IMPACT OF MISMATCHED STATISTICS ON CORRELATED MIMO CAPACITY

Vasanthan Raghavan, Akbar M. Sayeed, and Jayesh H. Kotecha

ABSTRACT

Many recent works that study the impact of spatial correlation on the performance of multi-input multi-output (MIMO) systems assume a separable (also known as the Kronecker) model where the variances of channel entries, upon decomposition onto the transmit and the receive eigen-bases, admit a separable form. If the true statistics of the channel coefficients are non-separable, the separability assumption leads to a flattening and spreading of the degrees of freedom (DoF) in the channel, and hence results in misleading estimates of capacity – an observation consistent with many measurement campaigns. Towards understanding this observation, we first elucidate the importance of channel power normalization, an often-ignored notion, in the capacity analysis of correlated channels. Using tools from random matrix theory, we characterize the mismatch in estimating capacity with the separable model, and thus provide a theoretical underpinning behind many measurement-based observations.

***Index Terms*—** Capacity, correlation, fading channels, MIMO systems, model mismatch, multiplexing, random matrix theory.

1. INTRODUCTION

Initial results on multi-input multi-output (MIMO) systems show that a linear growth in multiplexing gain and coherent capacity is possible with the number of antennas under the assumption of spatially independent and identically distributed (i.i.d.) Rayleigh fading between antenna pairs [1, 2]. However, the rich scattering assumption is idealistic and most physical channels encountered in practice exhibit clustered scattering and spatially correlated links. Correlated MIMO channels have been theoretically studied mainly in the contexts of the separable correlation model (also known as the Kronecker model) [3] and the virtual representation framework for uniform linear arrays (ULAs) [4]. The Kronecker model assumes separability in correlation induced by the transmitter and the receiver arrays which limits the degrees of freedom (DoF) in modeling the channel. Though this model is accurate in certain scattering environments, the separability assumption limits its applicability to more realistic settings. The virtual representation does not assume such separability, but is applicable only for ULAs.

A general *canonical* modeling framework which accommodates non-separable variances for the channel coefficients has been introduced in recent works [5, 6, 7]. Here, the channel is decomposed into a canonical channel matrix via two *statistics-dependent* unitary matrices that correspond to the eigen-modes of the transmit and the receive covariance matrices. The entries of the canonical channel matrix embody the statistically independent DoF that govern channel capacity and diversity. Experimental evidence [6] shows that in most realistic channels, the Kronecker model consistently estimates true capacity poorly while the canonical model predicts capacity more accurately. The main focus of this paper is on understanding the theoretical basis behind this observation.

Towards this goal, we first illustrate the importance of channel power normalization (where channel power is defined as $\rho_c = E[\mathrm{Tr}(\mathbf{H}\mathbf{H}^H)]$ with $\mathbf{H}$ denoting the channel matrix) in capacity analysis. The most commonly used normalization requires ρ_c to remain equal, irrespective of the environment or antenna dimensions. We argue that this normalization is inadequate and unfair for all performance comparisons. Recently reported observations on channel capacity like "A correlated channel performs better than an i.i.d. channel at low-SNR," critically depend on how the channel power is normalized. We present a new channel power normalization that constrains the variances of channel entries to be bounded. With the proposed normalization, we observe that a richer channel results in a higher capacity over a more correlated channel at any SNR.

We then assume that the true statistics of the channel coefficients are non-separable and study the capacity mismatch with a Kronecker model fitted to this channel under the proposed and the standard channel power normalizations. This is done by computing the means and variances of the capacity random variable for the two model fits using tools from random matrix theory. With the proposed normalization, our results show that for non-regular[1] channels at high- to medium-SNRs, the Kronecker model underestimates the outage capacity at all reliability levels (and also the reliability at all data rates). On the other hand, for any channel in the low-SNR regime, and regular channels in the high- to medium-SNR regime, the Kronecker model overestimates capacity at high levels of reliability (and reliability at low data rates) and *vice versa*. While experimental/measurement evidence has been put forth for some of these trends elsewhere [6], our work provides the first rigorous and systematic framework to explain these observations.

The first author is with the University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA. The second author is with the University of Wisconsin, Madison, WI 53706, USA. The third author is with the Freescale Semiconductor Inc., Austin, TX 78721, USA. Email: `vasanth@uiuc.edu`.

This work was partly supported by the NSF under grant #CCF-0431088 through the University of Wisconsin.

