DIMACS Series in Discrete Mathematics
and Theoretical Computer Science

Optimality of Beamforming in Multiple Transmitter Multiple Receiver Communication Systems with Partial Channel Knowledge

Steven H. Simon and Aris L. Moustakas

Abstract. In this paper we consider a narrowband point-to-point MIMO communication system with n_{T} transmit antennas and n_{R} receive antennas, and we assume that the receiver has perfect knowledge of the channel. We consider two canonical cases of partial channel feedback: (a) where the mean channel is fed back to the transmitter and the channel covariance is assumed to be white with known magnitude, and (b) the covariance (assumed factorizable) of the channel is fed back to transmitter, and we assume the mean channel is zero. In both cases we are able to derive an exact analytic condition (both necessary and sufficient) that determines when beamforming is optimal for maximizing the average throughput.

1. Introduction

The large capacities available to communication systems employing multiple antennas has been the subject of much research in recent years. Interesting questions arise in cases where the receiver has perfect knowledge of the channel (as might be achieved by pilots in practice) but the transmitter has only partial information about the channel (as would have to be gained from some sort of feedback). It has been seen in many cases that exploiting such partial channel information at the transmitter can increase link capacities substantially. In this paper we will focus on two canonical cases of channel feedback : (a) mean feedback, where the mean of the channel is known to the transmitter and the covariance of the channel is assumed to be white with known magnitude, and (b) covariance feedback, where the full (factorizable) covariance of the channel is known and the average channel is assumed to be zero. These two cases appear to be realistic and have been considered before in a number of publications[**1, 2, 3**]. In this paper, we focus on the question of whether beamforming transmission strategies are optimal for maximizing the ergodic (average) channel capacity given the partial channel knowledge. (Some very limited results have been previously obtained on the harder question

2000 *Mathematics Subject Classification.* Primary 94A15, 94A17; Secondary 94A05.

Key words and phrases. Transmit Diversity; Scheduling; Multiple Antennas, Side Information.

©2003 American Mathematical Society

of whether beamforming transmission is optimal for maximizing some particular outage capacity[**4, 5**]).

The question of beamforming optimality has been addressed in a number of publications. Much of this work has been concerned with the simple case where there are $n_{\rm T}$ transmit antennas and a single receive antenna[**6, 7, 8, 4**]. More recently certain limited results have been obtained for the more general case with multiple transmit and multiple receive antennas. Jafar et al[**3**] considered the case of covariance feedback in the special case where the channel is assumed to have i.i.d. rows and correlated columns, and managed to derive upper and lower bounds on the optimality condition for beamforming. Later, the current authors[**5**] derived the exact optimality condition in this special case. In the present paper we shall obtain a result that will apply to a general channel covariance which can be written as a direct product of arbitrarily correlated columns (transmitting antenna elements) and correlated rows (receiving antenna elements). In addition, we will also obtain an exact beamforming optimality condition for the case of mean channel feedback where the covariance of the channel is assumed to be white.

2. Definitions

In the narrowband MIMO problem - $n_{\rm T}$ transmitters and $n_{\rm R}$ receivers - the channel is defined by

$$y_\alpha = \sum_{i=1}^{n} G^*_{i\alpha} x_i + \eta_\alpha \tag{1}$$

where y_α is α^{th} component of the vector of $n_{\rm R}$ complex received signals, x_i is the i^{th} component of the $n_{\rm R}$ dimensional complex vector of transmitted signal, $G^*_{i\alpha}$ is the complex channel from the i^{th} transmitter to the α^{th} receiver, and η_α is the complex noise at the α^{th} receiver. We will try to keep the notation that italic roman indices $i, j, \ldots$ represent the transmitter antennas and take values $1 \ldots n_{\rm T}$ whereas Greek indices $\alpha, \beta, \ldots$ represent the receiver end and go from $1, \ldots, n_{\rm R}$. In this paper, boldface quantities denote matrices or vectors (hopefully usage will make clear which one is which) so we can also write $\mathbf{y} = \mathbf{G}^\dagger \mathbf{x} + \boldsymbol{\eta}$ (with the $\dagger$ indicating transpose-conjugate). We will assume that the noise is an i.i.d. Gaussian random vector, $E\{\eta^*_\alpha \eta_\beta\} = \nu \delta_{\alpha\beta}$ where $E\{\cdot\}$ denotes the expectation value, ν is the noise power per receiver, and δ is the Kronecker delta function. (Generalizing the calculations below to a nontrivial noise correlation matrix (in place of the δ) is trivial in the covariance feedback problem, but is nontrivial for mean feedback). For simplicity, we assume in this paper that the noise power per receiver is of unit amplitude $\nu = 1$ (all other quantities need to be measured in units of the noise power).

The transmitted signal covariance matrix is given by $E\{x^*_i x_j\} = Q_{ij}$ with Q_{ij} a non-negative definite Hermitian matrix, with eigenvalues q_i, for $i = 1, \ldots, n_{\rm T}$. The constraint of fixed total transmitted power P is expressed in terms of $\mathbf{Q}$ as $\mathrm{Tr}\{\mathbf{Q}\} = \sum_{i=1}^{n_{\rm T}} q_i = P$.

