For large n_T, it needs twice the bandwidth.

The suboptimality of beamforming is largely upheld even if the dominating eigenvalues of $E[\mathbf{H}^\dagger \mathbf{H}]$ are distinct but only modestly dissimilar: except at very low SNR, beamforming remains suboptimal. The precise SNR below which beamforming is strictly optimal when the eigenvalues of $E[\mathbf{H}^\dagger \mathbf{H}]$ are distinct has been determined for certain Rayleigh-fading channels (Simon and Moustakas, 2003; Jafar *et al.*, 2004; Jorswieck and Boche, 2004a).

The fallacy in Myth 3, revealed in Verdú (2002b), is evidenced by a low-SNR expansion of the capacity as function of

$$\frac{E_b}{N_0} = \frac{P/R}{N_0}, \tag{5.7}$$

which, letting $x|_{3\text{ dB}} = (10 \log_{10} x)/(10 \log_{10} 2)$, yields

$$\mathsf{C}\left(\tfrac{E_b}{N_0}\right) = \max_{\mathbf{\Phi_x} : \mathrm{Tr}\{\mathbf{\Phi_x}\} = n_\mathsf{T}} \left(\frac{E_b}{N_0}\Big|_{3\text{ dB}} - \frac{E_b}{N_0}_{\min}\Big|_{3\text{ dB}} \right) S_0 + \epsilon, \tag{5.8}$$

where $\frac{E_b}{N_0}_{\min}$ is the minimum required energy per bit, while S_0 is the capacity slope therein, in bits/s/Hz/(3 dB), and ϵ is a lower-order term. Both key measures, $\frac{E_b}{N_0}_{\min}$ and S_0, are functions of $\mathbf{\Phi_x}$.

The first-order optimization in (5.6) is tantamount to minimizing $\frac{E_b}{N_0}_{\min}$, with no regard for S_0, and it leads to wrong conclusions regardless of how low the SNR is.

At high SNR, the following is a common misconception:

Myth 4 *Isotropic inputs are optimal at high* SNR.

The rationale that leads to Myth 4 is to approximate $C(\mathsf{SNR})$ at high SNR as

$$C(\mathsf{SNR}) = \max_{\mathbf{\Phi_x}:\mathrm{Tr}\{\mathbf{\Phi_x}\}=n_\mathsf{T}} E\left[\log_2 \det\left(\frac{\mathsf{SNR}}{n_\mathsf{T}}\mathbf{H}^\dagger\mathbf{H}\right) + \log_2 \det(\mathbf{\Phi_x})\right] + o(1),$$

and then conclude that, because of the concavity of $\log_2 \det(\cdot)$, the input covariance $\mathbf{\Phi_x} = \mathbf{I}$ is optimal. Reinforced by Myth 2 and the limiting uniformity of the waterfilling solution at high SNR, the statement in Myth 4 fails to hold if $\mathbf{H}^\dagger\mathbf{H}$ is singular with positive probability, as for example when $n_\mathsf{T} > n_\mathsf{R}$ (Hoesli and Lapidoth, 2004b; Tulino *et al.*, 2004).

Example 5.3 *Consider a Rayleigh-fading channel with $n_\mathsf{T} = 2$ and $n_\mathsf{R} = 1$, and with λ_1 and λ_2 the (nonzero) eigenvalues of the transmit correlation matrix. Denote by p_1 and p_2 the eigenvalues of $\mathbf{\Phi_x}$ that achieve capacity. In the high-SNR limit (Tulino* et al.*, 2004),*

$$p_1 = \lambda_2 - \frac{\lambda_1\lambda_2}{2}\log_e \frac{p_2\lambda_2}{p_1\lambda_1} \qquad \text{and} \qquad p_2 = \lambda_1 + \frac{\lambda_1\lambda_2}{2}\log_e \frac{p_2\lambda_2}{p_1\lambda_1}.$$

If $\lambda_1 \neq \lambda_2$, then $p_1 \neq p_2$.

Scant progress has been made in the optimization of the input covariances that maximize the rate for a given allowed outage probability. Even for the canonical channel, Telatar's conjecture (Telatar, 1999) that the power is to be equally distributed among a subset of the transmit antennas (that decreases as the outage probability increases) remains open.

5.4 Low SNR

In addition to Myth 3, a number of misleading observations can be made on the basis of a first-order expansion of the coherent capacity at low SNR. For the canonical channel,

$$C(\mathsf{SNR}) = \mathsf{SNR}\, n_\mathsf{R} \log_2 e + o(\mathsf{SNR}). \tag{5.9}$$

Fix R, B, and n_{R}. From (5.7), (5.8), (5.11), and (5.12), the increase in transmit power with $n_{\mathsf{T}} = 1$ relative to arbitrary n_{T} is, in 3 dB units,

$$\Delta P|_{3\ \mathrm{dB}} = \frac{R}{2B}\left(1 - \frac{1}{n_{\mathsf{T}}}\right). \tag{5.10}$$

The first-order series expansion in (5.9) and the fact that (5.10) vanishes for $R/B \to 0$ buttresses the following widespread misconception.

Myth 5 *At low* SNR*, the capacity is unaffected by* n_{T}*.*

While the minimum transmit energy per bit is indeed unaffected by n_{T}, the following example characterizes its impact on the low-SNR power-bandwidth trade-off within the simple context of the canonical channel.

