Capacity Bounds of MIMO Channels with Asymmetric Channel State Information at Transmitter

An Liu, Xinming Huang, Youjian Liu, Haige Xiang, and Wu Luo

***Abstract*—Using two receive antennas at a mobile phone improves downlink performance. But because of the correlation of the two antennas, using one of them to transmit in the uplink is a good performance-complexity tradeoff. This creates an interesting scenario of asymmetric channel state information at the transmitter in a time division duplex system, where the downlink channel coefficients related to one of the two receive antennas can be learned by the base station from the uplink training due to reciprocity. We provide near optimal transmitter design by obtaining upper and lower bounds of the downlink ergodic capacity.**

***Index Terms*—Capacity, MIMO channels, asymmetric channel state information at transmitter.**

I. Introduction

THE capacity of wireless systems can be significantly improved by employing multiple uncorrelated transmit and receive antennas [1]. However, the antennas at the mobile phone are usually correlated due to its small size. For a mobile phone with two antennas, it is a good performance-complexity tradeoff in practice [2] to use both antennas to receive signal in the downlink and use one of the two antennas to transmit signal in the uplink. Correspondingly, in a time division duplex (TDD) system, one of the two rows of downlink channel state matrix can be learned by the base station from the uplink training due to reciprocity. This special partial channel state information at the transmitter (CSIT) is called asymmetric CSIT (ACSIT) in this letter.

Although the capacity with ACSIT can be solved by convex optimization numerically, it is difficult to obtain insight into the structure of the optimal transmission strategy. Using the concavity of the logarithm function, we obtain upper and lower bounds of the capacity. Maximizing the upper bound gives a near optimal transmission strategy in closed form. Assume the first row of the channel state matrix is known at the transmitter. The transmission strategy is to spend more power in the channel direction that matches the first row and spend less power in other channel directions. The power difference is a function of the receive antenna correlation. The power for other channel directions is evenly divided among them. We also discuss the extension to the case of more than two receive antennas, suggested by the structure of the strategy.

Manuscript received April 29, 2009. The associate editor coordinating the review of this letter and approving it for publication was G. Mazzini.

The work was supported in part by US-NSF Grant CCF-0728955 and ECCS-0725915, and in part by China 863 Program (No.2008AA7010205).

A. Liu, H. Xiang, and W. Luo (corresponding author) are with the State Key Laboratory of Advanced Optical Communication Systems & Networks, School of EECS, Peking University, China (e-mail: wendaolstr@gmail.com).

X. Huang is with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, USA.

Y. Liu is with the Department of Electrical, Computer, and Energy Engineering, University of Colorado at Boulder, USA.

Digital Object Identifier 10.1109/LCOMM.2009.090991

II. System Model

The conjugate transpose and determinant of a matrix $\mathbf{A}$ are denoted as $\mathbf{A}^\dagger$ and $|\mathbf{A}|$ respectively. The $(i,j)^{\text{th}}$ element of a matrix $\mathbf{A}$ is $\mathbf{A}_{i,j}$. The cardinality of a set $\mathcal{A}$ is $|\mathcal{A}|$. The expectation is denoted as $\mathrm{E}[\cdot]$ and $\|\mathbf{v}\|$ denotes the 2-norm of a vector $\mathbf{v}$.