For a given instantiation of $\mathbf{G}$, if the receiver knows the channel (which in practice is achieved by transmitting pilot signals), the mutual information $I[(\mathbf{y}, \mathbf{G}); \mathbf{x}]$ is given by

$$I = \log\det(\mathbf{1} + \mathbf{G}^\dagger \mathbf{Q} \mathbf{G}) = \mathrm{Tr}\log(\mathbf{1} + \mathbf{G}^\dagger \mathbf{Q} \mathbf{G}) \tag{2}$$

where **1** is the identity matrix. Throughout this paper we will measure information in nats/sec/Hz where one nat is equal to e bits ($e = 2.718\ldots$). That is, the log in this equation is defined to be a natural log.

The quantity we are concerned with is the so-called ergodic or mean average of this information over instantiations of the channel within some ensemble of channels which we write as $\langle I \rangle$. We will write brackets $\langle\rangle$ to denote an average over instantiations in a given ensemble of instantiations. We would like to maximize $\langle I \rangle$ over all possible transmission covariances $\mathbf{Q}$. More specifically, in this paper, we would like to know when $\mathbf{Q}$ being of rank one (ie, a beamforming solution) is optimal, and what that particular beamforming solution is.

3. Concavity and Optimality Condition

An essential part of the following calculations rely on the concavity of the mutual information (Eq. 2) with respect to the transmission covariance. To show this, we vary the transmission covariance slightly by writing

$$\mathbf{Q} = \mathbf{Q}^0 + \epsilon \mathbf{Q}^1 \tag{3}$$

where both $\mathbf{Q}$ and $\mathbf{Q}_0$ are nonnegative definite Hermitian matrices with $\mathrm{Tr}\mathbf{Q} = \mathrm{Tr}\mathbf{Q}^0 = P$, and $\mathbf{Q}^1$ has trace zero. We can then write Eq. 2 as

$$\begin{aligned} I &= \mathrm{Tr}\log(\mathbf{1} + \mathbf{G}^\dagger[\mathbf{Q}^0 + \epsilon\mathbf{Q}^1]\mathbf{G}) \\ &= \mathrm{Tr}\log\left([\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G}][\mathbf{1} + (\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})^{-1}\epsilon\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G})]\right) \\ &= I_0 + \mathrm{Tr}\log\left([\mathbf{1} + (\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})^{-1}\epsilon\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G})]\right) \end{aligned}$$

where $I_0 = \mathrm{Tr}\log(\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})$ is the value of I at $\epsilon = 0$. We now expand to second order in ϵ to obtain

$$\begin{aligned} I &= I_0 + \epsilon\mathrm{Tr}\left[(\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})^{-1}\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right] \\ &- \frac{\epsilon^2}{2}\mathrm{Tr}\left[(\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})^{-1}\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}(1 + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})^{-1}\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right] + \ldots\ . \end{aligned}$$

We see that the trace in the second order term is of the form $\mathrm{Tr}[\mathbf{MNMN}]$ where both $\mathbf{M}$ and $\mathbf{N}$ are Hermitian matrices. Using the cyclic property of the trace, this term can be rewritten as $\mathrm{Tr}[(\sqrt{\mathbf{M}}\,\mathbf{N}\sqrt{\mathbf{M}}\,)^2]$ which is nonnegative definite. Therefore, we have the mutual information being a concave function of $\mathbf{Q}$ at every allowable value of $\mathbf{Q}$. Thus, the average of the mutual information over any ensemble is also everywhere concave.

Since the mutual information is everywhere concave, the optimality condition for maximizing the average of the mutual information can be written as follows: A nonnegative definite Hermitian matrix $\mathbf{Q}^0$ is optimal if and only if

$$\frac{d\langle I\rangle}{d\epsilon} = \left\langle \mathrm{Tr}\left[(\mathbf{1} + \mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G})^{-1}\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right]\right\rangle \leq 0 \tag{4}$$

for all zero trace Hermitian matrices $\mathbf{Q}^1$ such that $Q^0 + \epsilon Q_1$ remains nonnegative definite for infinitesimal ϵ. Here, we are simply saying that any allowable small change of $\mathbf{Q}^0$ at fixed power decreases the mutual information (allowable means that the new $\mathbf{Q}$ is still Hermitian and nonnegative definite). The concavity property tells us that optimality with respect to small changes in $\mathbf{Q}$ implies global optimality. Note that the $\langle\rangle$ brackets in Eq. 4 represent average over the ensemble of channels. The derived optimality condition will, of course, depend on the definition of the ensemble.

If $\mathbf{Q}^0$ corresponds to beamforming (ie, has rank one) we can use the identities of Appendix A to reduce the optimality condition to

$$\frac{d\langle I\rangle}{d\epsilon} = \left\langle \mathrm{Tr}\left[\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right] - \frac{\mathrm{Tr}\left[\mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G}\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right]}{1+\mathrm{Tr}[\mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G}]}\right\rangle \leq 0. \tag{5}$$

The calculational difficulty here is that the channel matrix appears in the denominator. The first step in attacking this problem is to move the denominator into the exponent to give the following equivalent beamforming optimality condition

$$\left\langle \mathrm{Tr}\left[\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right] - \int_0^\infty dx\; e^{-x(1+\mathrm{Tr}[\mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G}])}\;\mathrm{Tr}\left[\mathbf{G}^\dagger\mathbf{Q}^0\mathbf{G}\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right]\right\rangle \leq 0 \tag{6}$$

which will be our calculational starting point below.