Example 5.4 *(Verdú, 2002a) Denoting by* $B_{n_{\mathsf{T}}}$ *the bandwidth required to sustain rate* R *with power* P *and* n_{T} *transmit antennas, for vanishing* P*,*

$$\frac{B_1}{B_{n_{\mathsf{T}}}} = n_{\mathsf{T}} \frac{1 + n_{\mathsf{R}}}{n_{\mathsf{T}} + n_{\mathsf{R}}}.$$

With $n_{\mathsf{T}} = 1$*, we require* $(n_{\mathsf{R}} + 1)/2$ *times the bandwidth needed with* $n_{\mathsf{T}} = n_{\mathsf{R}}$*, and* $(n_{\mathsf{R}} + 1)$ *times the bandwidth needed with large* n_{T}*.*

This result is easily obtained from (5.8) with the minimum energy per bit and slope of the canonical channel:

$$\frac{E_b}{N_0}_{\min} = \frac{\log_e 2}{g\, n_{\mathsf{R}}} \tag{5.11}$$

$$S_0 = 2 \frac{n_{\mathsf{T}} n_{\mathsf{R}}}{n_{\mathsf{T}} + n_{\mathsf{R}}}. \tag{5.12}$$

Myth 5 stems from the fact that, in terms of power at low SNR, the value of having multiple antennas resides exclusively at the receiver[†] but it fails to recognize that transmit and receive antennas are equally valuable in terms of bandwidth efficiency at a given power.

Consider now a channel with correlated entries, given by

$$\mathbf{H} = \mathbf{\Theta}_{\mathrm{R}}^{1/2} \mathbf{H}_w \mathbf{\Theta}_{\mathrm{T}}^{1/2}, \tag{5.13}$$

where $\mathbf{\Theta}_{\mathrm{T}}$ and $\mathbf{\Theta}_{\mathrm{R}}$ are deterministic transmit and receive correlation matrices, while $\mathbf{H}_w$ is a canonical channel matrix. The first-order expansion in (5.6) yields

$$C(\mathsf{SNR}) = \mathsf{SNR}\, n_{\mathsf{R}} \lambda_{\max}(\mathbf{\Theta}_{\mathrm{T}}) \log_2 e + o(\mathsf{SNR}). \tag{5.14}$$

[†]This stems directly from the fact that the captured power grows with n_{R} but not with n_{T}.

Myth 6 *At low* SNR, *receive correlation has no impact on the capacity.*

This misconception stems again from the fact that the notion of bandwidth is lacking in (5.6) and (5.14). The impact of correlation is indeed small (and it vanishes with the SNR) in terms of power increase for constant R and B, but it may be sizeable in terms of bandwidth expansion with constant R and P (Lozano *et al.*, 2003).

Example 5.5 *The bandwidth B_{c} required to sustain a rate R with power P and with receive correlation $\mathbf{\Theta}_{\mathrm{R}}$ relative to the bandwidth B_{u} required with no receive correlation is, for vanishing P*

$$\frac{B_{\mathrm{c}}}{B_{\mathrm{u}}} = \frac{n_{\mathsf{R}} + \mathrm{Tr}\{\mathbf{\Theta}_{\mathrm{R}}^2\}/n_{\mathsf{R}}}{n_{\mathsf{R}} + 1},$$

which approaches 2 if n_{R} is large and the correlations therein strong.

5.5 High SNR

At high SNR, the capacity with coherent reception expands as

$$C(\mathsf{SNR}) = S_\infty \, \mathsf{SNR}|_{3\ \mathrm{dB}} + O(1), \tag{5.15}$$

where S_∞ in bits/s/Hz/(3 dB) is known variously as the (maximum) multiplexing gain, the pre-log, the high-SNR slope, or the number of degrees of freedom. For the canonical channel,

$$S_\infty = \min(n_{\mathsf{T}}, n_{\mathsf{R}}),$$

an observation that ignited the enthusiasm in multiantenna communication (Foschini, 1996; Telatar, 1999).

Myth 7 *At high* SNR, *the multiantenna capacity is determined by S_∞.*

The reality is that almost all coherent multiantenna channels† have $S_\infty = \min(n_{\mathsf{T}}, n_{\mathsf{R}})$ regardless how they deviate from the canonical model. Yet, the power required to achieve a given capacity is highly sensitive to antenna correlation, line-of-sight components, noise color, the fading distribution, etc. Pragmatically, at any given SNR a certain number of nonzero singular values of the channel matrix are "dormant" (i.e., well below the noise), and

†Possible exceptions are channels with singular correlation and the keyhole channel (whose matrix is the outer product of two vectors (Chizhik *et al.*, 2000)), where $S_\infty \leq \min(n_{\mathsf{T}}, n_{\mathsf{R}})$.

thus at such SNR it is not the rank of the channel matrix that matters but the number of nonnegligible singular values (Tse and Viswanath, 2005). This more pragmatic interpretation of S_∞, nonetheless, still fails to capture the impact on required power at high SNR of various channel features. An expansion, which unlike (5.15), is able to capture such impact is (Shamai and Verdú, 2001; Tulino *et al.*, 2004)

$$C(\mathsf{SNR}) = S_\infty(\mathsf{SNR}|_{3\text{ dB}} - \mathcal{L}_\infty) + o(1), \tag{5.16}$$

where $\mathcal{L}_\infty$ represents the power offset, in 3 dB units, with respect to a reference channel having the same S_∞ but with unfaded and orthogonal dimensions (i.e., such that $(1/n_{\mathsf{T}})\mathbf{H}\mathbf{H}^\dagger = \mathbf{I}$). The power offset plays a chief role at SNR values of operational interest in current high spectral efficiency applications.

