# SOME ASYMPTOTIC CAPACITY RESULTS FOR MIMO WIRELESS WITH AND WITHOUT CHANNEL KNOWLEDGE AT THE TRANSMITTER

*Mai Vu, Arogyaswami Paulraj*

Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA 94305-9510, USA
E-mails: {mhv, apaulraj}@stanford.edu

***Abstract–*** **We study the asymptotic capacity of MIMO wireless channels predicated on channel knowledge (or channel state information - CSI) at the transmitter in the high SNR regime. The transmitter CSI studied includes instantaneous and transmit correlation types. The receiver is assumed to know the channel perfectly. Analytical expressions for the asymptotic ergodic capacity are derived in each case, which closely match the simulations. The results show that transmitter CSI in MIMO wireless can have a strong influence on capacity.**

## 1. INTRODUCTION

Recently channel knowledge at the transmitter in MIMO wireless have become a subject of interests due to its potential practicality. Studies in MIMO wireless capacity have incorporating some transmitter CSI, either in instantaneous forms or in long term statistics, and derive the transmission characteristics to achieve this capacity [4, 5, 6, 7].

In this paper, we examine MIMO wireless capacity with transmitter CSI and the effect of channel knowledge upon capacity in the high SNR regime. We study two types of channel state information at the transmitter: instantaneous CSI and correlation CSI. In instantaneous CSI case, two extremes of CSI, channel fully known and not known, are considered. In the correlation CSI case, we use a simplified correlation model where the transmit antennas exhibit the same correlation to all the receive antennas. Again two extremes of the correlation matrix $R_t$ full-rank and rank-one are considered, with $R_t$ known and not known at the transmitter. The purpose of studying these cases is to obtain the bound on the performance in practical situations when Tx-CSI will lie between these extremes.

In each case, asymptotic analytical formula for evaluating the capacity are given, which matches closely with numerical simulation results with a finite number of antennas at sufficiently high SNRs (from above 15dB in most cases). It is shown that in the instantaneous CSI case, there is a capacity gain from channel knowledge if the number of transmit antennas is larger than the number of receive antennas, and this gain grows linearly with the number of receive antennas while keeping the antenna ratio constant. In the correlation CSI case, a gain is obtained from knowing the transmit correlation matrix if this matrix is rank-deficient. The formulae also give some insights for the effect of transmit correlation on channel capacity, which is detrimental in most cases. However, at low SNRs, transmit correlation can sometimes help to increase the capacity over i.i.d channels without transmitter CSI.

The paper is organized as follows: The next Section sets up the channel models and the capacity framework. Section 3 presents analysis and simulation results for the instantaneous CSI case. Correlation CSI is studied in Section 4. We close with some concluding remarks in Section 5.

## 2. CHANNEL MODELS AND CAPACITY

### 2.1. Channel models

We consider frequency-flat Rayleigh fading MIMO wireless channels with $M_t$ transmit and $M_r$ receive antennas. The channel response can be represented by a matrix $\mathbf{H}$ of size $M_r \times M_t$ with random entries, which are zero mean complex Gaussian distributed.

When no transmit correlation exists, the elements of $\mathbf{H}$ are independent and identically distributed with variance normalized to one, where the real and imaginary components are independent and have equal variances.

When there is some transmit correlation given by a correlation matrix $\mathbf{R}_t$, the channel can be written as

$$\mathbf{H} = \mathbf{H}_w \mathbf{R}_t^{\frac{1}{2}} , \quad (1)$$

where $\mathbf{H}_w$ contains i.i.d complex Gaussian entries with zero mean and unit variance. To enable comparison between different correlation structures, we normalize the total power in the correlation matrix so that $\text{tr}(\mathbf{R}_t) = M_t$.

We define here some quantities which will be used later:

$$\begin{aligned} M &= \max(M_t, M_r) \\ k &= \text{rank}(R_t) \\ r &= \min(k, M_r) \\ m &= \max(k, M_r) . \end{aligned} \quad (2)$$

Note that $1 \le k \le M_t$.