Consider a TDD MIMO downlink with L_T transmit antennas at the base station (BS) and two receive antennas at the mobile phone. The received signal $\mathbf{Y} \in \mathbb{C}^{2\times 1}$ at the mobile phone is $\mathbf{Y} = \mathbf{HX} + \mathbf{W}$, where $\mathbf{X} \in \mathbb{C}^{L_T\times 1}$ is the transmit signal, whose average transmit power is upper bounded by ρ_0, $\mathbf{W} \in \mathbb{C}^{2\times 1}$ is the circularly symmetric complex Gaussian noise vector with i.i.d. entries of zero mean and unit variance. The channel state $\mathbf{H} \in \mathbb{C}^{2\times L_T}$ is i.i.d. across time[1] and has the following correlation structure $\mathbf{H} = \left(\mathbf{\Psi}^R\right)^{\frac{1}{2}} \mathbf{H}_w \left(\mathbf{\Psi}^T\right)^{\frac{1}{2}}$, where $\mathbf{H}_w \in \mathbb{C}^{2\times L_T}$ has zero-mean i.i.d. complex Gaussian entries with unit variance. The matrices $\mathbf{\Psi}^R \in \mathbb{C}^{2\times 2}$ and $\mathbf{\Psi}^T \in \mathbb{C}^{L_T\times L_T}$ describe the receive and transmit correlation respectively. This channel model is the well known Kronecker model [4]. The transmit antennas at the BS are assumed to be well separated. Thus we let $\mathbf{\Psi}^T = \mathbf{I}$ in this letter. In [5], exact expressions of the mutual information (MI) distributions under isotropic Gaussian input of the above MIMO channel is derived. In this paper, we derive the ergodic capacity and capacity bounds with ACSIT described as follows. Let $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2]^\dagger$, where $\mathbf{h}_i^\dagger \in \mathbb{C}^{1\times L_T}$ is the channel *vector* corresponding to the i^{th} receive antenna. We assume perfect CSI at the receiver for downlink. In the uplink, without loss of generality, only the first antenna of mobile phone is allowed to transmit signals. Thus, the BS is assumed to learn $\mathbf{h}_1$ perfectly for the downlink by channel reciprocity. The receive correlation matrix $\mathbf{\Psi}^R$ is also assumed known at the BS.

III. Near Optimal Transmission

We first give the ergodic capacity of MIMO channels with ACSIT. Then simple upper and lower bounds of the capacity are provided. The bounds lead to a near optimal transmission strategy.

A. The Capacity and its Upper and Lower Bounds

Proposition 1: The ergodic capacity of the MIMO channels with ACSIT is

[1]It is equivalent to a block fading channel as discussed in [3].

$$C_{\text{ACSIT}} = \max_{\boldsymbol{\Sigma}_{(.)}} \mathrm{E}_{\mathbf{h}_1}\left[\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\ln\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]\right] \quad (1)$$
$$\text{s.t. } \mathrm{E}_{\mathbf{h}_1}\left[\mathrm{Tr}\left(\boldsymbol{\Sigma}_{\mathbf{h}_1}\right)\right] \le \rho_0,$$

where the optimal input distribution is circularly symmetric complex Gaussian with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{h}_1} \in \mathbb{C}^{L_T\times L_T}$, which is a function of CSIT $\mathbf{h}_1$; and ρ_0 is the average power constraint.

Proof: The channel can be viewed as an equivalent channel with $\mathbf{Y}, \mathbf{h}_1$ and $\mathbf{h}_2$ as the output and $\mathbf{h}_1$ as the equivalent channel state [6]. Then the transmitter knows perfect CSIT about this equivalent channel. The corresponding capacity without input constraint is given in [6]. With input constraint, the capacity can be shown to be the minimum of the Lagrange dual function of the optimization problem (1) using the method in [7]. The remaining step is to show that (1) is a concave function of power constraint ρ_0 so that (1) is equal to the minimum of its Lagrange dual. But the concavity follows immediately from the concavity of $\log\det$ function on the domain of positive definite matrices [1]. ■

The capacity (1) can be calculated numerically. For each $\mathbf{h}_1$, generate a set $\mathcal{H}_{2|1}$ of $\left|\mathcal{H}_{2|1}\right|$ samples drawn from the distribution of $\mathbf{h}_2$ conditioned on $\mathbf{h}_1$. The average mutual information $\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\ln\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]$ can be approximated by

$$\frac{1}{\left|\mathcal{H}_{2|1}\right|}\sum_{\mathbf{h}_2\in\mathcal{H}_{2|1}} \ln\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|.$$

Then we can calculate $\boldsymbol{\Sigma}_{\mathbf{h}_1}$ using convex optimization algorithm similar to that used in [8]. However, the complexity is high and does not reveal insight to the solution structure. Therefore, we provide a simple near-optimal transmission scheme in closed form by solving an upper bound of the capacity.