[1] Let $\mathbf{H}$ be an $N_r \times N_t$ random matrix with independent entries and let the variance of $\mathbf{H}[i,j]$ be given by $\mathbf{P}_c[i,j]$. A channel is called column-regular if $\{\sum_{i=1}^{N_r} \mathbf{P}_c[i,j]\}$ is independent of j, row-regular if the above condition is true for $\mathbf{H}^T$, and regular if it is both row- and column-regular [8]. Otherwise, it is non-regular.

2. SYSTEM MODEL

We consider a narrowband, fading MIMO channel with N_t transmit and N_r receive antennas. The $N_t \times 1$ transmitted vector $\mathbf{x}$ and the $N_r \times 1$ received vector $\mathbf{y}$ are related by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{H}$ is the $N_r \times N_t$ channel matrix and $\mathbf{n}$ is the independent, white Gaussian noise added at the receiver. It has been well-documented that the assumption of zero mean Rayleigh fading is an accurate model for $\mathbf{H}$ in a non line-of-sight setting. Thus the complete channel statistics are described by the second-order moments. Rich scattering environments are accurately modeled by the commonly used *i.i.d. model* where the channel entries are i.i.d. $\mathcal{CN}(0,1)$. However, the i.i.d. model is not accurate in describing realistic propagation environments. Various statistical models have been proposed to overcome the deficiencies associated with the i.i.d. model.

The most general framework of *canonical modeling* [5, 6, 7] assumes that the auto- and cross-covariance matrices on both the transmitter and the receiver sides have the same eigen-bases, and exploits this redundancy to decompose $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{U}_r \mathbf{H}_{\sf ind} \mathbf{U}_t^H \tag{2}$$

where $\mathbf{H}_{\sf ind}$ has independent, but not necessarily identically distributed entries. $\mathbf{U}_r$ and $\mathbf{U}_t$ are eigenvector matrices corresponding to the receive and the transmit covariance matrices which are defined as $\mathbf{\Sigma}_r = E[\mathbf{H}\mathbf{H}^H]$ and $\mathbf{\Sigma}_t = E[\mathbf{H}^H\mathbf{H}]$, respectively.

A special case of the above model is the Kronecker model where $\mathbf{H}_{\sf ind}$ is assumed to be $\mathbf{\Lambda}_r^{1/2}\mathbf{H}_{\sf iid}\mathbf{\Lambda}_t^{1/2}$ with $\mathbf{\Lambda}_t$ and $\mathbf{\Lambda}_r$ diagonal. While the canonical model is characterized by N_tN_r parameters corresponding to $\{\mathbf{P}_c[i,j] = E[|\mathbf{H}_{\sf ind}[i,j]|^2]\}$, the Kronecker model is parameterized by fewer parameters: the $N_t + N_r$ diagonal entries of $\mathbf{\Lambda}_t$ and $\mathbf{\Lambda}_r$. Given a channel $\mathbf{H}_c$ with $\mathbf{H}_c[i,j] \sim \mathcal{CN}(0, \mathbf{P}_c[i,j])$ where $\mathbf{P}_c[i,j]$ are non-separable, a channel $\mathbf{H}_k$ whose entries have separable variances can be fitted according to the relationship: $\mathbf{H}_k[i,j] \sim \mathcal{CN}(0, \mathbf{P}_k[i,j])$ with

$$\mathbf{P}_k[i,j] = \frac{\sum_k \mathbf{P}_c[i,k] \cdot \sum_l \mathbf{P}_c[l,j]}{\sum_{kl} \mathbf{P}_c[k,l]}. \tag{3}$$

Note that the mapping in (3) may considerably increase the DoF in the Kronecker representation of a scattering environment described by (2). In general, the Kronecker model spreads the degrees of freedom across the resulting $\mathbf{P}_k$ and thereby 'flattens' it since its statistics are based only on column and row sum statistics of the true spatial power matrix.

We assume that the receiver can estimate $\mathbf{H}$ (perfectly) at essentially zero cost. Under the assumption that the true statistics of the channel coefficients are non-separable and the channel statistics do not change over a sufficiently long duration, long-term averages of the sample covariance matrix and the squared magnitude of the channel entries lead to reliable estimates for the eigen-matrices and the variances of channel entries at the receiver. Since the eigen-directions carry more information than the variances, we assume that these are fed back perfectly. To minimize the cost of statistical feedback overhead, the receiver may convey only a subset of the information in the variances to the transmitter. As described above, a simple method to reduce the number of variance parameters is to fit a Kronecker model to a canonical channel as in (3). In this work, we assume that the independent parameters in $\mathbf{P}_k$ are fed back to the transmitter perfectly.