4. Covariance Feedback

If the covariance of the channel is fed back to the transmitter, we then assume a channel model where $\mathbf{G}$ has zero mean and is normally distributed with that given covariance. In this paper we assume a factorizable form of the covariance $\mathbf{T}\otimes\mathbf{R}$ with $\mathbf{T}$ being an n_{T} dimensional nonnegative definite Hermitian matrix and $\mathbf{R}$ being an n_{R} dimensional nonnegative definite Hermitian matrix (so that $\mathbf{T}\otimes\mathbf{R}$ is an $n_{\mathrm{T}}n_{\mathrm{R}}$ dimensional matrix, see Eq. 8 below). This factorizable form has been used by many authors[**1, 2**] and has been found to be a good representation of MIMO channels in highly scattering environments[**2**]. Further, it can be shown that this factorizable form is the first term in a controlled approximation for diffusive environments[**1**]. The matrices $\mathbf{T}$ and $\mathbf{R}$ become nontrivial when there are non-negligible antenna correlations at the transmitter or receiver respectively. We will sometimes use the notation that T_i and R_α are the eigenvalues of the matrices $\mathbf{T}$ and $\mathbf{R}$.

Thus we have $\mathbf{G}$ being $\mathcal{N}(\mathbf{0},\mathbf{T}\otimes\mathbf{R})$ such that

$$\langle \mathbf{G}\rangle = 0 \tag{7}$$

and

$$\left\langle G_{i\alpha}G^*_{j\beta}\right\rangle = R_{\alpha\beta}T_{ij} \quad . \tag{8}$$

Equivalently, we could write have written $\mathbf{G} = \sqrt{\mathbf{T}}\;\mathbf{Z}^{(iid)}\;\sqrt{\mathbf{R}}$ where $\mathbf{Z}$ is an i.i.d. Gaussian random matrix with unit covariance.

The ensemble average of any function F of $\mathbf{G}$, which we write as $\langle F(\mathbf{G})\rangle$, can be written more explicitly as

$$\langle F(\mathbf{G})\rangle = \int d\mu(\mathbf{G}) e^{-\mathrm{Tr}[\mathbf{G}^\dagger\mathbf{T}^{-1}\mathbf{G}\mathbf{R}^{-1}]}\,F(\mathbf{G}) \tag{9}$$

which is shorthand for

$$\frac{1}{\det[\mathbf{R}\otimes\mathbf{T}]}\prod_{i=1}^{n_{\mathrm{T}}}\prod_{\alpha=1}^{n_{\mathrm{R}}}\left[\frac{1}{\pi}\int d\mathrm{Re}G_{i\alpha}\int d\mathrm{Im}G_{i\alpha}\right] e^{-\mathrm{Tr}[\mathbf{G}^\dagger\mathbf{T}^{-1}\mathbf{G}\mathbf{R}^{-1}]}\,F(\mathbf{G}) \tag{10}$$

We also note that the integration measure is correctly defined so that $\langle 1\rangle = 1$, and so that Eq. 8 is satisfied.

For this situation of covariance feedback, we will use to our advantage the fact that the optimal transmission covariance $\mathbf{Q}$ must be diagonal in the same basis as $\mathbf{T}$ (ie, $\mathbf{Q}$ and $\mathbf{T}$ have the same eigenvectors). This fact has been proven previously in

references [**3, 5**]. Although these particular proofs considered the case of $\mathbf{R} = \mathbf{1}$, it is easy to see that they apply equally well to the case of nontrivial $\mathbf{R}$. Furthermore, it is proven in these works that an optimal $\mathbf{Q}$ must have values along the diagonal that are ordered in the same manner as the eigenvalues of $\mathbf{T}$. In other words the eigenvector corresponding to the largest eigenvalue of $\mathbf{T}$ is also the eigenvector corresponding to the largest eigenvalue of $\mathbf{Q}$. Similarly, the eigenvector corresponding to the second largest eigenvalue of $\mathbf{T}$ is also the eigenvector corresponding to the second largest eigenvalue of $\mathbf{Q}$; and so forth. Thus, if a rank one (beamforming) $\mathbf{Q}^0$ is optimal, then it must be of the form

$$\mathbf{Q}^0 = P\,\mathbf{t}_1 \otimes \mathbf{t}_1^\dagger \tag{11}$$

where $\mathbf{t}_1$ is the normalized n_{T}-dimensional eigenvector of the matrix $\mathbf{T}$ corresponding to the largest eigenvalue T_1 of this matrix. (In more explicit notation $Q^0_{ij} = P[t_1]_i[t_1]^*_j$).

Since the optimal $\mathbf{Q}$ must be diagonal in the $\mathbf{T}$ basis, in the beamforming condition Eq. 6, one need only consider $\mathbf{Q}^1$'s that are also diagonal in this basis. Furthermore, since we must have $\mathbf{Q}^0 + \epsilon\mathbf{Q}^1$ nonnegative definite, we need only consider $\mathbf{Q}^1$'s of the form (recall $\mathbf{Q}^1$ must be traceless)

$$\mathbf{Q}^1 = \sum_{i=2}^{n_{\mathrm{T}}} q_i \left\{ \mathbf{t}_i \otimes \mathbf{t}_i^\dagger - \mathbf{t}_1 \otimes \mathbf{t}_1^\dagger \right\} \tag{12}$$

where $\mathbf{t}^i$ are the eigenvectors of $\mathbf{T}$ and all $q_i > 0$.