Mai Vu's work was supported in part by the Rambus Stanford Graduate Fellowship.

This work was also supported in part by grants from the National Science Foundation (Grant Nos. CCR-0241919 and CCR-0241921), Intel Corp. and Sprint Corp.

### 2.2. Channel ergodic capacity

Assuming that the receiver always knows the channel perfectly. The channel ergodic capacity, under a total average transmit power constraint, is then achieved by Gaussian input signal with zero mean and a normalized covariance matrix $\mathbf{Q}$ that is the solution of a optimization problem [1]. With various assumptions of the transmitter CSI, the capacity is

$$\begin{aligned} \mathcal{C} &= \max_{\mathbf{Q}} E[\log\det(\mathbf{I}+\gamma\mathbf{HQH}^*)] \quad (3) \\ \text{s.t.} \quad & \text{tr}(\mathbf{Q}) = 1\,, \quad \mathbf{Q} \geq 0 \\ & \text{transmitter CSI}\,, \end{aligned}$$

where $\gamma$ is the SNR.

Under the assumption of high SNR, the capacity expression can be approximated by

$$\mathcal{C} \approx \max_{\mathbf{Q}} E[\log\det(\gamma\mathbf{HQH}^*)]\,. \quad (4)$$

Note that this is always lower that the actual capacity and only equals asymptotically (as $\gamma \to \infty$). The valid range of SNR for this assumption will be justified through numerical simulations later. In most cases we found that the approximation works well with a SNR from above 15dB.

## 3. INSTANTANEOUS TX-CSI

In this section, we will derive formulae and compare the channel capacity for two cases of Tx-CSI: full channel knowledge and no channel knowledge at the transmitter. The underlying channel is assumed to have no transmit correlation (i.e. $\mathbf{R}_t = \mathbf{I}$).

### 3.1. Full Tx-CSI

When the transmitter knows the CSI fully, the instantaneous capacity is given by the well-known water-filling algorithm. This distributes the available transmit power over the eigenmodes of the channel realization based on the eigenvalues of the channel [3]. Since the channel is i.i.d, each realization is full-rank and there will be $r$ non-zero eigenvalues. The ergodic capacity is then the expected value of the instantaneous capacity over the channel eigenvalues. At sufficiently high SNR, the optimum power allocation along channel eigenmodes are approximately equal, and the capacity can be approximated by

$$\begin{aligned} \mathcal{C} &= \mathbf{E}_{\lambda_i}\Big[\sum_{i=1}^{r}\log\big(\frac{\gamma}{r}\lambda_i\big)\Big] \quad (5) \\ &= r\log(\gamma) + r\log\Big(\frac{M}{r}\Big) + \mathbf{E}_{\lambda_i}\Big[\sum_{i=1}^{r}\log\Big(\frac{\lambda_i}{M}\Big)\Big]\,, \end{aligned}$$

where $\lambda_i$ are the non-zero eigenvalues of $\mathbf{HH}^*$.

To evaluate the last expectation term in (5), we use the following result from random matrix theory. Suppose that the number of transmit and receive antennas are very large such that in the limit, the ratio between them $\lim_{M\to\infty}\frac{r}{M} = c$ is a constant (note that $c \leq 1$ in our setup here). Then the spectral distribution of the eigenvalues of $\mathbf{HH}^*/M$ will converge to a deterministic function, which can be given explicitly in this case by Marcenko and Pastur [2] as

