Beamforming Capacity and SNR Maximization for Multiple Antenna Systems

Syed Ali Jafar and Andrea Goldsmith
Stanford University, USA

Abstract

We explore the optimal input distributions to maximize average received SNR and to achieve beamforming capacity on the downlink for two different channel models. First we model a scenario where the mobile unit is surrounded by local scatterers while the base station (BS) is relatively unobstructed, so that the channel matrix consists of i.i.d. rows and correlated columns. For this model we show that the average SNR maximizing input distribution also achieves the beamforming capacity. Then we model a scenario where the transmit antennas at the BS are placed significantly far apart, so that the fades are independent but not identically distributed. For this model we obtain necessary and sufficient conditions under which the average SNR maximizing solution also achieves beamforming capacity. We find that for typical mobile unit antenna spacings, increasing the BS antenna spacing not only improves the capacity but achieves this higher capacity using just one transmit antenna at a time, thus saving significantly on the cost of power amplifiers at the BS.

1 Introduction

One of the most promising approaches towards improving the data rates achievable in wireless systems is to use multiple transmit and receive antennas. By utilizing space as an additional dimension, these multiple input multiple output (MIMO) systems easily surpass the data rates achievable by conventional systems utilizing time and frequency alone. An interesting aspect of these systems is the tradeoff between the feedback quality and coding complexity, explored in [2]. It is shown in [2] that a good feedback quality allows us to achieve a higher Shannon capacity with a lower decoding complexity. However for a MIMO channel a sufficiently accurate channel state feedback is hard to implement due to the multiple channel coefficients that need to be estimated and the limited processing power at the mobile unit. The quality of feedback gets worse when the channel exhibits fast fading. A more practical scenario, assumed in this work, is a MIMO system where the transmitter knows only the statistics of the channel. The receiver is still assumed to have perfect channel state information. This assumption is reasonable since the receiver can wait until it has enough received symbols to form a good channel estimate.

From the Shannon capacity point of view, in the absence of Channel State Information at Transmitter (CSIT), vector codes are needed to achieve the highest data rates possible. Alternatively, in theory, with a BLAST-like approach [1], one could transmit several scalar codes in parallel to achieve the same capacity. In practice however both these approaches have severe drawbacks in terms of the decoding complexity required at the mobile unit. Also, in contrast to the well established scalar codec technology, good vector codes have not been found. In the light of these observations, beamforming becomes an interesting alternative. Beamforming allows scalar code transmission with multiple antennas. The numerical results in [4] show that beamforming is usually close to the capacity achieving strategy. Moreover, if the channel covariance matrix exhibits disparate modes then beamforming is the capacity achieving strategy [2][3].

We focus on optimizing the input to achieve either Shannon capacity C_{bf} or the maximum average SNR using beamforming, i.e. using input covariance matrices with unit rank. Achieving C_{bf} and maximizing average SNR represent different objectives and therefore in general lead to different optimal input distributions. However we find that for several common fading models the two objectives lead to the same optimal input distribution. Typically, it is harder to optimize input to achieve capacity as the capacity expression tends to be less tractable, and therefore finding the input distribution to maximize average SNR provides a simpler alternative problem. Next, we further analyze the relationship between these two objectives.

2 System Model and Problem Definition

A MIMO system using n_T transmit and n_R receive antennas (an (n_T, n_R) system) is characterized by $\mathbf{y} = H\mathbf{x} + \mathbf{n}$, where $\mathbf{y}$ is the n_R dimensional output vector, $\mathbf{x}$ is the n_T dimensional input vector, $\mathbf{n}$ is the n_R dimensional white

Gaussian noise vector, and $\mathbf{H}$ describes the channel matrix. The elements of H are modeled as zero mean complex Gaussian random variables (Rayleigh fading).