Working in a basis where both $\mathbf{R}$ and $\mathbf{T}$ are diagonal (and $\mathbf{Q}^0$ and $\mathbf{Q}^1$ therefore), and using the above forms (Eq. 11 and 12) for $\mathbf{Q}^0$ and $\mathbf{Q}^1$, the optimality condition Eq. 6) is rewritten as

$$\left\langle \mathrm{Tr}\left[\mathbf{G}^\dagger\mathbf{Q}^1\mathbf{G}\right]\right\rangle = \left\langle \sum_{\alpha=1}^{n_{\mathrm{R}}}\sum_{j=2}^{n_{\mathrm{T}}} q_j(|G_{j\alpha}|^2 - |G_{1\alpha}|^2)\right\rangle \leq \tag{13}$$

$$\left\langle \int_0^\infty dx\; e^{-x-Px\sum_{\alpha=1}^{n_{\mathrm{R}}}|G_{1\alpha}|^2} \sum_{\alpha,\beta=1}^{n_{\mathrm{R}}} PG^*_{1\alpha}G_{1\beta} \sum_{j=2}^{n_{\mathrm{T}}} q_j \left[G^*_{j\beta}G_{j\alpha} - G^*_{1\beta}G_{1\alpha}\right]\right\rangle .$$

(Here we have used the fact that the covariance of $\mathbf{G}$ is diagonal in both j and α). Taking some of the averages, can be reduced to

$$\sum_{\alpha=1}^{n_{\mathrm{R}}}\sum_{j=2}^{n_{\mathrm{T}}} q_j(T_jR_\alpha - T_1R_\alpha) \leq$$

$$\left\langle P\int_0^\infty e^{-x}\sum_{j=2}^{n_{\mathrm{T}}} q_j \left(\sum_{\alpha=1}^{n_{\mathrm{R}}}|G_{1\alpha}|^2T_jR_\alpha - \sum_{\alpha,\beta=1}^{n_{\mathrm{R}}}\frac{d^2}{d(Px)^2}\right) e^{-Px\sum_{\alpha=1}^{n_{\mathrm{R}}}|G_{1\alpha}|^2}\right\rangle . \tag{14}$$

Writing out the expectation on the right hand side explicitly as an integral we obtain

$$\mathrm{Tr}[\mathbf{R}]\sum_{j=2}^{n_{\mathrm{T}}} q_j(T_j - T_1) \leq P\int d\mu(\mathbf{G})e^{-\mathrm{Tr}[\mathbf{R}^{-1}\mathbf{G}^\dagger\mathbf{T}^{-1}\mathbf{G}]} \tag{15}$$

$$\int_0^\infty dx\; e^{-x}\sum_{j=2}^{n_{\mathrm{T}}} q_j \left(\sum_{\alpha=1}^{n_{\mathrm{R}}}|G_{1\alpha}|^2T_jR_\alpha - \frac{d^2}{d(Px)^2}\right) e^{-Px\sum_{\alpha=1}^{n_{\mathrm{R}}}|G_{1\alpha}|^2}$$

Writing in the exponent

$$\mathrm{Tr}[\mathbf{R}^{-1}\mathbf{G}^\dagger\mathbf{T}^{-1}\mathbf{G}] = \sum_{\alpha=1}^{n_\mathrm{R}}\sum_{j=1}^{n_\mathrm{T}} |G_{j\alpha}|^2 T_j^{-1} R_\alpha^{-1} \tag{16}$$

the integrals over $\mathbf{G}$ can be performed explictly to yield

$$\mathrm{Tr}[\mathbf{R}] \sum_{i=2}^{n_\mathrm{T}} q_i(T_i - T_1) \leq \tag{17}$$
$$P\int_0^\infty dx\ e^{-x} \sum_{j=2}^{n_\mathrm{T}} q_j \left\{ \sum_{\alpha=1}^{n_\mathrm{R}} \frac{T_j R_\alpha}{T_1^{-1}R_\alpha^{-1}+Px} - \frac{d^2}{d(Px)^2}\right\} \prod_{\gamma=1}^{n_\mathrm{R}} \left[\frac{T_1^{-1}R_\gamma^{-1}}{T_1^{-1}R_\gamma^{-1}+Px}\right].$$

There is only a single appearance of T_j on the right hand side for $j \neq 1$. The magnitude of this term can be easily bounded to be (see Appendix B)

$$P\int_0^\infty dx\ e^{-x} \prod_{\gamma=1}^{n_\mathrm{R}} \left[\frac{T_1^{-1}R_\gamma^{-1}}{T_1^{-1}R_\gamma^{-1}+Px}\right] \sum_{\alpha=1}^{n_\mathrm{R}} \frac{R_\alpha}{T_1^{-1}R_\alpha^{-1}+Px} < \mathrm{Tr}\mathbf{R} \tag{18}$$

which then allows us to symbolically write Eq. 17 as

$$\sum_{i=2}^{n_\mathrm{T}} q_i[AT_i - C] \leq 0 \tag{19}$$

with $A > 0$ due to the inequality Eq. 18 (where all dependence on q_i and T_i for $i \geq 2$ are explicit here). Since $T_2 \geq T_3 \geq T_4 \ldots$, using this form, one can see that if the the above optimality condition (Eq. 17) is true for the case of $q_2 = 1$ and all other q_i's equal to zero, then it is true for any selection of q_i's (recall that all $q_i > 0$) . It is then sufficient to only consider this one case of $q_2 = 1$. This is simply the statement that beamforming optimality is completely determined by testing whether transferring a small amount of weight from the strongest eigendirection (the $\mathbf{t}_1$ direction) to the next strongest direction (the $\mathbf{t}_2$ direction or q_2 direction) increases capacity or not (This principle has been deduced previously[**3, 5**]). Thus setting $q_2 = 1$ and all other $q_i = 0$, Eq. 17 is further simplified by integrating by parts twice (paying careful attention to the boundary terms) to give

$$T_2\mathrm{Tr}[\mathbf{R}] - \frac{1}{P} \leq P\int_0^\infty dx\ e^{-x} \left\{ \sum_{\alpha=1}^{n_\mathrm{R}} \frac{T_2 R_\alpha}{T_1^{-1}R_\alpha^{-1}+Px} - \frac{1}{P^2}\right\} \prod_{\gamma=1}^{n_\mathrm{R}} \left[\frac{T_1^{-1}R_\gamma^{-1}}{T_1^{-1}R_\gamma^{-1}+Px}\right]. \tag{20}$$