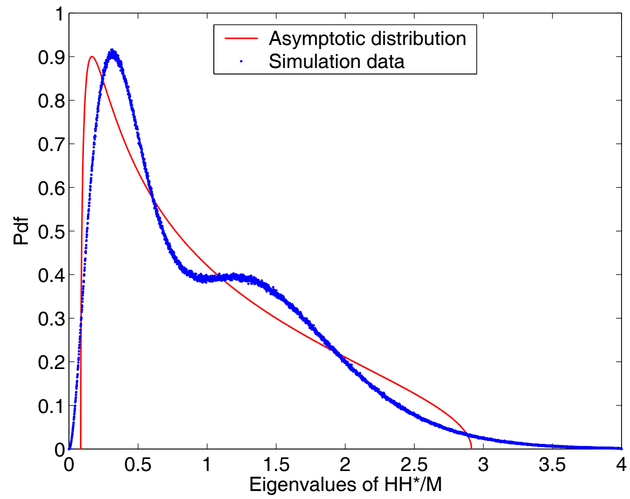


**Fig. 1. Spectral distribution of eigenvalues of $\frac{1}{4}\mathbf{HH}^*$, where $\mathbf{H}$ is a $2\times4$ random matrix with i.i.d Gaussian entries of zero mean and unit variance.**

$$f_\lambda(x) = \frac{\sqrt{\big((1+\sqrt{c})^2 - x\big)\big(x-(1-\sqrt{c})^2\big)}}{2\pi xc}\,, \quad (6)$$

where the support set of the eigenvalues is $(1-\sqrt{c})^2 \leq x \leq (1+\sqrt{c})^2$.

It turns out that the p.d.f of an eigenvalue of $HH^*/M$ for finite antenna cases can be approximated quite closely by this function. An example of the distribution for $r = 2$ and $M = 4$ using Monte-Carlo simulation, is plotted in Fig. 1 together with the asymptotic distribution. The approximation gets tighter as the number of antennas grows larger.

Thus using (6) as an approximation for the p.d.f of $\lambda_i/M$ in finite antenna cases, the capacity equation (5) can be evaluated analytically. The asymptotic calculations are plotted together with simulation results in Fig. 2. Note that in the simulations, we do not use the high SNR assumption (i.e. we use equation (3)), and the instantaneous capacity calculation is exact using the water-filling algorithm. The asymptotic calculation and the simulation results are indistinguishable at SNRs from above 15dB. It is noticed that this SNR threshold, above which the asymptotic calculation is the same as simulation, changes with the number of antennas. The more antennas used while keeping the transmit to receive antenna ratio constant, the higher the SNR required.

### 3.2. No Tx-CSI

When the channel is not known at the transmitter, the optimum transmit strategy is to send i.i.d Gaussian signals with

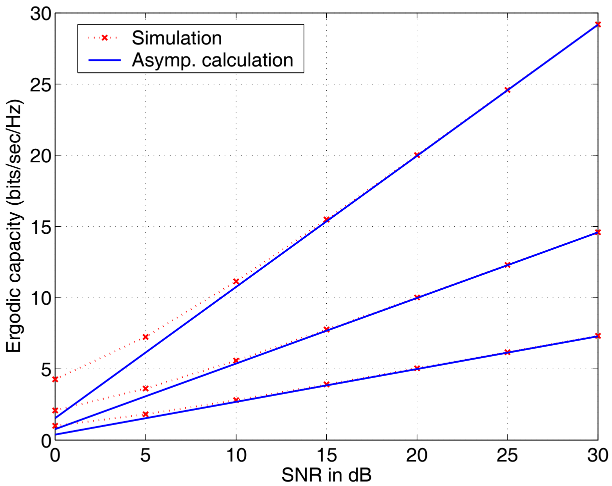


**Fig. 2**. **Capacity for full Tx-CSI i.i.d channels at a fixed transmit to receive antenna ratio of** 2**, and the numbers of receive antennas are** $[1, 2, 4]$**, going from bottom to top pairs.**

equal power in all directions [1]. The ergodic capacity at sufficiently high SNR can be approximated as:

$$\begin{aligned} \mathcal{C} &= \mathbf{E}_{\lambda_i}\Big[\sum_{i=1}^{r} \log\big(\frac{\gamma}{M_t}\lambda_i\big)\Big] \qquad (7) \\ &= r\log(\gamma) + r\log\Big(\frac{M}{M_t}\Big) + \mathbf{E}_{\lambda_i}\Big[\sum_{i=1}^{r}\log\Big(\frac{\lambda_i}{M}\Big)\Big] \ . \end{aligned}$$