The Shannon capacity of this system for a transmit power P can be expressed as

$$C = \max_{Q\,:\,\text{trace}(Q)=P} \mathrm{E}\left[\log\left|I_{n_R} + \frac{HQH^\dagger}{\sigma^2}\right|\right] \qquad (1)$$

Beamforming, by definition, implies a unit rank input covariance matrix $Q = \boldsymbol{\nu}\boldsymbol{\nu}^\dagger$ where $\boldsymbol{\nu}$ is an n_T dimensional vector. Using the proper ty $|I + AB| = |I + BA|$ we express the beamforming capacity as

$$C_{bf} \;=\; \max_{\nu\,:\,\boldsymbol{\nu}^\dagger\boldsymbol{\nu}=P} \mathrm{E}\left[\log\left(1 + \frac{\boldsymbol{\nu}^\dagger H^\dagger H\boldsymbol{\nu}}{\sigma^2}\right)\right] \qquad (2)$$

Note that $\frac{\boldsymbol{\nu}^\dagger H^\dagger H\boldsymbol{\nu}}{\sigma^2}$ is the received SNR. Thus the problems we tackle are:

1. **Beamforming Capacity** Determine the optimum ν to achieve the beamforming capacity C_{bf}.

2. **Average SNR** Determine the optimum ν to maximize the average SNR $\mathrm{E}[\frac{\boldsymbol{\nu}^\dagger H^\dagger H\boldsymbol{\nu}}{\sigma^2}]$.

We consider two different models for the covariances of the elements of the channel matrix $\mathbf{H}$.

• *I.i.d. rows and correlated columns:* In the first model the rows of $\mathbf{H}$ are assumed to be i.i.d. while the columns are correlated. This is a common assumption [3] for the downlink where the scatterers surrounding the mobile unit decorrelate the fades associated with different receive antennas, while the base station is relatively unobstructed so that the fades associated with different transmit antennas are correlated. Mathematically, the distribution of the i^{th} row of $\mathbf{H}$ is given by $H_{i\cdot} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ for $1 \le i \le n_R$, and the covariance matrix Σ has an eigendecomposition $\Sigma = U_\Sigma \Lambda_\Sigma U_\Sigma^\dagger$.

• *Independent but not identically distributed fades:* To achieve a richer scattering environment, i.e. to achieve independent fades we need significant spacing (of the order of meters for Gigahertz carrier frequency) between the transmit antennas [3]. While the base station does not have very stringent size constraints and can place the transmit antennas far enough apart to achieve independent fades, the path between each transmit antenna and the mobile unit is now sufficiently distinct so that the fades can not be assumed to be identically distributed. This forms the basis for our second model, where we assume that the elements of $\mathbf{H}$ are independent but not necessarily identically distributed. So the second model is more general in that we no longer constrain the rows to be identically distributed. However it is more constrained than the previous model in that all elements are assumed to be independent of each other. We model the elements of H as independent zero mean complex Gaussian random variables with variances $\mathrm{E}[H_{ij}H_{ij}^*] = \gamma_j^i$. So the i^{th} row of H has a covariance matrix $\Gamma^i = \mathrm{E}[\sum_{i=1}^{n_R} H_{i\cdot}^\dagger H_{i\cdot}]=$ $\mathrm{diag}\{\gamma_1^i, \gamma_2^i, \cdots, \gamma_{n_T}^i\}$. Without loss of generality we assume the rows are numbered so that $\gamma_1^1 \ge \gamma_1^2 \ge \cdots \ge \gamma_1^{n_R}$.

3 Solution

3.1 Channel matrix with i.i.d rows and correlated columns

Let us define the vector

$$v \triangleq \frac{H\nu}{\sqrt{\nu^\dagger \Sigma \nu}}. \qquad (3)$$

The i^{th} component of v is given by $v_i = \frac{H_{i\cdot}\nu}{\sqrt{\nu^\dagger\Sigma\nu}}$. Since the rows of H are i.i.d., so are the random variables v_i. Furthermore $\mathrm{E}[v_i] = 0$ and $E[v_i v_i^*] = \frac{\mathrm{E}[H_{i\cdot}\nu\nu^\dagger H_{i\cdot}^\dagger]}{\nu^\dagger\Sigma\nu} = \frac{\mathrm{E}[\nu^\dagger H_{i\cdot}^\dagger H_{i\cdot}\nu]}{\nu^\dagger\Sigma\nu} = 1$. Thus $v \sim \tilde{N}(\mathbf{0}, I)$ and the capacity can be expressed as

$$C = \max_{\nu\,:\,||\nu||=\sqrt{P}} \mathrm{E}\left[\log\left(1 + \nu^\dagger\Sigma\nu\frac{v^\dagger v}{\sigma^2}\right)\right] \qquad (4)$$