In the above setting, the ergodic (or average) capacity of a MIMO channel at a transmit SNR of ρ is given by [2]

$$C_{\sf erg}(\rho) = \sup_{\mathbf{Q}\,:\,\mathbf{Q}\geq \mathbf{0},\ \mathrm{Tr}(\mathbf{Q})\leq\rho} E_{\mathbf{H}}\left[\log\det\left(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\right]$$

where the optimization is over the set of trace-constrained, positive semi-definite matrices. It is known that [7, 9] the optimal $\mathbf{Q}$ reduces to beamforming along the statistically dominant eigen-mode in the low-SNR extreme and uniform power signaling in the high-SNR extreme. Thus, we have

$$\begin{aligned} C_{\sf low}(\rho) &= E\left[\log_2\left(1 + \rho\sum_i \left|\mathbf{H}_{\sf ind}[i, j_{\max}]\right|^2\right)\right] && (4) \\ &= \log_2(e)\cdot\rho\cdot\sum_i \mathbf{P}_c[i, j_{\max}]\cdot(1 + o(1)) && (5) \\ C_{\sf high}(\rho) &= E\left[\log_2\det\left(\mathbf{I}_{N_r} + \frac{\rho}{\mathrm{rank}(\mathbf{\Sigma}_t)}\mathbf{H}_{\sf ind}\mathbf{H}_{\sf ind}^H\right)\right] \end{aligned}$$

where $j_{\max} = \arg\max_j \sum_i \mathbf{P}_c[i,j]$. Since ergodic capacity is insufficient to completely characterize performance, the more reasonable notion of outage capacity at an outage probability of $q\%$ is necessary. Following a Gaussian approximation for the capacity random variable, this is given by

$$C_{{\sf out},\,q}(\rho) = C_{\sf erg}(\rho) - x_q\sqrt{V(\rho)} + o(1) \tag{6}$$

where x_q is the unique solution to ${\sf erfc}(x_q/\sqrt{2}) = 2q$ with $V(\rho)$ and ${\sf erfc}(\cdot)$ denoting the variance of the capacity random variable and the complementary error function, respectively.

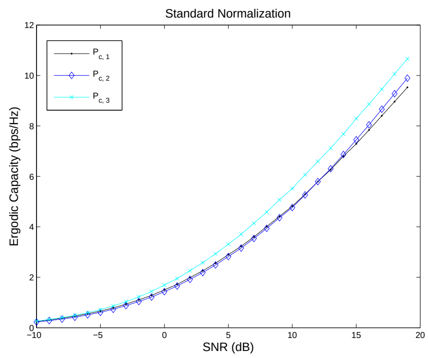

Fig. 1. Impact of the richness of scattering environments on the ergodic capacity under the standard channel power normalization. More correlated channels show a larger capacity in the low-SNR extreme with a cross-over SNR from which onwards the more richer channel shows higher capacity.

3. CHANNEL POWER NORMALIZATION

The standard channel power normalization used in the MIMO literature is $\rho_c = N_tN_r$, a legacy of the i.i.d. model so that the total channel power is identical irrespective of the environment. Under the above normalization, Fig. 1 illustrates the impact of progressively rich scattering on capacity when statistical information is available at the transmitter for the following channels:

$$\mathbf{P}_{c,1} = \frac{4}{2.5} \times \begin{bmatrix} 1 & 0 \\ 1 & 0.5 \end{bmatrix}, \mathbf{P}_{c,2} = \frac{4}{3.3} \times \begin{bmatrix} 1 & 0.8 \\ 1 & 0.5 \end{bmatrix},$$

and $\mathbf{P}_{c,3}$, the 2×2 i.i.d. channel. We see that in a highly correlated channel, the received power in the dominant column is higher than that of an i.i.d. channel where isotropic transmission is optimal. Hence, correlated channels have higher capacity than i.i.d. channels below some critical SNR value, as observed in many recent works [10, 9]. This suggests that the above normalization has to be examined more carefully.