Let $\boldsymbol{\Sigma}^*_{\mathbf{h}_1}$ be the optimal input covariance matrix, as a function of $\mathbf{h}_1$. We have

$$\begin{aligned} C_{\text{ACSIT}} &= \mathrm{E}_{\mathbf{h}_1}\left[\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\ln\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}^*_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]\right] && (2)\\ &\le \mathrm{E}_{\mathbf{h}_1}\left[\ln\left[\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}^*_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]\right] && (3)\\ &\le \max_{\boldsymbol{\Sigma}_{(.)}} \mathrm{E}_{\mathbf{h}_1}\left[\ln\left[\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]\right] && (4)\\ &\quad \text{s.t.} \mathrm{E}_{\mathbf{h}_1}\left[\mathrm{Tr}\left(\boldsymbol{\Sigma}_{\mathbf{h}_1}\right)\right] \le \rho_0, \end{aligned}$$

where inequality (3) follows from the concavity of the logarithm function. Let $\tilde{\boldsymbol{\Sigma}}_{\mathbf{h}_1}$ be the optimal input covariance for the modified optimization problem (4), the capacity is upper bounded by

$$C^{\text{up}}_{\text{ACSIT}} = \mathrm{E}_{\mathbf{h}_1}\left[\ln \mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\left|\mathbf{I}+\mathbf{H}\tilde{\boldsymbol{\Sigma}}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]\right]. \quad (5)$$

Substituting $\tilde{\boldsymbol{\Sigma}}_{\mathbf{h}_1}$ into (2) gives the lower bound

$$C^{\text{low}}_{\text{ACSIT}} = \mathrm{E}_{\mathbf{h}_1}\left[\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\ln\left|\mathbf{I}+\mathbf{H}\tilde{\boldsymbol{\Sigma}}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right]\right]. \quad (6)$$

B. Transmitter Design

The Lagrange dual function of the optimization (4) is

$$\max_{\boldsymbol{\Sigma}_{(.)}}\left[\ln \mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right] - \lambda\left(\mathrm{Tr}\left(\boldsymbol{\Sigma}_{\mathbf{h}_1}\right)-\rho_0\right)\right], \quad (7)$$

which can be solved for each fixed $\mathbf{h}_1$. Let

$$\tilde{\mathbf{U}} = [\tilde{\mathbf{u}}_1\cdots\tilde{\mathbf{u}}_{L_T}],\ \tilde{\mathbf{D}} = \mathrm{diag}(\tilde{d}_1,\cdots,\tilde{d}_{L_T}). \quad (8)$$

be eigenvectors and eigenvalues of $\tilde{\boldsymbol{\Sigma}}_{\mathbf{h}_1} = \tilde{\mathbf{U}}\tilde{\mathbf{D}}\tilde{\mathbf{U}}^\dagger$. Define

$$\begin{aligned} &\bar{\mathbf{h}}_1 = \mathbf{h}_1/\left\|\mathbf{h}_1\right\|^2, \nu_1 = \left\|\mathbf{h}_1\right\|^2,\\ &\phi_1 = \boldsymbol{\Psi}^R_{2,2} - \left|\boldsymbol{\Psi}^R_{1,2}\right|^2/\boldsymbol{\Psi}^R_{1,1}, \phi_2 = \boldsymbol{\Psi}^R_{1,2}/\boldsymbol{\Psi}^R_{1,1},\\ &\eta = \left(1+\left|\phi_2\right|^2\right)/\phi_1,\ b = \frac{1}{(\eta\nu_1+1)\phi_1}\\ &d^a(\lambda) = \left(\frac{\sqrt{\frac{1}{4}+\frac{\lambda^2}{\nu_1^2\phi_1}\left(\frac{1}{b}-\nu_1\right)}+\frac{1}{2}}{\lambda} - \frac{1}{\nu_1}\right)^+,\\ &d^b(\lambda) = \left(\frac{1}{\lambda}-b\right)^+, \end{aligned} \quad (9)$$

where $(x)^+ = \max(x,0)$. The transmitter design is as follow.

Proposition 2: The optimal solution to problem (4) is as follows. The first eigenvector is $\tilde{\mathbf{u}}_1 = \bar{\mathbf{h}}_1$ and other eigenvectors $\tilde{\mathbf{u}}_2\cdots,\tilde{\mathbf{u}}_{L_T}$ can be arbitrary as long as $\tilde{\mathbf{U}}$ is unitary; The optimal eigenvalues are

$$\begin{aligned} &\tilde{d}_1 = \begin{cases} d^a(\lambda), & d^b(\lambda) > \eta\\ d^b(\lambda), & d^b(\lambda) \le \eta \end{cases}\\ &\tilde{d}_i = \left(\tilde{d}_1-\eta\right)^+/\left(L_T-1\right),\ i = 2\cdots L_T, L_T > 1. \end{aligned} \quad (10)$$

The Lagrange multipliers λ satisfies the power constraint $\mathrm{E}_{\mathbf{h}_1}\left(\sum_{i=1}^{L_T}\tilde{d}_i\right) = \rho_0$.