Myth 8 *At high* SNR*, antenna correlation is immaterial.*

Nonsingular antenna correlations have no effect on S_∞ but they do shift the power offset.

Example 5.6 *Consider a correlated channel represented by* (5.13) *with nonsingular* $\mathbf{\Theta}_{\mathrm{T}}$ *and* $\mathbf{\Theta}_{\mathrm{R}}$*, and with* $n_{\mathsf{T}} = n_{\mathsf{R}}$*. Denoting the power offset in the absence of correlation by* $\mathcal{L}_\infty^{\mathsf{i.i.d.}}$*, the power penalty in 3-dB units incurred with correlation is*

$$\mathcal{L}_\infty - \mathcal{L}_\infty^{\mathsf{i.i.d.}} = -\frac{1}{n_{\mathsf{T}}}\sum_{\ell=1}^{n_{\mathsf{T}}}\log_2 \lambda_\ell(\mathbf{\Theta}_{\mathrm{T}}) - \frac{1}{n_{\mathsf{R}}}\sum_{\ell=1}^{n_{\mathsf{R}}}\log_2 \lambda_\ell(\mathbf{\Theta}_{\mathrm{R}}),$$

where $\lambda_\ell(\cdot)$ *indicates the* ℓ*th eigenvalue of a matrix. If some of the eigenvalues of* $\mathbf{\Theta}_{\mathrm{T}}$ *or* $\mathbf{\Theta}_{\mathrm{R}}$ *are small, this power penalty can be arbitrarily large.*

In nonergodic channels, it is sometimes desirable to sacrifice some rate in exchange for spatial diversity, so as to ensure a faster decay of the error probability. (In the presence of other diversity mechanisms or when enough delay is tolerable, the need for spatial diversity is far less acute and approaching capacity is the primary goal.) In the high SNR region, the trade-off between degrees of freedom (or multiplexing gain) and diversity was established by Zheng and Tse (2003) for the canonical channel. Note that, at the point of zero diversity, the multiplexing gain attains its maximum value S_∞. For the reasons discussed above, the diversity-multiplexing trade-off does not completely capture the high SNR behavior for noncanonical channels as it neglects the power offset.

5.6 Antenna correlation

In addition to the low- and high-SNR regions, explicit expressions for the coherent capacity of Rayleigh-fading channels whose correlation conforms to (5.13) have appeared in the literature (Kiessling and Speidel, 2004). The capacity for large numbers of antennas has also been reported as the solution of a fixed-point equation, for the correlation structure in (5.13) by Moustakas *et al.* (2000), Mestre *et al.* (2003), and for more general correlations by Tulino *et al.* (2003). All these results, however, require a separate optimization of the eigenvalues of the input covariance.

For $n_{\mathsf{R}} = 1$, with the channel realization known at the transmitter, it has been shown that antenna correlation lowers capacity at any SNR (Jorswieck and Boche, 2004b).[†] However, this is not true in general.

Myth 9 *Antenna correlation is detrimental at any* SNR.

The mutual information achievable with isotropic inputs is indeed lowered by antenna correlation (Jorswieck and Boche, 2004b). When the capacity is achieved by nonisotropic inputs, in contrast, correlation need not be detrimental. The effect is easily assessed when the correlations obey (5.13):

- Transmit correlation reduces the effective dimensionality of the transmitter, but it also enables power focusing. The net effect depends on the SNR and the ratio between n_{T} and n_{R}. Below a certain SNR, correlation is always advantageous; above, it is unfavorable if $n_{\mathsf{T}} \le n_{\mathsf{R}}$, but it may be again advantageous if $n_{\mathsf{T}} > n_{\mathsf{R}}$.[‡]
- Receive correlation reduces the effective dimensionality of the receiver without increasing the captured power and thus it lowers the mutual information at any SNR.

5.7 Ricean fading

The computation of the capacity of Ricean channels is substantially more involved than that of Rayleigh-fading channels, and fewer results are available (Kang and Alouini, 2002; Lebrun *et al.*, 2004; Alfano *et al.*, 2004).

Consider the Ricean channel

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\bar{\mathbf{H}} + \sqrt{\frac{1}{K+1}}\mathbf{H}_w, \tag{5.17}$$

[†] This also holds if the transmitter not only does not know the channel realization but only knows that the channel distribution satisfies certain symmetry constraints.

[‡] Transmit correlation is also beneficial at every SNR in the case of noncoherent reception (Jafar and Goldsmith, 2003).

where $\bar{\mathbf{H}}$ is deterministic while $\mathbf{H}_w$ is a canonical channel matrix and K is the Ricean K-factor. If $n_{\mathsf{T}} = n_{\mathsf{R}} = 1$, the capacity increases monotonically with K. In the multiantenna arena, however, line-of-sight components are sometimes perceived as being detrimental to the capacity.