The only difference between (5) and (7) is in the number of modes that the transmit power is distributed, which are $r$ and $M_t$ respectively. It can be inferred immediately that if $M_t \leq M_r$, there is no gain in instantaneous channel knowledge at sufficiently high SNRs. However, if the number of transmit antennas is larger than the number of receive antennas ($M_t > M_r$), there is an additional capacity gain of $M_r \log(M_t/M_r)$ obtained from knowing the channel at the transmitter. At a large number of antennas, this gain can be quite substantial.

$$\mathcal{C}_{\text{gain, inst CSI}} = \max\Big\{M_r \log\Big(\frac{M_t}{M_r}\Big), 0\Big\} \ . \qquad (8)$$

The capacity gain (8) is illustrated in Fig. 3, where it is added to the simulation results for no Tx-CSI and then compared with simulation results for full Tx-CSI. The plot shows that the gain is exact at high SNRs (from above 15dB), but is too optimistic at lower SNRs.

## 4. CORRELATION TX-CSI

In this part, we study the channel with some correlation structure between the transmit antennas. It is assumed that all receive antennas exhibit independent fading from each other, whereas all transmit antennas exhibit the same correlation structure to every receive antenna. The channel model (1)

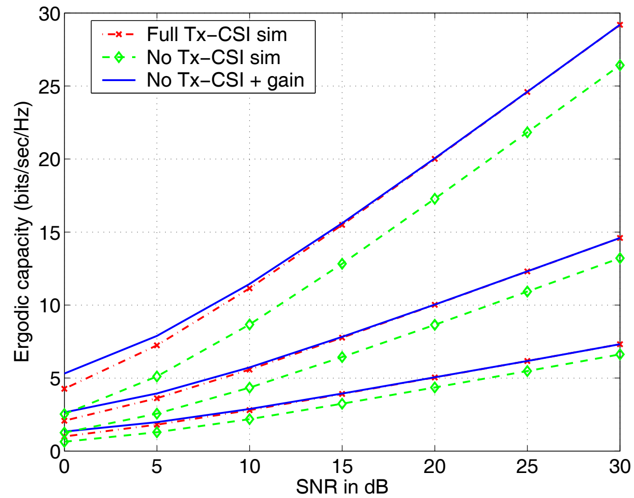


**Fig. 3**. **Capacity gain (8) when added to no Tx-CSI i.i.d channel capacity simulation results. The transmit to receive antenna ratio is** 2**, and the number of receive antennas are** $[1, 2, 4]$**, going from bottom to top groups.**

hence applies here. Again we will consider two cases of Tx-CSI: when the transmitter knows the correlation matrix $\mathbf{R}_t$ and when it does not.

### 4.1. Transmit correlation known

In this case, it can be shown that the optimum transmit strategy is to distribute power along the eigenmodes of the correlation matrix [5, 7]. At sufficiently high SNR, the ergodic capacity is given by

$$\begin{aligned} \mathcal{C} &\approx \max_{\mathbf{\Lambda}} E_{\mathbf{H}_\omega}\Big[\log\det\big(\gamma\mathbf{H}_\omega\mathbf{\Xi}\mathbf{\Lambda}\mathbf{H}_\omega^*\big)\Big] \qquad (9) \\ \text{s.t.} \quad & \text{tr}(\mathbf{\Lambda}) = 1\ , \quad \mathbf{\Lambda} \geq 0\ , \end{aligned}$$

where $\mathbf{\Xi}$ is the diagonal matrix of the non-zero eigenvalues of $\mathbf{R}_t$ and $\mathbf{\Lambda}$ is the diagonal transmit power allocation on the eigenmodes of $\mathbf{R}_t$.