We now propose an alternate normalization where we compare different environments by assuming equal fading gain per DoF in all cases. That is, $\mathbf{P}_c[i,j] \leq 1$ holds for all i, j, irrespective of the number of antennas or the scattering environment. The i.i.d. channel can be obtained by setting $\mathbf{P}_c[i,j] = 1$ for all i, j while a correlated/sparse scattering is obtained by setting $\mathbf{P}_c[i,j]$ to zero (or close to zero) for some indices of $\{i, j\}$. The more the indices that are set to zero, the more correlated/sparse the scattering environment is.

Fig. 2 illustrates the capacity of the above three channels with the proposed channel power normalization. Note that the slope of the capacity curves corresponding to the three channels are the same at high-SNR since the spatial multiplexing gain is the same for these channels ($\text{rank}(\mathbf{P}_{c,i}) = 2$). Since a richer scattering environment projects more power to the receiver than a sparser scattering environment, it is to be expected that the i.i.d. channel has the highest capacity among all channels, irrespective of the SNR. In fact, in [5] we claim that as a function of the statistics, $C_{\text{erg}}(\rho)$ is maximized under the proposed normalization when $\mathbf{H}$ is i.i.d., that is, $\mathbf{P}_c[i,j] = 1$ for all i, j.

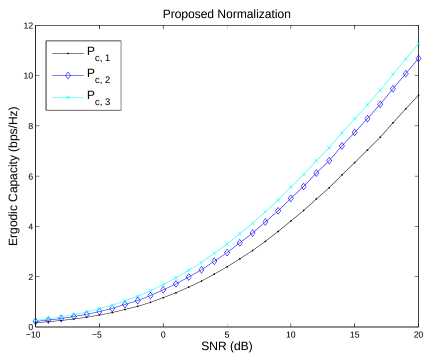

Fig. 2. Impact of the richness of scattering environments on the ergodic capacity with the proposed channel power normalization. Under this normalization, the richer channel always leads to larger capacity for any SNR.

4. KRONECKER VS. CANONICAL MODELS

Our focus in this section is on comparing the capacities (under the two normalizations) when the true statistics follow the canonical model (This channel is denoted by $\mathbf{H}_c$.) while we fit a Kronecker model for this channel according to (3) (This channel is denoted by $\mathbf{H}_k$.). While this study for an arbitrary choice of transmit SNR and N_t, N_r seems difficult, we can obtain fundamental insights by studying this problem at the low- and the high-SNR extremes in the antenna asymptotic limit. We expect these conclusions to be meaningful for reasonably large antenna dimensions. The main conclusions[2] are as follows.

Theorem 1 *In the low-SNR extreme, for the case of regular channels, the ergodic capacity under both models is the same for either normalization. In particular, the dominant terms satisfy:*

$$C_{\text{erg},c,1}(\rho) = C_{\text{erg},k,1}(\rho) = \frac{\log_2(e)\rho N_t N_r}{\rho_c} \cdot \sum_i \mathbf{P}_c[i, j_{\max}],$$
$$C_{\text{erg},c,2}(\rho) = C_{\text{erg},k,2}(\rho) = \log_2(e)\rho \cdot \sum_i \mathbf{P}_c[i, j_{\max}]$$

with equality under the two normalizations if and only if $\mathbf{H}_c = \mathbf{H}_k \leftrightarrow \mathbf{H}_{\text{iid}}$. *Furthermore, the variances satisfy:*

$$\left(\frac{\log_2(e)\rho N_t N_r}{\rho_c}\right)^2 \cdot \sum_i \left(\mathbf{P}_c[i, j_{\max}]\right)^2 = V_{c,1}(\rho) \geq$$
$$V_{k,1}(\rho) = \left(\frac{\log_2(e)\rho N_t N_r}{\rho_c}\right)^2 \cdot \sum_i \left(\mathbf{P}_k[i, j_{\max}]\right)^2,$$
$$(\log_2(e)\rho)^2 \cdot \sum_i \left(\mathbf{P}_c[i, j_{\max}]\right)^2 = V_{c,2}(\rho) \geq$$
$$V_{k,2}(\rho) = (\log_2(e)\rho)^2 \cdot \sum_i \left(\mathbf{P}_k[i, j_{\max}]\right)^2$$

with equality in either case if and only if $\mathbf{H}$ *is i.i.d. Also, the standard normalization always results in higher variance with equality if and only if* $\mathbf{H}$ *is i.i.d.*

The same expressions continue to hold for non-regular channels, and it can be observed that $V_{c,i}(\rho) > V_{k,i}(\rho)$ *under either normalization;* $V_{c,1}(\rho) > V_{c,2}(\rho)$*; and* $V_{k,1}(\rho) > V_{k,2}(\rho)$.