Myth 10 *The presence of a line-of-sight component reduces the capacity of a multiantenna channel.*

If $\mathbf{H}$ is normalized as per (5.2), then, depending on the singular values of $\bar{\mathbf{H}}$, we can find instances where the capacity is either improved or degraded by the line-of-sight component. The following examples illustrate these different behaviors:

Example 5.7 *Let $\bar{\mathbf{H}}$ be unit rank, and let $n_{\mathsf{T}} = n_{\mathsf{R}} = 2$. At high* SNR*, the capacity behaves as* (5.16) *with $S_\infty = 2$ and with Lozano* et al. *(2005a)*

$$\mathcal{L}_\infty(K) = 1 + \log_2 \frac{K+1}{2\sqrt{K}} - \frac{\log_2 e}{2}\left(\mathrm{E}_1(4K) - \gamma + \frac{1-e^{-4K}}{4K}\right), \qquad (5.18)$$

where γ is Euler's constant, and $\mathrm{E}_1(z) = \int_1^\infty e^{-z\xi}/\xi \, d\xi$ is an exponential integral. The power offset in (5.18) *increases monotonically with K and for large K it grows as $(1/2)\log_2 K$. With $K = 100$, for instance, the excess power offset due to the line-of-sight component amounts to 7.9 dB.*

Example 5.8 *For the full-rank line-of-sight matrix*

$$\bar{\mathbf{H}} = \begin{bmatrix} 1 & 0.4 \\ 0.9 & 0.1 \\ 0.2 & -2 \end{bmatrix}$$

it can be verified, using the expressions in Lozano et al. *(2005a), that*

$$\mathcal{L}_\infty(5) = \mathcal{L}_\infty(0) - 0.33,$$

or about a 1 dB gain at high SNR *thanks to a $K = 5$ line-of-sight component.*

Notice that, by using (5.17), we have evaluated the impact of the line-of-sight component at a fixed SNR. This requires that, in the presence of a line-of-sight component, the power received over the fading portion of the channel declines. Alternatively, we can consider the modified Ricean channel

$$\mathbf{H} = \sqrt{K}\bar{\mathbf{H}} + \mathbf{H}_w, \qquad (5.19)$$

which, in general, does not satisfy (5.2). The addition of a line-of-sight component does not alter the faded power, and the corresponding capacity increases monotonically with K (Hoesli and Lapidoth, 2004a). Therefore, the behavior is seen to depend critically on the channel model.

As both n_{T} and n_{R} go to infinity with constant ratio, the change in capacity caused by a rank-1 line-of-sight component takes a remarkably simple form. Making explicit the dependence of the mutual information on K, for the model in (5.17)

$$\lim_{n_{\mathsf{T}},n_{\mathsf{R}}\to\infty} \frac{\mathcal{I}(\mathsf{SNR},K)}{n_{\mathsf{R}}} = \lim_{n_{\mathsf{T}},n_{\mathsf{R}}\to\infty} \frac{\mathcal{I}(\frac{\mathsf{SNR}}{K+1},0)}{n_{\mathsf{R}}}, \tag{5.20}$$

whereas, for the model in (5.19),

$$\lim_{n_{\mathsf{T}},n_{\mathsf{R}}\to\infty} \frac{\mathcal{I}(\mathsf{SNR},K)}{n_{\mathsf{R}}} = \lim_{n_{\mathsf{T}},n_{\mathsf{R}}\to\infty} \frac{\mathcal{I}(\mathsf{SNR},0)}{n_{\mathsf{R}}}. \tag{5.21}$$

In either case, the mutual information behaves as if the line-of-sight component were absent. This is a direct manifestation of the fact that only a single eigenvalue of $\mathbf{H}\mathbf{\Phi}_{\mathbf{x}}\mathbf{H}^{\dagger}$ is perturbed by the line-of-sight component. Asymptotically, this perturbation is not reflected in the empirical eigenvalue distribution of $\mathbf{H}\mathbf{\Phi}_{\mathbf{x}}\mathbf{H}^{\dagger}$, which determines the mutual information.[†]

5.8 Asymptotic analysis

Asymptotic analyses as n_{T} and n_{R} go to infinity with a constant ratio are often feasible using results in random matrix theory (Tulino and Verdú, 2004) and provide valuable engineering insights. Furthermore, the number of antennas required for the asymptotics to be closely approached are often very small. However, this principle should not be taken for granted without numerical verification. For example, in the scenario considered in the last paragraph of Section 5.7 the limits therein are approached very slowly.

Myth 11 *Asymptotic capacity results are always closely approached for small numbers of antennas.*

If either n_{T} or n_{R} is held fixed while the other dimension grows without bound, application of the strong law of large numbers to a channel with zero-mean i.i.d. entries yields

$$\begin{aligned} \frac{1}{n_{\mathsf{R}}}\mathbf{H}^{\dagger}\mathbf{H} &\overset{a.s.}{\to} \mathbf{I} \qquad (n_{\mathsf{R}}\to\infty) \\ \frac{1}{n_{\mathsf{T}}}\mathbf{H}\mathbf{H}^{\dagger} &\overset{a.s.}{\to} \mathbf{I} \qquad (n_{\mathsf{T}}\to\infty). \end{aligned}$$

[†]The behavior in (5.20) and (5.21) extends to channels whose line-of-sight component has rank $r > 1$ as long as $\lim_{n_{\mathsf{R}}\to\infty} r/n_{\mathsf{R}} = 0$.