Since v does not depend on ν, the capacity is maximized by maximizing $\nu^\dagger\Sigma\nu$ subject to $||\nu|| = \sqrt{P}$. But this is the standard quadratic form which is maximized by choosing ν as the dominant eigenvector of Σ scaled to meet the power constraint. However note that $\nu^\dagger\Sigma\nu$ is also the average SNR with maximum ratio combining at the receiver. Hence we conclude that in this case the capacity achieving solution also maximizes the average received SNR.

3.2 Channel Matrix with Independent Elements

3.2.1 Maximizing Average SNR

First we maximize the average received SNR given by

$$\begin{aligned}\mathrm{E[SNR]} &= \mathrm{E}[\nu^\dagger H^\dagger H\nu] = \nu^\dagger E[\sum_{i=1}^{n_R} H_{i\cdot}^\dagger H_{i\cdot}]\nu \\ &= \nu^\dagger[\sum_{i=1}^{n_R}\Gamma^i]\nu = \sum_{j=1}^{n_T}\Gamma_j||\nu_j||^2, \qquad (5)\end{aligned}$$

where $\Gamma_j = \sum_{i=1}^{n_R}\gamma_j^i$. But (5) is maximized by choosing $\nu_i = \sqrt{P}$ for $i = i^*$ and 0 otherwise, where $i^* = \arg\max_j \Gamma_j$. Thus we make our first observation - **average SNR is maximized by using one transmit antenna alone**. This transmit antenna is the one corresponding to the maximum sum of channel gain powers from a transmit antenna to each of the receive antennas. To save space we write $\mathbf{A}_i$ **is S-optimal** when we mean that "average SNR is maximized by using transmit antenna i alone".

3.2.2 Beamforming Capacity

The capacity with beamforming $Q = \nu\nu^\dagger$ can be expressed as

$$\begin{aligned} C_{bf} &= \max_{\nu\,:\,||\nu||=\sqrt{P}} \mathrm{E}\left[\log\left(1+\frac{\nu^\dagger H^\dagger H\nu}{\sigma^2}\right)\right] \\ &= \max_{\nu\,:\,||\nu||=\sqrt{P}} \mathrm{E}\left[\log\left(1+\frac{\nu^\dagger\left(\sum_{i=1}^{n_R}\Gamma^i|v_i|^2\right)\nu}{\sigma^2}\right)\right]. \end{aligned}$$

where v is an n_R dimensional vector with components $v_i \triangleq \frac{H_{i.}\nu}{\sqrt{\nu^\dagger\Gamma^i\nu}}$ so that v_i are independent complex Gaussian random variables distributed as $v_i \sim \tilde{N}(0,1)$. To simplify notation we make the following definitions

$$w_i \triangleq |v_i|^2 \quad \text{and} \quad f_j \triangleq \sum_{i=1}^{n_R}\gamma_j^i w_i, \tag{6}$$

so that w_i are i.i.d. exponential random variables with unit mean. Now since Γ^i are diagonal matrices we can rewrite this as

$$C_{bf} = \max_{\nu\,:\,||\nu||=\sqrt{P}} \mathrm{E}\left[\log\left(1+\frac{\sum_{j=1}^{n_T}|\nu_j|^2 f_j}{\sigma^2}\right)\right]. \tag{7}$$

Since SNR is maximized by using only one transmit antenna, we wish to determine the conditions under which capacity is also achieved by using only one transmit antenna. Similar to the notation introduced earlier we express the statement "capacity is maximized by using transmit antenna i alone" as **$\mathbf{A}_i$ is C-optimal**. Note that the power allocated to transmit antenna i is given by $|\nu_i|^2$. Without loss of generality let us allocate a power $P-p$ to antenna 1, and divide the power p among the remaining antennas such that antenna j gets power $\alpha_j p$ where $\sum_{i=2}^{n_T}\alpha_i = 1$ and $\alpha_j \geq 0$ for $1 < j \leq n_T$. So the capacity with this power allocation becomes

$$C_{bf} = \mathrm{E}\left[\log\left(1+(P-p)\frac{f_1}{\sigma^2}+p\frac{\sum_{j=2}^{n_T}\alpha_j f_j}{\sigma^2}\right)\right]. \tag{8}$$