#### Case of $k \leq M_r$

When the rank of the correlation matrix $k$ is not larger than the number of receive antennas $M_r$, the determinant expression in (9) can be broken into a product of determinants. It can then be shown that at high SNRs, distributing power equally along all $k$ available modes is asymptotically optimum, and the capacity expression can be given explicitly as

$$\mathcal{C} = k\log(\gamma)\mathbf{\Xi}) + k\log\Big(\frac{m}{k}\Big) + \mathbf{E}_{\lambda_i}\sum_{i=1}^{k}\log\Big(\frac{\lambda_i}{m}\Big) \ . \qquad (10)$$

Again the last expectation term can be evaluated using the approximate p.d.f in (6). Comparing to (5), and noting that $k = r$ here, expression (10) shows that at high SNRs, knowing only the antenna correlation can be as good as having the

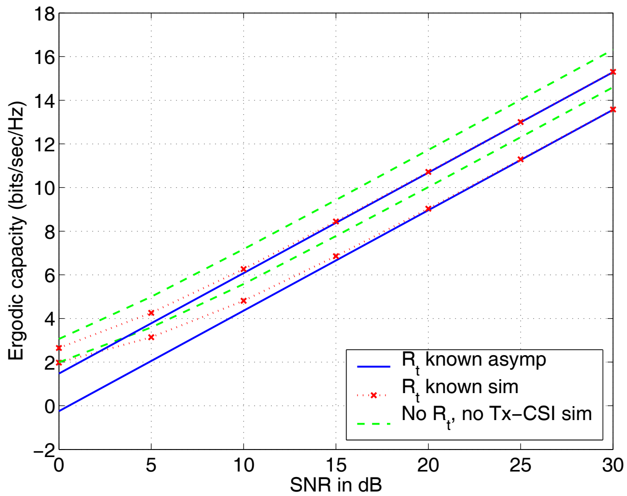


**Fig. 4. Capacity of systems with $2$ transmit antennas and full rank transmit correlation which has eigenvalues $[1.8\ 0.2]$. The receive to transmit antenna ratios are $[2, 4]$ from bottom to top groups. Simulation results of capacity for i.i.d channels with the same antenna configurations but no Tx-CSI are superimposed for comparison.**

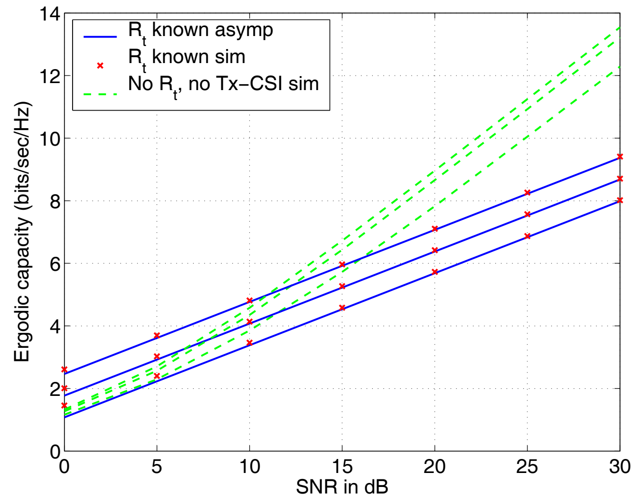


**Fig. 5. Capacity of systems with $2$ receive antennas and rank one transmit correlation. The transmit to receive antenna ratios are $[1, 2, 4]$ from bottom to top groups. Simulation results of capacity for i.i.d channels with the same antenna configurations but no Tx-CSI are superimposed for comparison.**

full channel knowledge when correlation exists (where $m$ is equivalent to $M$). There is an additional term of $\log\det(\mathbf{\Xi})$ in capacity caused by the correlation at the transmitter. When the transmit correlation matrix is full rank ($k = M_t$), this term is always non-positive since tr$(\mathbf{\Xi})$ = tr$(\mathbf{R}_t) = M_t$, thus represents a penalty in capacity due to transmit correlation. It is zero only when $\mathbf{R}_t = \mathbf{I}$. When transmit correlation is not full rank, the additional term can be positive. However, since $k \le M_r$ here, lower rank correlation matrix will reduce the number of effective modes in the channel. At sufficiently high SNRs, (which is usually much lower than the SNR required for the asymptotic capacity equations to hold), this leads to a reduction in capacity.