In these asymptotes, the capacity of the canonical channel (and of channels that can be expressed as function thereof) becomes particularly simple. If both n_{T} and n_{R} go to infinity with constant ratio, the (i, j) entries of both $(1/n_{\mathsf{R}})\mathbf{H}^{\dagger}\mathbf{H}$ and $(1/n_{\mathsf{T}})\mathbf{H}\mathbf{H}^{\dagger}$ also converge almost surely (a.s.) to δ_{i-j} for any fixed pair (i, j). The empirical eigenvalue distribution, however, does not converge to a mass at 1, but rather to the Marčenko–Pastur law (Marčenko and Pastur, 1967). Integration over this asymptotic distribution yields the corresponding asymptotic capacity (normalized by the number of antennas) in closed form (Verdú and Shamai, 1999; Rapajic and Popescu, 2000).

For most channels, the mutual information normalized by the number of antennas converges a.s. to its expectation asymptotically.[†] In contrast, the unnormalized mutual information, which determines the outage capacity, still suffers from nonvanishing random fluctuations. Numerous authors have observed (through simulation) that the distribution of the unnormalized mutual information resembles a Gaussian law once n_{T} and n_{R} become large. Asymptotic normality has been rigorously established (Tulino and Verdú, 2005) for arbitrary SNR in the presence of correlation at either transmitter or receiver as a direct application of recent random matrix results (Bai and Silverstein, 2004). For other multiantenna channels, asymptotic normality has been conjectured (Moustakas *et al.*, 2000; Kang and Alouini, 2003).

5.9 Intercell and multiuser issues

A popular way to deal with multiuser interference in general, and out-of-cell interference in particular, is to model it as white Gaussian noise. It is well known that, for single-antenna systems, such a model leads to severe capacity underestimations. This conclusion holds when the interference emanates from multiantenna arrays and the receiver is also equipped with an array.

Myth 12 *Out-of-cell interference can be modeled as additional white Gaussian noise.*

It is possible to effectively exploit the structure of the out-of-cell interference even if the codebooks used in other cells are unknown. To that end, consider the scenario where the MIMO channel from the strongest out-of-cell base station array to an in-cell mobile array is known (e.g., through pilot monitoring). The following example illustrates the kind of improvements achievable in the low-SNR regime by taking this knowledge into account.

[†]Exceptions are channels with finite-rank correlations and the keyhole channel.

Example 5.9 *Let E_b and E_b^{awgn} be the minimum energy per bit achievable in the presence of one out-of-cell interferer (and no thermal noise), and in the presence of white Gaussian noise of identical power, respectively. Both the desired transmitter and the interferer have n_{T} antennas. From Lozano* et al. *(2003), if $n_{\mathsf{T}} > n_{\mathsf{R}}$, then*

$$\frac{E_b}{E_b^{\mathsf{awgn}}} = 1 - \frac{n_{\mathsf{R}}}{n_{\mathsf{T}}}. \tag{5.22}$$

Thus, for $n_{\mathsf{T}} = 2\,n_{\mathsf{R}}$, $\frac{E_b}{N_0}_{\min}$ is 3 dB lower than if the interference was white Gaussian noise with the same power. For $n_{\mathsf{T}} \leq n_{\mathsf{R}}$, the $\frac{E_b}{N_0}_{\min}$ is the same as if the interference did not exist.

In the high-SNR regime, with a large ratio between in-cell and out-of-cell powers, the following evidences how the power offset can be considerably reduced by exploiting the structure of the interference.

Example 5.10 *Let $n_{\mathsf{T}} = n_{\mathsf{R}} = n$, and denote by $\mathcal{L}_\infty^{\mathsf{awgn}}$ the high-SNR power offset with white Gaussian noise. In the presence of one interferer equipped with n_{T} antennas (and no thermal noise),*

$$\mathcal{L}_\infty = \mathcal{L}_\infty^{\mathsf{awgn}} + \log_2 n + \left(\gamma - \sum_{\ell=2}^{n} \frac{1}{\ell}\right) \log_2 e$$

in 3 dB units. For $n = 4$, for instance, the difference in power required to achieve a certain capacity approaches 3.8 dB as the SNR grows.

Another important concept in multiuser problems is the principle of multiuser diversity. One of its earliest embodiments (for a flat-fading single-antenna setting) is that, to maximize the uplink sum capacity with fading known to the scheduler, only the strongest user should be allowed to transmit (Knopp and Humblet, 1995). Likewise for the downlink, the sum capacity is maximized by transmitting only to the strongest user (Tse, 1999), a result that was a central theme in the design of CDMA2000® 1xEV-DO and UMTS-HSDPA third generation data-only wireless systems. Time sharing on the strongest user ceases to be optimal in a multiantenna setting. A less coarse misconception, but perhaps more generalized, is

---

Myth 13 *In either uplink or downlink, it is optimal to schedule as many simultaneous transmissions as the number of antennas at the base station.*

This belief may originate partially from Myth 7 and from the fact that, with a linear decorrelating receiver, each additional antenna enables to maintain the same performance, but with one additional interferer (e.g., Verdú, 1998, Section 5.7.2).

For both uplink and downlink with an n_{T}-antenna base station, Yu and Rhee (2004) show that it is strictly suboptimal to transmit from or to only n_{T} mobiles if $n_{\mathsf{T}} > 1$. A (generally loose) bound in Yu and Rhee (2004) indicates that maximizing the sum rate may entail scheduling up to n_{T}^2 users at once.

Despite the fact that scheduling up to n_{T} users is strictly suboptimal, it generally incurs a small loss in sum capacity. For example, consider an uplink with four statistically equivalent Rayleigh-fading mobiles and a base station with two independently faded antennas. To achieve a sum capacity of 7.5 bits/s/Hz, the optimum policy schedules more than two simultaneous users in 55% of the realizations. Using the suboptimal policy that allows only the two strongest users to transmit, achieving that same sum capacity requires an additional 0.25 dB power expenditure per user, in addition to incurring higher latency and reduced fairness.