For the simple case where $\mathbf{R} = \mathbf{1}$, a case that can be solved[**5**] with other methods (and previously studied by [**3**]), one can evaluate these integrals explicitly using the fact that for $A/P > 0$ we have [**9**]

$$P\int_0^\infty dx \frac{e^{-x}}{(A+Px)^m} = e^{A/P}P^{(1-m)}\Gamma[-m+1, A/P] \tag{21}$$

with Γ the incomplete gamma function. With some algebra, using the identity[**9**] $\Gamma[m+1,x] = m\Gamma[m,x] + x^m e^{-x}$ the result can be simplified to the trivial form

$$\frac{e^{-1/(T_1P)}T_2P(T_1P)^{n_\mathrm{R}}}{1+T_2P} \leq \Gamma\left[-n_\mathrm{R}, 1/(T_1P)\right]. \tag{22}$$

which agrees with previous derived results[**5**] for the special case of $\mathbf{R} = \mathbf{1}$. The current derivation, however, is simpler and far more general.

In the more general case where $\mathbf{R} \neq \mathbf{1}$, the integrals in Eq. 20 can still be performed analytically using Eq. 21 after decomposing products into partial fractions. We write

$$\prod_{\gamma=1}^{n_{\mathrm{R}}} \left[\frac{T_1^{-1} R_\gamma^{-1}}{T_1^{-1} R_\gamma^{-1} + Px} \right] = \frac{1}{\det[\mathbf{R}]} \sum_{\gamma=1}^{n_{\mathrm{R}}} \frac{T_1^{-1}}{T_1^{-1} R_\gamma^{-1} + Px} \prod_{\beta=1, \beta\neq\gamma}^{n_{\mathrm{R}}} \left[\frac{1}{R_\beta^{-1} - R_\gamma^{-1}} \right] \tag{23}$$

which enables us to write Eq. 20 as

$$\begin{aligned} T_2 \mathrm{Tr}[\mathbf{R}] - \frac{1}{P} \quad \leq \quad & \frac{P T_1^{-1}}{\det[\mathbf{R}]} \int_0^\infty dx \; e^{-x} \sum_{\gamma=1}^{n_{\mathrm{R}}} \left(\prod_{\beta=1, \beta\neq\gamma}^{n_{\mathrm{R}}} \left[\frac{1}{R_\beta^{-1} - R_\gamma^{-1}} \right] \right) \\ & \left[\frac{T_2 R_\gamma}{(T_1^{-1} R_\gamma^{-1} + Px)^2} - \frac{1}{P^2 (T_1^{-1} R_\gamma^{-1} + Px)} \right. \\ & \left. + \sum_{\alpha=1, \alpha\neq\gamma}^{n_{\mathrm{R}}} \frac{T_2 R_\alpha}{(T_1^{-1} R_\gamma^{-1} + Px)(T_1^{-1} R_\alpha^{-1} + Px)} \right]. \end{aligned} \tag{24}$$

Writing the last term as

$$\begin{aligned} & \frac{1}{(T_1^{-1} R_\gamma^{-1} + Px)(T_1^{-1} R_\alpha^{-1} + Px)} = \\ & \frac{T_1}{R_\alpha^{-1} - R_\gamma^{-1}} \left[\frac{1}{(T_1^{-1} R_\gamma^{-1} + Px)} - \frac{1}{(T_1^{-1} R_\alpha^{-1} + Px)} \right] \end{aligned}$$

and using Eq. 21 we then simplify Eq. 24 to

$$\begin{aligned} P T_2 \mathrm{Tr}[\mathbf{R}] - 1 \leq & \frac{1}{P T_1 \det[\mathbf{R}]} \sum_{\gamma=1}^{n_{\mathrm{R}}} \left(\prod_{\beta=1, \beta\neq\gamma}^{n_{\mathrm{R}}} \left[\frac{1}{R_\beta^{-1} - R_\gamma^{-1}} \right] \right) \\ & \left[e^{1/(T_1 R_\gamma P)} \left(P R_\gamma T_2 \; \Gamma[-1, 1/(T_1 R_\gamma P)] - \; \Gamma[0, 1/(T_1 R_\gamma P)] \right) \right. \\ & \left. + \sum_{\alpha\neq\gamma} \frac{P^2 T_1 T_2 R_\alpha}{R_\alpha^{-1} - R_\gamma^{-1}} \left(e^{1/(T_1 R_\gamma P)} \; \Gamma[0, 1/(T_1 R_\gamma P)] - e^{1/(T_1 R_\alpha P)} \; \Gamma[0, 1/(T_1 R_\alpha P)] \right) \right] \end{aligned} \tag{25}$$

which is our final analytic result for the general optimality condition for beamforming in a MIMO system with covariance feedback (where T_1 and T_2 are the largest and second largest eigenvalues of the transmission correlation matrix $\mathbf{T}$, and R_γ are the eigenvalues of the receiver correlation matrix $\mathbf{R}$). It is interesting to note that the beamforming condition at the transmitter depends not only on the eigenvalues of the correlation matrix at the transmitter T, but in general it also depends on the correlations of the receiving array.