Now if A1 is C-optimal we must have (necessary condition)

$$\begin{aligned} &\left.\frac{\partial C_{bf}}{\partial p}\right|_{p=0} \leq 0 \\ \Rightarrow\quad &\mathrm{E}\left[\frac{\sum_{j=2}^{n_T}\alpha_j f_j - f_1}{\sigma^2+f_1}\right] \leq 0 \qquad (9) \\ \Rightarrow\quad &\sum_{i=1}^{n_R}\gamma_1^i x_i^1 \geq \sum_{j=2}^{n_T}\alpha_j\left(\sum_{i=1}^{n_R}\gamma_j^i x_i^1\right), \qquad (10) \end{aligned}$$

where

$$x_l^m \triangleq \mathrm{E}\left[\frac{w_l}{\sigma^2+\sum_{i=1}^{n_R}\gamma_m^i w_i}\right]. \tag{11}$$

(10) has to be true for all non-negative choices of α_i that sum up to unity. So the necessary condition becomes

$$\sum_{i=1}^{n_R}\gamma_1^i x_i^1 \geq \sum_{i=1}^{n_R}\gamma_j^i x_i^1, \;\; 1 \leq j \leq n_T \tag{12}$$

Further, taking the second derivative of C_{bf} it is easily seen that $\frac{\partial^2 C_{bf}}{\partial p^2} \leq 0$. Thus (12) is both necessary and sufficient for A1 to be C-optimal.

Using this necessary and sufficient condition one can easily come up with numerical examples that lead to the following observation - **Beamforming capacity cannot always be achieved by using just one transmit antenna alone.** As an example consider a (2, 2) channel with the variances given by $\gamma_1^1 = 10, \gamma_2^1 = 15, \gamma_1^2 = 4$ and $\gamma_2^2 = 1$. Numerically evaluating the x_i^j and plugging into (12) we find that neither A1 nor A2 is C-optimal. Similarly consider another (2,2) channel, this time with variances given by $\gamma_1^1 = 12, \gamma_2^1 = 16, \gamma_1^4 = 4$ and $\gamma_2^2 = 1$. Obviously for this channel A2 is S-optimal. However numerically testing the necessary and sufficient condition we find that A1 is C-optimal. Thus we note that - **It is possible that the S-optimal transmit antenna is not used at all in the beamforming capacity maximizing solution**.

Next we want to obtain simpler conditions under which the S-optimal antenna is also C-optimal. Without loss of generality suppose A1 is S-optimal, i.e. (13) holds. Recall that the average SNR is maximized by using transmit antenna 1 alone if and only if

$$\Gamma_1 = \sum_{i=1}^{n_R}\gamma_1^i = \max_{1\leq j\leq n_T}\Gamma_j = \max_{1\leq j\leq n_T}\sum_{i=1}^{n_R}\gamma_j^i. \tag{13}$$

Now notice that if the gains from transmit antenna 1 to all receive antennas are identically distributed, i.e. if the variances $\gamma_1^i = \gamma_1$ for $1 \leq i \leq n_R$ then from the definition of x_i^1 and since w_i are i.i.d. it follows that all x_i^1 are also equal, i.e. $x_1^1 = x_2^1 = \cdots = x_{n_R}^1 = x$. Substituting into (12) and using (13) we see that the necessary and sufficient condtion is always satisfied. Hence we make our next observation - **If A1 is S-optimal and the channel gains from transmit antenna 1 to all receive antennas are i.i.d. then A1 is also C-optimal**.

In order to gain further insight into the relationship between the C-optimal and S-optimal solutions we need the following propositions. The proofs for propositions 1 and 2 are given in the appendix and proposition 3 is proved in [3].