5.10 Transmit-receive architectures

One of the earliest multiantenna transmit-receive techniques is the widely popular Alamouti scheme designed for $n_{\mathsf{T}} = 2$ (Alamouti, 1998).

Myth 14 *In conjunction with a scalar code, the Alamouti scheme achieves capacity if* $n_{\mathsf{T}} = 2$ *and* $n_{\mathsf{R}} = 1$.

A signal constructed according to the Alamouti scheme is spatially isotropic. Accordingly, the statement in Myth 14 fails to hold for those channels where the identity input covariance is suboptimal. For those channels, a linear interface should be inserted to properly color the signal.

Another technique that helped propel the interest in multiantenna communication is the BLAST layered architecture, which, in both its diagonal and vertical versions (D-BLAST and V-BLAST), was designed to operate with scalar codes and without transmit side information (Foschini, 1996; Foschini *et al.*, 1999).

Myth 15 *The vertical layered architecture* (V-BLAST) *cannot achieve the capacity of the canonical channel.*

If the rates at which the various transmit antennas operate are not constrained to be identical, and these rates can be communicated to the transmitter, then those rates can be chosen such that V-BLAST with MMSE filtering achieves capacity (Ariyavisitakul, 2000; Chung *et al.*, 2004); see also Varanasi and Guess (1998). By feeding the independently encoded signals onto appropriate signaling eigenvectors, the optimality of V-BLAST can be extended to noncanonical channels. In contrast, a vertically layered architecture cannot attain the diversity-multiplexing trade-off of Zheng and Tse (2003). D-BLAST, on the other hand, can achieve both capacity and the diversity-multiplexing trade-off.

5.11 Noncoherent communication

When the sequence of channel matrices $\mathbf{H}_k$ is not known exactly at the receiver, the study of the capacity is much more challenging, and neither a counterpart to the log-determinant formula nor the capacity-achieving signaling are yet known, even for $n_{\mathsf{T}} = n_{\mathsf{R}} = 1$.

At high SNR, the analysis of the capacity of both scalar and multiantenna noncoherent channels has received considerable attention, under various models of the fading dynamics. For scalar channels with memoryless fading, Taricco and Elia (1997) showed that the capacity grows as

$$C(\mathsf{SNR}) = \log_2 \log \mathsf{SNR} + O(1), \tag{5.23}$$

a behavior that has been shown to hold for a broad class of scalar and vector channels where the fading coefficients cannot be perfectly predicted from their past (Lapidoth and Moser, 2003). According to (5.23), transmit power is essentially wasted in the noncoherent high-SNR regime, suggesting that the system perhaps ought to be redesigned to operate in a wider-bandwidth, lower-SNR regime. In contrast, when the fading coefficients are perfectly predictable from their past, the capacity grows—as in the coherent regime—logarithmically with SNR (Lapidoth, 2003, 2005). One such noncoherent model, which has attracted much attention in the literature, is the nonstationary block fading model, where the channel coefficients remain constant for blocks of T symbols (Marzetta and Hochwald, 1999). In that case, Zheng and Tse (2002) showed

$$S_\infty = m\left(1 - \frac{m}{T}\right),$$

with $m = \min(n_\mathsf{R}, n_\mathsf{T}, \lfloor T/2 \rfloor)$. Thus, although slight modeling variations usually lead to small disparities in capacity at any given SNR, for $\mathsf{SNR} \to \infty$ the qualitative behavior of the capacity is quite sensitive to the model.

One could espouse the view that every wireless channel is noncoherent and that any measure (such as pilot symbols) taken to provide the receiver with an estimate of the channel should be modeled as part of the input. This view is particularly acute at low SNR, where the channel coefficients are harder to estimate (Rao and Hassibi, 2004).

Myth 16 *At low* SNR*, coherent communication is not feasible.*

Although, as $\mathsf{SNR} \to 0$, the channel can indeed be learned with diminishing precision, coherent operation also requires diminishing accuracy (Lapidoth and Shamai, 2002). The classical analysis of capacity in the low-power regime predicts that, regardless of whether coherence is feasible or not, it is useless, since it does not decrease the minimum energy per bit. In contrast, the view propounded by Verdú (2002b) is that in the low-power regime coherence plays a key role in reducing the bandwidth required for reliable communication. In terms of the power-bandwidth trade-off, a quasi-coherent channel (known with little error at the receiver) can indeed be judiciously analyzed as a coherent channel using the tools developed by Verdú (2002b).

Acknowledgment

The authors gratefully acknowledge valuable comments from Giuseppe Caire, Jerry Foschini, Andrea Goldsmith, Babak Hassibi, Shlomo Shamai, David Tse, and Wei Yu.

References

Alamouti, S. M. (1998). A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Commun.*, **16** (8), 1451–1458.

Alfano, G., Lozano, A., Tulino, A. M., and Verdú, S. (2004). Mutual information and eigenvalue distribution of MIMO Ricean channels. In *Proc. Int. Symp. Inf. Theory and its Applications*, 1040–1045.

Ariyavisitakul, S. L. (2000). Turbo space-time processing to improve wireless channel capacity. *IEEE Trans. Commun.*, **48** (8), 1347–1358.