A brief proof is given in the Appendix.

Remark 1: The solution has a waterfilling interpretation. The power is poured into $\tilde{d}_1$ first until $d^b(\lambda) = \eta = d^a(\lambda)$. Then $\tilde{d}_1$ grows according to $d^a(\lambda)$ with reduced power $\left(\tilde{d}_1-\eta\right)^+$ for the rest of the eigenvectors. This can be generalized to systems with more than two receive antennas by optimizing the power difference for the two groups of rows of the channel state matrix.

Remark 2: The solution of the problem (7) is not unique. The upper bound can be achieved by different choice of $d_i, i = 2\cdots L_T$ as long as they satisfy $\sum_{i=2}^{L_T} d_i = \left(d_1-\eta\right)^+, d_i \ge 0$. However, making $d_i, i = 2\cdots L_T$, equal gives the largest average rate because the problem is convex.

Remark 3: If we add the constraint of a single beam transmission ($d_i = 0, i = 2\cdots L_T$) to the problem (4), the optimal beamforming vector is $\bar{\mathbf{h}}_1$ and the optimal power allocation is obtained as in (9). This scheme is called optimal beamforming in this letter. Note that although we use the name optimal beamforming, it maximizes the upper bound rather than the capacity.

Remark 4: For the special case where $L_T = 1$, we only need to obtain the optimal power allocation, which takes the form of water-filling as in (9). For the zero probability case of $\nu_1 = 0$, eigenvectors $\tilde{\mathbf{U}}$ is arbitrary as long as it is unitary and the optimal eigenvalues are $\tilde{d}_i = d^b(\lambda)/L_T,\ i = 1\cdots L_T, L_T > 1$.

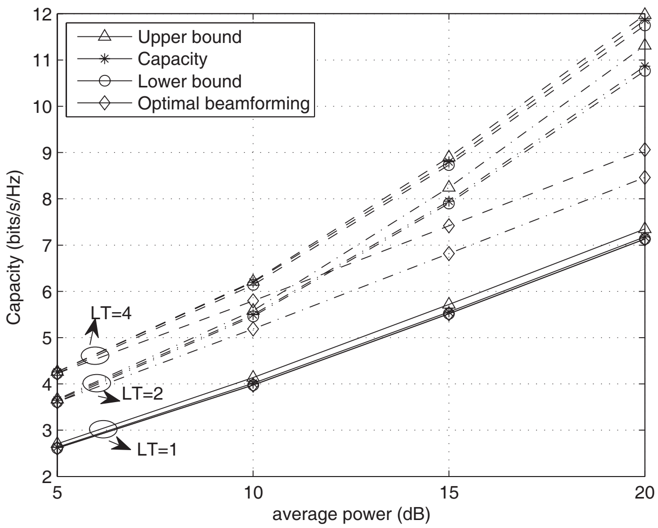

Fig. 1. Capacity bounds for MIMO channel with ACSIT.

IV. Simulation Results

We demonstrate that the transmission strategy is near optimal by simulation. The complex spatial cross-correlation between the receive antennas is given by the Kronecker model [9]

$$\rho_{1,2} = \int_0^{2\pi} \exp\left[j2\pi\frac{d_s}{\lambda_c}\sin(\theta_{AoA})\right]\mathrm{PAS}(\theta_{AoA})d\theta_{AoA},$$

where λ_c is the carrier wavelength; d_s is the distance between the two antennas at the mobile phone; θ_{AOA} is the angle of arrival at the mobile phone; and $\mathrm{PAS}(\theta_{AoA})$ is the power azimuth spectrum (PAS). We use the Gaussian shaped PAS because it enables the Kronecker model to best fit the 3GPP spatial channel model (SCM) [10]. We use the same parameters as in Fig. 2 of Section 4.1 in [10]. The normalized distance d_s/λ_c is set as 0.5. The resulting cross-correlation coefficient between the two receive antennas is $-0.5949+0.439j$. In Fig. 1, we compare the upper and lower bounds with the capacity for different number of transmit antennas and various average power constraint ρ_0, which is also the signal-to-noise-power-ratio (SNR). The fact that the lower bound is very close to the capacity demonstrates the transmission strategy is near optimal. The gap between the upper and lower bounds is small. When SNR is low, the rate of the optimal beamforming is close to the lower bound. However, the gap between them increases when SNR increases as expected.