Proposition 1 *For all* $i, j \in \{1, 2, \cdots, n_T\}$, *if* $\gamma_1^i \leq \gamma_1^j$ *then* $x_i^1 \geq x_j^1$.

Proposition 2 *For*

$$g(\Delta) \triangleq \Bigg[\sigma^2 + (\frac{\gamma_1^i + \gamma_1^j}{2} + \Delta)w_i + (\frac{\gamma_1^i + \gamma_1^j}{2} - \Delta)w_j + \sum_{l=1, l\neq i,j}^{n_R} \gamma_1^l w_l\Bigg]^{-1} \quad (14)$$

and $1 \leq i, j \leq n_R$, i.e. when the coefficient of w_i is greater than or equal to that of w_j,

$$G(\Delta) \triangleq E[w_i g(\Delta)] \quad (15)$$

is convex and attains its minimum at $\Delta = 0$. Thus as the coefficients of w_i in the denominator become more spread out, or less equal, the value of $G(\Delta)$ increases.

Proposition 3 *For any two given positive real vectors $\alpha = [\alpha_i], \beta = [\beta_i] \in \mathbb{R}_+^n$ the permutation π^* that minimizes the sum $\sum_{i=1}^n \frac{\alpha_{\pi(i)}}{\beta_i}$ is such that $\alpha_{\pi^*(i)}$ and β_i are in the same order. That is, $\forall i, j \in \{1, 2, \cdots, n\}$, if $\alpha_{\pi^*(i)} < \alpha_{\pi^*(j)}$, then $\beta_i < \beta_j$.*

Since we assumed γ_1^i are arranged in decreasing order, from proposition 1 it follows that we must have $x_1^1 \leq x_2^1 \leq \cdots \leq x_{n_R}^1$. From proposition 3, of all the permutations π of $[\gamma_j^i]_i$ the one that minimizes the sum $\sum_{i=1}^{n_R} \gamma_j^{\pi(i)} x_i^1$ is the one for which $\gamma_j^{\pi(i)}$ are arranged in decreasing order. Note that in general $[\gamma_j^i]_i$ can be in any order for $1 \leq i \leq n_R$. However γ_1^i are arranged such that the LHS of (12) is minimized. Since the LHS has to be greater than or equal to the RHS for A_1 to be C-optimal and the RHS depends on the way $[\gamma_j^i]_i$ are arranged, we note that the C-optimality of A_1 is a function of the ordering of $[\gamma_j^i]_i$. However if we impose the additional constraint that all $[\gamma_j^i]_i$ are in decreasing order, i.e. if we assume $\Gamma^1 \geq \Gamma^2 \geq \cdots \geq \Gamma^{n_R}$ and further that $[\gamma_1^i]$ majorizes $[\gamma_j^i]_i = \{\gamma_j^1, \gamma_j^2, \cdots, \gamma_j^{n_R}\}$ then it follows from proposition 3 that the condition for C-optimality of A_1 is satisfied. Thus we note that **if the covariance matrices of the rows of H can be arranged in decreasing order and, for some l, $[\gamma_l^i]_i$ weakly majorizes $[\gamma_j^i]_i$ for all j such that $1 \leq j \leq n_T$ then A_l is both S-optimal as well as C-optimal.** Note that although proposition 3 was proved for strong majorization the same proof holds for weak majorization too. Recall that a real vector $\alpha = [\alpha_i] \in \mathbb{R}^n$ majorizes another real vector $\beta = [\beta_i] \in \mathbb{R}^n$ if and only if the sum of the k smallest entries of α is greater than or equal to the sum of the k smallest entries of β for $k = 1, 2, \cdots, n-1$. The majorization is strong if $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i$ and weak if $\sum_{i=1}^n \alpha_i > \sum_{i=1}^n \beta_i$.