Bai, Z. D. and Silverstein, J. W. (2004). CLT of linear spectral statistics of large dimensional sample covariance matrices. *Ann. Probab*, **32** (1), 553–605.

Biglieri, E., Caire, G., and Taricco, G. (2001). Limiting performance of block-fading channels with multiple antennas. *IEEE Trans. Inf. Theory*, **47** (4), 1273–1289.

Brandenburg, L. H. and Wyner, A. D. (1974). Capacity of the Gaussian channel with memory: The multivariate case. *Bell Syst. Tech. J.*, **53** (5), 745–778.

Bronk, B. V. (1965). Exponential ensembles for random matrices. *J. Math. Physics*, **6** (2), 228–237.

Chizhik, D., Foschini, G. J., and Valenzuela, R. A. (2000). Capacities of multi-element transmit and receive antennas: Correlations and keyholes. *IEE Electr. Lett.*, **36** (13), 1099–1100.

Chung, S. T., Lozano, A., Huang, H. C., Sutivong, A., and Cioffi, J. M. (2004). Approaching the MIMO capacity with a low-rate feedback channel in V-BLAST. *EURASIP J. Appl. Signal Process.*, **2004** (5), 762–771.

Cover, T. M. and Thomas, J. A. (1990). *Elements of Information Theory* (New York, Wiley).

Foschini, G. J. (1996). Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, **1** (2), 41–59.

Foschini, G. J., Golden, G. D., Valenzuela, R. A., and Wolnianski, P. W. (1999). Simplified processing for high spectral efficiency wireless communication employing multi-element arrays. *IEEE J. Sel. Areas Commun.*, **17** (11), 1841–1852.

Goldsmith, A. and Varaiya, P. (1997). Capacity of fading channels with channel side information. *IEEE Trans. Inf. Theory*, **43** (6), 1986–1997.

Hoesli, D. and Lapidoth, A. (2004a). The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean. In 5*th Int. ITG Conf. Source and Channel Coding, Erlangen, Germany.*

—— (2004b). How good is an isotropic input on a MIMO Ricean channel? In *Proc. IEEE Int. Symp. Inf. Theory*, 291.

Ivrlac, M. T., Utschick, W., and Nossek, J. A. (2003). Fading correlations in wireless MIMO communication systems. *IEEE J. Sel. Areas Commun.*, **21** (5), 819–828.

Jafar, S. A. and Goldsmith, A. J. (2003). Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information. In *Proc. IEEE Int. Symp. Inf. Theory*, 470.

Jafar, S. A., Vishwanath, S., and Goldsmith, A. J. (2004). Transmitter optimization and optimality of beamforming for multiple antenna systems. *IEEE Trans. Wireless Commun.*, **3** (4), 1165–1175.

Jorswieck, E. and Boche, H. (2004a). Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback. *IEEE Trans. Wireless Commun.*, **3** (5), 1543–1553.

—— (2004b). Optimal transmission strategies and impact of correlation in multiantenna systems with differeny types of channel state information. *IEEE Trans. Signal Process.*, **52** (12), 3440–3453.

Kang, M. and Alouini, M. S. (2002). On the capacity of MIMO Rician channels. In *Proc. 40th Allerton Conf. Commun., Contr., Comput.*, 936–945.

—— (2003). Impact of correlation on the capacity of MIMO channels. In *Proc. IEEE Int. Conf. Commun.*, vol. 4, 2623–2627.

Kiessling, M. and Speidel, J. (2004). Exact ergodic capacity of MIMO channels in correlated Rayleigh fading environments. In *Proc. Int. Zurich Seminar Commun.*, 128–131.

Knopp, R. and Humblet, P. A. (1995). Information capacity and power control in single-cell multi-user communications. In *Proc. IEEE Int. Conf. Commun.*, 331–335.

Lapidoth, A. (2003). On the high SNR capacity of stationary Gaussian fading channels. In *Proc. Allerton Conf. Commun., Contr., Comput.*, 410–419.

—— (2005). On the asymptotic capacity of stationary Gaussian fading channels. *IEEE Trans. Inf. Theory*, **51** (2), 437–446.

Lapidoth, A. and Moser, S. M. (2003). Capacity bounds via duality with applications to multi-antenna systems on flat-fading channels. *IEEE Trans. Inf. Theory*, **49** (10), 2426–2467.

Lapidoth, A. and Shamai, S. (2002). Fading channels: How perfect need "perfect side information" be? *IEEE Trans. Inf. Theory*, **48** (5), 1118–1134.

Lebrun, G., Faulkner, M., Shafi, M., and Smith, P. J. (2004). MIMO Ricean channel capacity. In *Proc. IEEE Int. Conf. Commun.*, 2939–2943.

Lozano, A., Tulino, A. M., and Verdú, S. (2003). Multiple-antenna capacity in the low-power regime. *IEEE Trans. Inf. Theory*, **49** (10), 2527–2544.

—— (2005a). High-SNR power offset in multi-antenna in multi-antenna Ricean channels. In *Proc. IEEE Int. Conf. Commun.*

—— (2005b). Mercury/waterfilling: Optimum power allocation with arbitrary input constellations. In *Proc. IEEE Int. Symp. Inf. Theory.*

Marčenko, V. A. and Pastur, L. A. (1967). Distributions of eigenvalues for some sets of random matrices. *Math USSR-Sbornik*, **1**, 457–483.

Marzetta, T. L. and Hochwald, B. H. (1999). Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading. *IEEE Trans. Inf. Theory*, **45** (1), 139–157.