From the above theorem, it is to be noted that the ergodic capacities remain the same under the canonical and the Kronecker models for either normalization, irrespective of whether the channel is regular or non-regular. Thus, the dominant factors in understanding outage capacity in (6) are the variances of capacity. And, except for the i.i.d. channel, the outage capacity under the canonical model is less steeper than the outage capacity under the Kronecker model.

To understand the high-SNR regime, we need to make two assumptions on the random matrix channel $\mathbf{H}_c$ to aid in capacity analysis: 1) $N_t = N_r = N$, and 2) $\text{rank}(\mathbf{H}_c) = N$ a.s. Note that the second condition is equivalent to the assumption that none of the transmit or the receive eigenvalues ($\{\sum_i \mathbf{P}_c[i,j]\}$ or $\{\sum_j \mathbf{P}_c[i,j]\}$, respectively) are zero. Also, note that since $\mathbf{P}_k[i,j]$ are determined by the sum statistics of $\mathbf{P}_c[i,j]$, the non-redundancy assumption above means that $\mathbf{P}_k[i,j] > 0$ for all i, j. Then, it follows that $\text{rank}(\mathbf{H}_k) = N$ a.s.. In this setting, the capacity random variables under the two models (and normalization i) are given by

$$C_{c,i}(\rho, \mathbf{H}) = \log_2 \det\left(\mathbf{H}_c \mathbf{H}_c^H\right) + \kappa_i$$
$$C_{k,i}(\rho, \mathbf{H}) = \log_2 \det\left(\mathbf{H}_k \mathbf{H}_k^H\right) + \kappa_i \quad (7)$$

where κ_i is a constant independent of the statistics (but dependent on the normalization) up to an $\mathcal{O}(\frac{1}{\rho})$ term. Hence the

[2] See [5] for proofs.

statistics of $C_{c,i}(\rho, \mathbf{H})$ and $C_{k,i}(\rho, \mathbf{H})$ at high-SNR are related to the moments of $\log\det\left(\mathbf{H}_c\mathbf{H}_c^H\right)$ and $\log\det\left(\mathbf{H}_k\mathbf{H}_k^H\right)$, respectively. Towards studying the statistics of these log determinants, we need the following result.

Lemma 1 *Let $\mathbf{H}_c[i,j]$ be independent and distributed as $\mathcal{CN}(0, \mathbf{P}_c[i,j])$. Then, there exist independent random variables $\widetilde{\mathbf{Z}}_i$, $i = 1 \cdots N$ on some probability space such that $\widetilde{\mathbf{Z}}_i \sim i\,\frac{\sum_{j=1}^{N} |\mathbf{H}_c[i,j]|^2}{N}$ and $\det\left(\mathbf{H}_c\mathbf{H}_c^H\right)$ can be well approximated by $\prod_{i=1}^{N} \widetilde{\mathbf{Z}}_i$.*

Using the above lemma, we have the following main result.

Theorem 2 *In the high-SNR extreme and the asymptotics of N, for the case of regular channels, we have $C_{\mathsf{erg},c,i}(\rho) = C_{\mathsf{erg},k,i}(\rho)$ for either normalization. Furthermore,*

$$C_{\mathsf{erg},c,1}(\rho) \geq C_{\mathsf{erg},c,2}(\rho), \;\; C_{\mathsf{erg},k,1}(\rho) \geq C_{\mathsf{erg},k,2}(\rho)$$

with equality in either case if and only if $\mathbf{H}$ is i.i.d. In the non-regular case, $C_{\mathsf{erg},c,i}(\rho) > C_{\mathsf{erg},k,i}(\rho)$.

Further, for both regular as well as non-regular channels, $V_{c,i}(\rho)$ and $V_{k,i}(\rho)$ are independent of the normalization.