The necessary and sufficient condition for C-optimality of A_1 given in (12) requires a numerical computation to evaluate the x_1^i. A closed form expression for the inequality is hard to obtain. However we can obtain a sufficient condition in closed form by bounding techniques. Starting with (12) and using the properties $\gamma_1^1 \geq \gamma_1^2 \geq \cdots \geq \gamma_1^{n_R}$ and $x_1^1 \leq x_2^1 \leq \cdots \leq x_{n_R}^1$ we bound

$$\text{LHS of (12)} \geq x_1^1 \sum_{i=1}^{n_R} \gamma_1^i = x_1^1 \Gamma_1, \quad (16)$$

$$\text{and RHS of (12)} \leq x_{n_R}^1 \sum_{i=1}^{n_R} \gamma_j^i = x_{n_R}^1 \Gamma_j, \quad (17)$$

to obtain the following sufficient condition:

$$A_1 \text{ is C-optimal if } \frac{x_1^1}{x_{n_R}^1} \geq \frac{\Gamma_j}{\Gamma_1}, \quad 1 \leq j \leq n_T. \quad (18)$$

But we have

$$\begin{aligned} x_1^1 &= \mathrm{E}\left[\frac{w_1}{\sigma^2 + \sum_{j=1}^{n_R} \gamma_1^j w_j}\right] \\ &\geq \mathrm{E}\left[\frac{w_1}{\sigma^2 + \sum_{j=1}^{n_R} \gamma_1^j w_1}\right] \quad (19) \\ &= \frac{1}{\Gamma_1} - \sigma^2 \frac{e^{\frac{\sigma^2}{\Gamma_1}} \Gamma\left(0, \frac{\sigma^2}{\Gamma_1}\right)}{(\Gamma_1)^2} \quad (20) \end{aligned}$$

where $\Gamma(.,.)$ is the incomplete Gamma function. (19) follows from proposition 1 and (20) is obtained by straightforward integration. Also we have

$$\begin{aligned} x_{n_R}^1 &= \mathrm{E}\left[\frac{w_{n_R}}{\sigma^2 + \sum_{j=1}^{n_R} \gamma_1^j w_j}\right] \\ &\leq \mathrm{E}\left[\frac{w_{n_R}}{\sigma^2 + (\gamma_1^1 + \gamma_1^{n_R})w_1 + \sum_{j=2}^{n_R-1} \gamma_1^j w_j}\right] \quad (21) \\ &= \mathrm{E}\left[\frac{1}{\sigma^2 + (\gamma_1^1 + \gamma_1^{n_R})w_1 + \sum_{j=2}^{n_R-1} \gamma_1^j w_j}\right] \quad (22) \\ &\leq \mathrm{E}\left[\frac{1}{\sigma^2 + w_1 \sum_{j=1}^{n_R} \gamma_1^j}\right] \quad (23) \\ &= \frac{e^{\frac{\sigma^2}{\Gamma_1}} \Gamma\left(0, \frac{\sigma^2}{\Gamma_1}\right)}{\Gamma_1}, \quad (24) \end{aligned}$$

where (21) follows from proposition 1 and (22) makes use of the fact that w_{n_R} is independent of $w_1, \cdots, w_{n_R-1}$ and has unit mean. Proposition 2 implies (23) and (24) follows from straightforward integration.

Combining (24), (20), and (18) we obtain the following sufficient condition

$$A_1 \text{ is C-optimal if } \frac{\Gamma_j}{\Gamma_1} \leq \frac{e^{-\frac{\sigma^2}{\Gamma_1}}}{\Gamma\left(0, \frac{\sigma^2}{\Gamma_1}\right)} - \frac{\sigma^2}{\Gamma_1} \quad 1 \leq j \leq n_T.$$

4 Conclusions

We obtain the SNR maximizing input and the beamforming capacity achieving input for a channel with correlated columns and i.i.d. rows. We find that the SNR maximizing input also achieves the beamforming capacity for this channel. However, for a channel model where the elements of the channel matrix are independent but not identically distributed, we find that while the average SNR is maximized by transmitting on one antenna alone, beamforming capacity may require more than one transmit antenna. Moreover, the transmit antennas used to achieve beamforming capacity may in some cases not even include the average SNR maximizing transmit antenna. Thus the problems are quite distinct and the solutions in general reflect that. However, we also noticed that when the fades from a transmit antenna to the receive antennas are identically distributed, the SNR maximizing input also achieves beamforming capacity. Since the identical fading distribution property is true for typical antenna spacings at the mobile unit, we conclude that for practical multiple antenna systems with only the channel statistics available to the transmitter, the average SNR maximizing input also achieves the beamforming capacity.