Mestre, X., Fonollosa, J. R., and Pagès-Zamora, A. (2003). Capacity of MIMO channels: asymptotic evaluation under correlated fading. *IEEE J. Sel. Areas Commun.*, **21** (5), 829–838.

Moustakas, A., Baranger, H., Balents, L., Sengupta, A., and Simon, S. (2000). Communication through a diffusive medium: Coherence and capacity. *Science*, **287** (5451), 287–290.

Negi, R. and Cioffi, J. M. (2002). Delay-constrained capacity with causal feedback. *IEEE Trans. Inf. Theory*, **48** (9), 2478–2494.

Pinsker, M. S. (1964). *Information and Information Stability of Random Variables and Processes* (Holden-Day, San Francisco, CA).

Raleigh, G. and Cioffi, J. M. (1998). Spatio-temporal coding for wireless communications. *IEEE Trans. Commun.*, **46** (3), 353–366.

Rao, C. and Hassibi, B. (2004). Analysis of multiple-antenna wireless links at low SNR. *IEEE Trans. Inf. Theory*, **50** (9), 2123–2130.

Rapajic, P. and Popescu, D. (2000). Information capacity of a random signature multiple-input multiple-output channel. *IEEE Trans. Commun.*, **48** (8), 1245–1248.

Root, W. L. and Varaiya, P. P. (1968). Capacity of classes of Gaussian channels. *SIAM J. Appl. Math.*, **16** (6), 1350–1393.

Shamai, S. and Verdú, S. (2001). The impact of frequency-flat fading on the spectral efficiency of CDMA. *IEEE Trans. Inf. Theory*, **47** (4), 1302–1327.

Shannon, C. E. (1949). Communication in the presence of noise. *Proc. IRE*, **37** (1), 10–21.

Shin, H. and Lee, J. H. (2003). Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering and keyhole. *IEEE Trans. Inf. Theory*, **49** (10), 2636–2647.

Simeone, O. and Spagnolini, U. (2003). Combined linear pre-equalization and BLAST equalization with channel correlation feedback. *IEEE Commun. Lett.*,

7 (10), 487–489.

Simon, S. H. and Moustakas, A. L. (2003). Optimizing MIMO antenna systems with channel covariance feedback. *IEEE J. Sel. Areas Commun.*, **21** (3), 406–417.

Taricco, G. and Elia, M. (1997). Capacity of fading channels with no side information. *IEE Electr. Lett.*, **33** (16), 1368–1370.

Telatar, I. E. (1999). Capacity of multi-antenna Gaussian channels. *Eur. Trans. Telecommun.*, **10**, 585–595.

Tse, D. (1999). Forward-link multiuser diversity through rate adaptation and scheduling.

Tse, D. and Viswanath, P. (2005). *Fundamentals of Wireless communications* (Cambridge Univ. Press).

Tsybakov, B. S. (1965). The capacity of a memoryless Gaussian vector channel. *Probl. Inf. Transmission*, **1**, 18–29.

Tulino, A., Lozano, A., and Verdú, S. (2004). High-SNR power offset in multi-antenna communication. In *Proc. IEEE Int. Symp. Inf. Theory*, 288.

—— (2006). Capacity-achieving input covariance for single-user multi-antenna channels. *IEEE Trans. Wireless Commun.* In press.

Tulino, A. and Verdú, S. (2004). Random matrix theory and wireless communications. *Foundations & Trends in Commun. and Inf. Theory*, **1** (1), 1–182.

—— (2005). Asymptotic outage capacity of multiantenna channels. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 825–828.

Tulino, A., Verdú, S., and Lozano, A. (2003). Capacity of antenna arrays with space, polarization and pattern diversity. In *Proc. IEEE Inf. Theory Workshop.*

Varanasi, M. K. and Guess, T. (1998). Optimum decision-feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel. In *Proc. Asilomar Conf. Signals, Syst., Comput.*, 1405–1409.

Venkatesan, S., Simon, S. H., and Valenzuela, R. A. (2003). Capacity of a Gaussian MIMO channel with nonzero mean. In *Proc. IEEE Vehic. Tech. Conf.*, vol. 3, 1767–1771.

Verdú, S. (1986). Capacity region of Gaussian CDMA channels: The symbol synchronous case. In *Proc. Allerton Conf. Commun., Contr., Comput.*, 1025–1034.

—— (1998). *Multiuser Detection* (Cambridge Univ. Press).

—— (2002a). Saving bandwidth in the wideband regime. Plenary talk, Intl. Symp. Inf. Theory.

—— (2002b). Spectral efficiency in the wideband regime. *IEEE Trans. Inf. Theory*, **48** (6), 1319–1343.

Verdú, S. and Shamai, S. (1999). Spectral efficiency of CDMA with random spreading. *IEEE Trans. Inf. Theory*, **45** (3), 622–640.

Yu, W. and Rhee, W. (2004). Degrees of freedom in multi-user spatial multiplex systems with multiple antennas. *IEEE Trans. Commun.* Submitted.

Zheng, L. and Tse, D. (2003). Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEEE Trans. Inf. Theory*, **49** (5), 1073–1096.

Zheng, L. and Tse, D. N. C. (2002). Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel. *IEEE Trans. Inf. Theory*, **48** (2), 359–383.

Zhou, S. and Giannakis, G. B. (2003). Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations. *IEEE Trans. Inf. Theory*, **49** (7), 1673–1690.