V. Conclusion

The results provide fundamental limit and insight for the design of MIMO systems with ACSIT. Future work is to generalize the results to the case where the transmit antennas are also correlated and the mobile has more than two receive antennas.

Appendix

A Brief Proof of Proposition 2

Conditioned on $\mathbf{h}_1$, $\mathbf{h}_2$ is a Gaussian random vector with mean $\mathrm{E}(\mathbf{h}_2|\mathbf{h}_1) = \phi_2^*\mathbf{h}_1$ and covariance matrix $\phi_1\mathbf{I}$. Calculating the determinant and the expectation gives

$$\mathrm{E}_{\mathbf{h}_2|\mathbf{h}_1}\left[\left|\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{h}_1}\mathbf{H}^\dagger\right|\right] = \sum_{i=1}^{L_T}\left(\eta+\sum_{j\neq i}^{L_T} d_j\right)\phi_1 d_i\left\|\mathbf{h}_1^\dagger\mathbf{u}_i\right\|^2+\phi_1\sum_{i=1}^{L_T}d_i+1, \quad (11)$$

where d_i and $\mathbf{u}_i$ are defined in (8). For fixed d_i, $i = 1\cdots L_T$, let $i^* = \arg\max_i\left(\left(\eta+\sum_{j\neq i}^{L_T} d_j\right)\phi_1 d_i\right)$. Since $\sum_{i=1}^{L_T}\left\|\mathbf{h}_1^\dagger\mathbf{u}_i\right\|^2 = \|\mathbf{h}_1\|^2$, to maximize (11), we should let $\left\|\mathbf{h}_1^\dagger\mathbf{u}_{i^*}\right\|^2 = \|\mathbf{h}_1\|^2$ and $\left\|\mathbf{h}_1^\dagger\mathbf{u}_i\right\|^2 = 0$, $i \neq i^*$. Without loss of generality, let $i^*=1$. Thus the optimal first eigenvector is $\tilde{\mathbf{u}}_1 = \bar{\mathbf{h}}_1$. And other eigenvectors can be chosen arbitrarily without affecting the value of (11). Define the total power of other eigenvectors as $d^o \triangleq \sum_{i=2}^{L_T} d_i$. Problem (7) for each fixed $\mathbf{h}_1$ becomes

$$\max_{d_1,d^o}\left\{\ln\left[(\eta+d^o)\phi_1 d_1\|\mathbf{h}_1\|^2+\phi_1(d_1+d^o)+1\right] -\lambda(d_1+d^o-\rho_0)\right\}.$$

One may solve for the optimal power distribution between d_1 and d^o, and then solve for d_1+d^o by making the derivative zero. Then it can be verified that the solution in proposition (2) is one of the optimal solutions of problem (7).

References

[1] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommu.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[2] Long Term Evolution of the 3GPP radio technology, 3GPP, 2006. [Online]. Available: http://www.3gpp.org/Highlights/LTE/LTE.htm

[3] V. K. N. Lau, Y. Liu, and T.-A. Chen, "Capacity of memoryless channels and block fading channels with designable cardinality-constrained channel state feedback," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2038–2049, 2004.

[4] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[5] A. Maaref and S. Aissa, "Mutual information statistics for dual MIMO systems in correlated rayleigh fading," *IEEE Commun. Lett.*, vol. 10, no. 8, pp. 591–593, Aug. 2006.

[6] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2007–2019, 1999.

[7] Y. Liu, "A simple converse proof and a unified capacity formula for channels with input constraints," pp. 1–8, June 2008. [Online]. Available: http://arxiv.org/abs/0807.0042

[8] T. Kim, M. Bengtsson, E. Larsson, and M. Skoglund, "Combining long-term and low-rate short-term channel state information over correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2409–2414, July 2008.

[9] L. Schumacher, K. Pedersen, and P. Mogensen, "From antenna spacings to theoretical capacities—guidelines for simulating MIMO systems," in *Proc. PIMRC '02*, vol. 2, pp. 587–592, Sept. 2002.

[10] C.-X. Wang, X. Hong, H. Wu, and W. Xu, "Spatial-temporal correlation properties of the 3GPP spatial channel model and the Kronecker MIMO channel model," *EURASIP J. Wireless Commun. and Networking*, 2007.