As explained in [5], the cost of a transmitter is dominated by that of power amplifiers. Antennas on the other hand are cheaper by typically two orders of magnitude. So it is economically advantageous to use a small number of amplifiers and a larger number of antennas and connect the amplifiers to a selected set of antennas to achieve maximum possible system capacity. This is particularly interesting in the light of our result that by placing the transmit antennas at the base station far enough apart to achieve independent fades we can achieve the beamforming capacity using just one transmit antenna. So we conclude that increasing the antenna spacing at the transmitter not only improves the capacity of the channel by providing a richer fading environment but also saves on the cost of power amplifiers by achieving this higher capacity with just one transmit antenna.

5 Appendix

5.1 Proof of Proposition 1

Consider the functions

$$x_i^1(\Delta) = \mathrm{E}\left[w_i g(\Delta)\right] \text{ and } x_j^1(\Delta) = \mathrm{E}\left[w_j g(\Delta)\right]. \qquad (25)$$

For $\Delta = \frac{\gamma_1^j - \gamma_1^i}{2}$ we have $x_i^1(\Delta) = x_i^1$ and $x_j^1(\Delta) = x_j^1$ as defined earlier. Notice that $x_i^1(0) = x_j^1(0)$. We wish to prove that as Δ increases from 0, $x_i^1(\Delta)$ increases monotonically and by the same token $x_j^1(\Delta)$ decreases monotonically so that $x_i^1 = x_i^1(\frac{\gamma_1^j - \gamma_1^i}{2}) \geq x_j^1(\frac{\gamma_1^j - \gamma_1^i}{2}) = x_j^1$. Differentiating (25) with respect to Δ we obtain

$$\left.\frac{\partial x_i^1(\Delta)}{\partial \Delta}\right|_{\Delta=0} = \mathrm{E}\left[w_i(w_i - w_j)g(\Delta)^2\right] \qquad (26)$$

$$= \mathrm{E}\left[X_i^2 - X_i X_j\right] \qquad (27)$$

where $X_i = w_i g(\Delta)$ and $X_j = w_j g(\Delta)$. From symmetry it follows that X_i and X_j are identically distributed random variables. Therefore it follows from the Cauchy-Schwartz inequality that $\mathrm{E}[X_i^2] \geq \mathrm{E}[X_i X_j]$. This proves that $\left.\frac{\partial x_i^1(\Delta)}{\partial \Delta}\right|_{\Delta=0} \geq 0$. It is easily verified that the second derivative $\frac{\partial^2 x_i^1(\Delta)}{\partial \Delta^2} \geq 0$ and the result in proposition 1 follows.

5.2 Proof of Proposition 2

The proof is very similar to the proof of proposition 1. It is easily seen that $\left.\frac{\partial G(\Delta)}{\partial \Delta}\right|_{\Delta=0} = 0$ and $\frac{\partial^2 G(\Delta)}{\partial \Delta^2} \geq 0$. That is all we need to establish the statement of proposition 2.

References

[1] P.W. Wolniansky, G. J. Foschini, G. D. Golden, R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel", ISSSE 98, International Symposium on URSI, 1998, Pages: 295-300.

[2] A. J. Syed, A. J. Goldsmith, "On Optimality of Beamforming for Multiple Antenna Systems with Imperfect Feedback", Submitted.

[3] A. J. Syed, S. Vishwanath, A. J. Goldsmith, "Channel Capacity and Beamforming for Multiple Transmit and Receive Antennas with Covariance Feedback", To appear in Proceedings of ICC 2001, Helsinki, Finland.

[4] E. Visotsky, U. Madhow, "Space-Time Transmit Precoding with Imperfect Feedback", Proceedings of ISIT 2000.

[5] R. Nabar, D. Gore, A. Paulraj, "Selection and Use of Optimal Transmit Antennas in Wireless Systems", ICT'2000, Acapulco, Mexico, May 2000.