5. Mean Feedback

If the mean of the channel is fed back to the transmitter, we then assume a channel model that has nonzero mean $\mathbf{H}$ and white covariance around this mean

with some amplitude σ^2 which we assume is also known by the transmitter. Thus we have $\mathbf{G}$ being $\mathcal{N}(\mathbf{H}, \sigma^2 \mathbf{1} \otimes \mathbf{1})$ such that

$$\langle \mathbf{G} \rangle = \mathbf{H} \tag{26}$$

and

$$\langle (G_{i\alpha} - H_{i\alpha})(G_{j\beta} - H_{j\beta})^* \rangle = \sigma^2 \delta_{ij} \delta_{\alpha\beta} \tag{27}$$

Equivalently, we could write

$$\mathbf{G} = \mathbf{H} + \sigma \mathbf{Z}^{(iid)} \tag{28}$$

where $\mathbf{Z}$ is an i.i.d. Gaussian random matrix with unit covariance.

Again, the ensemble average of any function F of $\mathbf{G}$ we write as $\langle F(\mathbf{G}) \rangle$ can be written more explicitly as

$$\langle F(\mathbf{G}) \rangle = \int d\mu(\mathbf{Z}) e^{-\mathrm{Tr}[\mathbf{Z}^\dagger \mathbf{Z}]} F(\mathbf{H} + \sigma \mathbf{Z}) \tag{29}$$

which here is shorthand for

$$\prod_{i=1}^{n_\mathrm{T}} \prod_{\alpha=1}^{n_\mathrm{R}} \left[\frac{1}{\pi} \int d\mathrm{Re} Z_{i\alpha} \int d\mathrm{Im} Z_{i\alpha} \right] e^{-\mathrm{Tr}[\mathbf{Z}^\dagger \mathbf{Z}]} F(\mathbf{H} + \sigma \mathbf{Z}) \tag{30}$$

We note again that the integration measure is correctly defined so that $\langle 1 \rangle = 1$ and Eq. 27 is satisfied.

We now rewrite the beamforming optimality condition Eq. 6 as

$$\begin{aligned} \left\langle \mathrm{Tr}\left[(\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^1 (\mathbf{H} + \sigma \mathbf{Z}) \right] \right\rangle \leq \Bigg\langle \int_0^\infty dx \; e^{-x(1 + \mathrm{Tr}[(\mathbf{H}+\sigma\mathbf{Z})^\dagger \mathbf{Q}^0 (\mathbf{H}+\sigma\mathbf{Z})]} \\ \mathrm{Tr}\left[(\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^0 (\mathbf{H} + \sigma \mathbf{Z}) (\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^1 (\mathbf{H} + \sigma \mathbf{Z}) \right] \Bigg\rangle \end{aligned}$$

or

$$\begin{aligned} \mathrm{Tr}\left[\mathbf{H}^\dagger \mathbf{Q}^1 \mathbf{H} \right] \leq & \int_0^\infty dx \int d\mu(\mathbf{Z}) e^{-\mathrm{Tr}[\mathbf{Z}^\dagger \mathbf{Z}] - x(1 + \mathrm{Tr}[(\mathbf{H}+\sigma\mathbf{Z})^\dagger \mathbf{Q}^0 (\mathbf{H}+\sigma\mathbf{Z})]} \\ & \mathrm{Tr}\left[(\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^0 (\mathbf{H} + \sigma \mathbf{Z}) (\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^1 (\mathbf{H} + \sigma \mathbf{Z}) \right] \end{aligned}$$

where on the left hand side we have used the fact that $\mathrm{Tr}\mathbf{Q}^1 = 0$. With some algebraic manipulation we massage this into

$$\begin{aligned} \mathrm{Tr}\left[\mathbf{H}^\dagger \mathbf{Q}^1 \mathbf{H} \right] \leq & \int_0^\infty dx \, e^{-x(1 + \mathrm{Tr}[\mathbf{H}^\dagger \mathbf{Q}^0 \mathbf{M} \mathbf{H}])} \\ & \int d\mu(\mathbf{Z}) e^{-\mathrm{Tr}[\{\mathbf{Z}^\dagger + \mathbf{H}^\dagger x \sigma \mathbf{Q}^0 \mathbf{M}\} \mathbf{M}^{-1} \{\mathbf{Z} + \mathbf{M} x \sigma \mathbf{Q}^0 \mathbf{H}\}]} \\ & \mathrm{Tr}\left[(\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^0 (\mathbf{H} + \sigma \mathbf{Z}) (\mathbf{H} + \sigma \mathbf{Z})^\dagger \mathbf{Q}^1 (\mathbf{H} + \sigma \mathbf{Z}) \right] \end{aligned}$$

where we have defined the Hermitian matrix (see appendix A)

$$\mathbf{M} = (\mathbf{1} + x\sigma^2 \mathbf{Q}^0)^{-1} = \mathbf{1} - \frac{x\sigma^2 \mathbf{Q}^0}{1 + x\sigma^2 \mathrm{Tr}[\mathbf{Q}^0]}. \tag{31}$$