It is also interesting to compare capacity for known transmit correlation (10) with capacity for an i.i.d channel without transmitter CSI (7), where the latter may apply when the channel is varying too fast to be known at the transmitter. With a smaller number of transmit than receive antennas and full rank correlation, (10) is always less than (7). This is illustrated in Fig. 4, where simulation results for i.i.d channels capacity are used for comparison instead of the asymptotic capacity in (7) for a better accuracy. The simulated capacity results for the correlated channels are exact using the optimum transmit power distribution over the eigenvectors of $\mathbf{R}_t$. This optimum power distribution is found via numerical convex optimization, however, it is very close to the approximation of water-filling over the eigenvalues of $\mathbf{R}_t$ with accuracy improves at higher SNRs and more antennas. Again from Fig. 4, the asymptotic calculation for capacity with correlation Tx-CSI matches well with the simulation results for SNRs from above 15dB. The SNR required for the asymptotic results to apply decreases as the number of transmit or receive antennas increases relative to the correlation rank.

With a larger number of transmit than receive antennas and a rank-deficient transmit correlation, correlation may gain capacity over i.i.d channels for some SNRs. Fig. 5 gives an example of this case when the correlation is rank one. However, at sufficiently high SNRs, correlation will always reduce the ergodic capacity regardless of the correlation rank.

**Case of $k > M_r$**

When the rank of the transmit correlation matrix $\mathbf{R}_t$ is more than the number of receive antennas ($k > M_r$), it does not yet appear that an explicit expression for the capacity can be obtained. It is known that as the number of antennas grows large, the spectral distribution of the eigenvalues of $\frac{1}{k}\mathbf{H}_\omega \mathbf{\Xi} \mathbf{D} \mathbf{H}_\omega^*$ will also converge to a deterministic function [2, 8]. Obtaining the explicit form of this distribution function involves solving a polynomial equation of degree $k+1$ analytically to find the Stieltjes transform of the c.d.f, and then carrying out the inverse transform. While distribution functions for $k \le 3$ may be obtained in closed form [9], no general solution appears to exist.

### 4.2. Transmit correlation unknown

When transmit correlation exists but is not known at the transmitter, the best the transmitter can do is to distribute power equally in all directions, that is $\mathbf{\Lambda} = \mathbf{I}/M_t$. In this case the asymptotic capacity at sufficiently high SNRs is given by

$$\mathcal{C} \approx E_{\mathbf{H}_\omega}\left[\log\det\left(\frac{\gamma}{M_t}\mathbf{H}_\omega \mathbf{\Xi} \mathbf{H}_\omega^*\right)\right] \qquad (11)$$