While Lemma 1 results in closed-form expressions for means of capacity [5], the same for the variances seems difficult. However, numerical studies suggest that for most scattering environments $(V_{c,i}(\rho))^{\frac{1}{2}}$ and $(V_{k,i}(\rho))^{\frac{1}{2}}$ are sub-dominant when compared with $C_{\mathsf{erg},c,i}(\rho)$ and $C_{\mathsf{erg},k,i}(\rho)$, respectively. Thus the outage capacities with the two models are primarily determined by $C_{\mathsf{erg},c,i}(\rho)$ and $C_{\mathsf{erg},k,i}(\rho)$ for most generic (non-regular) scattering environments. More importantly, the smoothing effect of the Kronecker model as can be seen from (3), and the low-SNR trends of $V_{\bullet}(\rho)$ lend credence to the conjecture that $V_{c,i}(\rho) \geq V_{k,i}(\rho)$, even in the medium- to high-SNR regime. In the ensuing discussion, we will assume these two observations without proof.

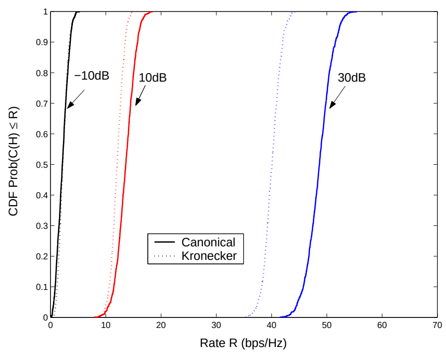

Fig. 3. Comparison of capacity CDFs of a sparse scattering environment with canonical and Kronecker models at -10, 10 and 30 dB SNRs.

5. DISCUSSION

We now assume the proposed channel power normalization. In the case of non-regular channels, the fact that $C_{\mathsf{erg},\,c}(\rho) > C_{\mathsf{erg},\,k}(\rho)$ (Theorem 2) and the sub-dominance assumption of $V_{\bullet}(\rho)$ implies that the Kronecker model underestimates capacity confirming the observations made in recent measurement campaigns [6]. This result is also illustrated in Fig. 3 where the cumulative distribution functions (CDFs) of capacity (with a sparse scattering environment) for each model at -10 dB, 10 dB and 30 dB SNR are plotted. Note that the Kronecker model CDF curves are steeper, indicating smaller variance than that predicted by the canonical model. Also, note that low outage probability corresponds to high levels of operational reliability and hence the Kronecker model also underestimates reliability for all data rates.

For any channel in the low-SNR regime (as well as regular channels in the high-SNR regime), $C_{\mathsf{erg},\,c}(\rho) = C_{\mathsf{erg},\,k}(\rho)$ and the precise outage-reliability characterization is determined by $V_c(\rho)$ and $V_k(\rho)$. Using $V_c(\rho) \geq V_k(\rho)$, we see that at high levels of operational reliability, the Kronecker model overestimates capacity while it switches roles and underestimates capacity at low levels of reliability. Rephrasing this observation, the Kronecker model overestimates reliability at low data rates and it underestimates reliability at high rates. However, these trends are not that prominent in Fig. 3 due to the smallness of capacity values at low-SNRs. Furthermore, as the scattering becomes more richer (and the channel more close to being regular), the gap in the ergodic capacities between the two models becomes smaller. Thus, our analysis provides a theoretical foundation for many recent measurement-based observations.

6. REFERENCES

[1] Í. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–596, Nov. 1999.

[2] G. J. Foschini, "Layered Space-Time Architechture for Wireless Communication in a Fading Environment when Using Multi-Element Antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] C. N. Chuah, J. M. Kahn, and D. N. C. Tse, "Capacity Scaling in MIMO Wireless Systems under Correlated Fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[4] A. M. Sayeed, "Deconstructing Multi-Antenna Fading Channels," *IEEE Trans. Sig. Proc.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[5] V. Raghavan, J. H. Kotecha, and A. M. Sayeed, "Canonical Statistical Modeling and Capacity Analysis of Correlated MIMO Fading Channels," *Submitted to IEEE Trans. Inform. Theory*, Oct. 2007, Available: [Online] `http://www.ifp.uiuc.edu/~vasanth`.

[6] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A Stochastic MIMO Channel Model with Joint Correlation of Both Link Ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan 2006.

[7] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of Antenna Correlation on the Capacity of Multiantenna Channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2491–2509, July 2005.

[8] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-Achieving Input Covariance for Correlated Multi-Antenna Channels," *Allerton Conf. Commun. Cont. and Comp.*, 2003.

[9] V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO Rayleigh Fading Channels: Capacity, Optimal Signaling and Asymptotics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2058–2072, June 2005.

[10] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE Journ. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.