Making the change of variables

$$\tilde{\mathbf{Z}} = \mathbf{Z} + \mathbf{M} x \sigma \mathbf{Q}^0 \mathbf{H} \tag{32}$$

this then becomes

$$\text{(33)}\qquad \begin{aligned}\mathrm{Tr}\left[\mathbf{H}^\dagger\mathbf{Q}^1\mathbf{H}\right] \le \int_0^\infty dx\, e^{-x(1+\mathrm{Tr}[\mathbf{H}^\dagger\mathbf{Q}^0\mathbf{M}\mathbf{H}])} \int d\mu(\tilde{\mathbf{Z}}) e^{-\mathrm{Tr}[\tilde{\mathbf{Z}}^\dagger\mathbf{M}^{-1}\tilde{\mathbf{Z}}]} \\ \mathrm{Tr}\left[(\mathbf{H}^\dagger\mathbf{M}+\sigma\tilde{\mathbf{Z}}^\dagger)\mathbf{Q}^0(\mathbf{M}\mathbf{H}+\sigma\tilde{\mathbf{Z}})(\mathbf{H}^\dagger\mathbf{M}+\sigma\tilde{\mathbf{Z}}^\dagger)\mathbf{Q}^1(\mathbf{M}\mathbf{H}+\sigma\tilde{\mathbf{Z}})\right].\end{aligned}$$

This can now be integrated using Wick's theorem[**10, 11**] to yield the expression

$$\begin{aligned}\mathrm{Tr}\left[\mathbf{H}^\dagger\mathbf{Q}^1\mathbf{H}\right] \le \int_0^\infty dx\, \frac{e^{-x(1+\mathrm{Tr}[\mathbf{H}^\dagger\mathbf{Q}^0\mathbf{M}\mathbf{H}])}}{\det[\mathbf{M}^{-1}]} \Bigg\{ & \mathrm{Tr}\left[\mathbf{H}^\dagger\mathbf{M}\mathbf{Q}^0\mathbf{M}\mathbf{H}\mathbf{H}^\dagger\mathbf{M}\mathbf{Q}^1\mathbf{M}\mathbf{H}\right] \\ + \quad & \sigma^2\mathrm{Tr}\left[\mathbf{Q}^0\mathbf{M}\right]\mathrm{Tr}\left[\mathbf{H}^\dagger\mathbf{M}\mathbf{Q}^1\mathbf{M}\mathbf{H}\right] + \sigma^2\mathrm{Tr}\left[\mathbf{Q}^1\mathbf{M}\right]\mathrm{Tr}\left[\mathbf{H}^\dagger\mathbf{M}\mathbf{Q}^0\mathbf{M}\mathbf{H}\right] \\ + \quad & \sigma^2 n_{\mathrm{R}}\mathrm{Tr}\left[\mathbf{M}\mathbf{Q}^0\mathbf{M}\mathbf{H}\mathbf{H}^\dagger\mathbf{M}\mathbf{Q}^1\right] + \sigma^2 n_{\mathrm{R}}\mathrm{Tr}\left[\mathbf{M}\mathbf{Q}^1\mathbf{M}\mathbf{H}\mathbf{H}^\dagger\mathbf{M}\mathbf{Q}^0\right] \\ + \quad & \sigma^4 n_{\mathrm{R}}\mathrm{Tr}\left[\mathbf{Q}^0\mathbf{M}\right]\mathrm{Tr}\left[\mathbf{Q}^1\mathbf{M}\right] + \sigma^4 {n_{\mathrm{R}}}^2\mathrm{Tr}\left[\mathbf{Q}^0\mathbf{M}\mathbf{Q}^1\mathbf{M}\right] \Bigg\}.\end{aligned}$$

We now define H_i to be the eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ and $\mathbf{h}_i$ to be the corresponding eigenvectors. We also assume the eigenvalues are ordered $H_1 > H_2 > \ldots$. Analogous to above covariance feedback case, the beamforming solution for this mean feedback case should be of the form

$$\text{(34)}\qquad \mathbf{Q}^0 = P\mathbf{h}_1 \otimes \mathbf{h}_1^\dagger \quad .$$

We now write $\mathbf{Q}^1$ in the general form

$$\text{(35)}\qquad \mathbf{Q}^1 = \sum_{i=2}^{n_{\mathrm{T}}} q_i \left\{\mathbf{h}_i \otimes \mathbf{h}_i^\dagger - \mathbf{h}_1 \otimes \mathbf{h}_1^\dagger\right\} + \sum_{i,j=1;i\neq j}^{n_{\mathrm{T}}} q_{ij}\mathbf{h}_i \otimes \mathbf{h}_j^\dagger$$

with $q_{ij} = q_{ji}^*$. (Note that we have not taken $\mathbf{Q}^1$ to be diagonal yet for generality. Nonetheless, the restriction $q_i > 0$ still applies to assure that $\mathbf{Q}^0 + \epsilon\mathbf{Q}^1$ remains nonnegative definite).

Using this form, the above expression can be simplified (using the results of Appendix A) to yield

$$\text{(36)}\qquad \begin{aligned}&\sum_{i=2}^{n_{\mathrm{T}}} q_i(H_i - H_1) \le P\int_0^\infty dx \frac{e^{-x[1+H_1P/(1+x\sigma^2P)]}}{(1+x\sigma^2P)^2} \\ &\sum_{i=2}^{n_{\mathrm{T}}} q_i\left\{\sigma^2 H_i - \frac{H_1^2}{(1+x\sigma^2P)^3} + \frac{\sigma^2H_1[x\sigma^2P-(1+2n_{\mathrm{R}})]}{(1+x\sigma^2P)^2} + \frac{\sigma^4 n_{\mathrm{R}}[x\sigma^2P - n_{\mathrm{R}}]}{1+x\sigma^2P}\right\}.\end{aligned}$$

Note that the off-diagonal terms of $\mathbf{Q}^1$ don't enter and therefore need not be considered (as one might have guessed).