Again when $k \le M_r$, the capacity can be evaluated ex-

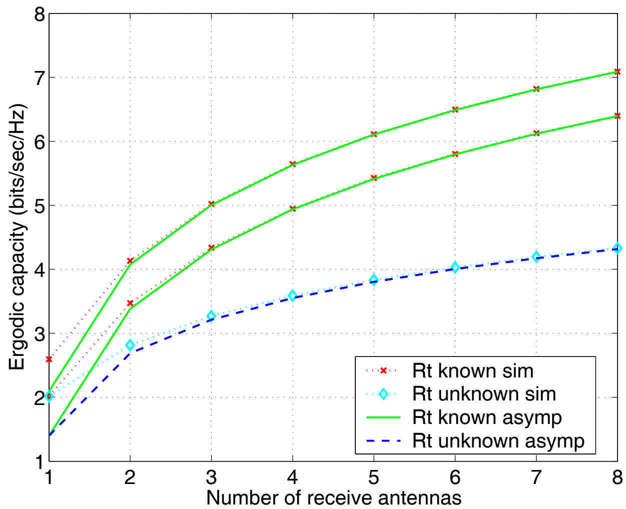


**Fig. 6. Capacity with rank one transmit correlation at SNR = 10dB. The transmit to receive antenna ratios are $[1, 2]$ from bottom to top pairs when $\mathbf{R}_t$ is known at the transmitter. The same antenna configurations are used when $\mathbf{R}_t$ is unknown; however, with rank one transmit correlation, more transmit antennas do not gain capacity without the correlation knowledge.**

plicitly as

$$\mathcal{C} = k\log(\gamma) + \log\det(\boldsymbol{\Xi}) + k\log(\frac{m}{M_t}) + \mathbf{E}_{\lambda_i}\sum_{i=1}^{k}\log\Big(\frac{\lambda_i}{m}\Big)\ . \tag{12}$$

Comparing (10) and (12), the gain in capacity obtained by knowing the correlation at the transmitter is $k\log(M_t/k)$. This gain is realized only when the correlation matrix is not full-rank, that is $k < M_t$. Thus with $k \le M_r$, we have

$$\mathcal{C}_{\text{gain, corr CSI}} = \max\Big\{k\log\frac{M_t}{k}, 0\Big\}\ . \tag{13}$$

Fig. 6 illustrates the capacity gain in knowing transmit correlation when the correlation matrix has rank one structure. The predicted gain is very accurate in this case, even at 10dB SNR. The gain increases with increasing number of transmit antennas, and is quite substantial at large number of antennas.

## 5. CONCLUSION

We have analyzed MIMO wireless channel capacity with different transmitter CSI information at high SNRs. Asymptotic formulae for evaluating the capacity are given, which are good approximations at finite number of antennas and sufficiently high SNRs. The results show that channel state information at the transmitter plays an important role in increasing the channel capacity.

## REFERENCES


[1] I.E.Telatar, *"Capacity of Multi-antenna Gaussian Channels"*, AT&T-Bell Laboratories Internal Technical Memorandum, June 1995.

[2] Marcenko and Pastur, *"Distribution of Eigenvalues for Some Sets of Random Matrices"*, Math.USSR-Sbornik, Vol.1, pp 457-483, 1967.

[3] T.M.Cover and J.A.Thomas, *"Elements of Information Theory"*, John Wiley & Sons, Inc., 1991.

[4] A. Narula, M. J. Lopez, M. D. Trott, G. W. Wornell, *"Efficient Use of Side Information in Multiple-Antenna Data Transmission over Fading Channels"*, IEEE J. Sel. Areas Comm., pp. 1423-1436, Vol. 16, No. 8, Oct. 1998.

[5] A.Gorokhov, *"Capacity of Multiple-Antenna Rayleigh Channel with a Limited Transmit Diversity"*, ISIT 2000, Sorrento, Italy, June 2000.

[6] E.Visotsky and U.Madhow, *"Space-Time Transmit Precoding with Imperfect Feedback"*, IEEE Trans. on Info. Theory, pp. 2632-2639, Vol. 47, No. 6, Sep. 2001.

[7] S. Jafar, A. Goldsmith, *"Transmitter Optimization and Optimality of Beamforming for Multiple Antenna Systems"*, to appear in Tran. Wireless Comm.

[8] J. Silverstein, *"Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices"*, J. Multivariate Analysis 55(2), pp. 331-339, 1995.

[9] Z.D. Bai, J. Silverstein, *"CLT of Linear Spectral Statistics of Large Dimensional Sample Covariance Matrices"*, accepted for pub., Annals of Probability.

[10] S.Boyd and L.Vandenberghe, *"Convex Optimization"*, Cambridge University Press 2003. Avail. at http://www.stanford.edu/ boyd/cvxbook.html.