Analogous to the argument given near Eq. 18, we use the inequality

$$\text{(37)}\qquad \sigma^2 P\int_0^\infty dx \frac{e^{-x[1+H_1P/(1+x\sigma^2P)]}}{(1+x\sigma^2P)^2} \le \int_0^\infty dx \frac{e^{-x}\sigma^2P}{(1+x\sigma^2P)^2} \le 1$$

(where the last inequality is demonstrated in Eq. 44) to again allow us to symbolically write Eq. 36 as

$$\sum_{i=2}^{n_{\rm T}} q_i(AH_i - C) \leq 0 \tag{38}$$

with $A > 0$. Once again, we see that this inequality holds for arbitrary $\mathbf{Q}^1$ if and only if it holds for the case of $q_2 = 1$ and all other $q_i = 0$. As with the above case, we can then reduce our beamforming optimality condition to

$$(H_2 - H_1) \leq P \int_0^\infty dx \frac{e^{-x[1+H_1P/(1+x\sigma^2P)]}}{(1+x\sigma^2P)^2} \left\{ \sigma^2 H_2 - \frac{H_1^2}{(1+x\sigma^2P)^3} + \frac{\sigma^2 H_1[x\sigma^2P - (1+2n_{\rm R})]}{(1+x\sigma^2P)^2} + \frac{\sigma^4 n_{\rm R}[x\sigma^2P - n_{\rm R}]}{1+x\sigma^2P} \right\} \tag{39}$$

which is our final analytic result for optimality of beamforming in the mean feedback case. (Here H_1 and H_2 are the largest and second largest eigenvalues of the mean channel squared $\mathbf{H}\mathbf{H}^\dagger$, and σ is the fluctuation of the channel around this average).

6. Summary

We have shown how one can determine whether or not beamforming is optimal for MIMO systems when partial information about the channel is known at the transmitter. We have given analytic results for the cases of channel covariance feedback and mean channel feedback. Both of these cases seem realizable for practical systems. Further discussion and interpretation of these results will be presented in a forthcoming paper.

Appendix A. Identities for Rank One Matrices

A rank one Hermitian matrix $\mathbf{M}$ can be written as $M_{ij} = v_i v_j^*$. We then have the identity

$$\mathbf{M}^2 = \mathbf{M}\,{\rm Tr}\mathbf{M} \tag{40}$$

which is easily seen since $\sum_j M_{ij}M_{jk} = v_i v_k^* \sum_j |v_j|^2 = M_{ik}{\rm Tr}\mathbf{M}$. Using this result, we can establish that

$$(\mathbf{1}+\mathbf{M})^{-1} = \mathbf{1} - \frac{\mathbf{M}}{1+{\rm Tr}\mathbf{M}} \tag{41}$$

which can easily be checked since

$$(\mathbf{1}+\mathbf{M})\left[\mathbf{1} - \frac{\mathbf{M}}{1+{\rm Tr}\mathbf{M}}\right] = \mathbf{1} + \mathbf{M} - \frac{\mathbf{M}+\mathbf{M}^2}{1+{\rm Tr}\mathbf{M}} = \mathbf{1} \tag{42}$$

Finally, we note another useful identity that

$$\det[\mathbf{1}+\mathbf{M}] = 1 + {\rm Tr}\mathbf{M} \tag{43}$$

which is most easily proven by simply working in a basis where $\mathbf{M}$ is diagonal.

Appendix B. A Bound

$$\begin{aligned} & P\int_0^\infty dx\ e^{-x}\prod_{\gamma=1}^{n_{\mathrm{R}}}\left[\frac{T_1^{-1}R_\gamma^{-1}}{T_1^{-1}R_\gamma^{-1}+Px}\right]\frac{R_\alpha}{T_1^{-1}R_\alpha^{-1}+Px} \\ < \quad & P\int_0^\infty dx\ e^{-x}\frac{T_1^{-1}}{(T_1^{-1}R_\alpha^{-1}+Px)^2} \\ = \quad & R_\alpha\int_0^\infty dx\ e^{-x}\frac{PR_\alpha T_1}{(1+PxR_\alpha T_1)^2} < R_\alpha \end{aligned}$$

The last inequality follows since by integration by parts

$$\begin{aligned} \int_0^\infty dx\frac{e^{-x}A}{(1+Ax)^2} &= 1-\int_0^\infty dx\frac{e^{-x}}{1+Ax} \\ &= \int_0^\infty dx\frac{e^{-x}Ax}{1+Ax} < \int_0^\infty dx\, e^{-x} = 1 \end{aligned} \tag{44}$$

with the inequality holding for $A > 0$.

References

[1] A. L. Moustakas *et al.*, "Communication through a diffusive medium: Coherence and capacity," *Science*, vol. 287, pp. 287–290, January 2000.

[2] F. Ling, D. Chizhik, F. Rashid-Fanokhi, and A. Lozano, "Antenna separation and capacity of blast in correlated channels," *IEEE-APS Conference on Antennas and Propagation for Wireless Communication*, pp. 183–185, 2000.

[3] S. Vishwanath S. A. Jafar and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," *ICC2001, Helsinki Finland*, 2001.

[4] S. H. Simon and A. L. Moustakas, "Optimizing MISO systems with partial channel knowledge," *submitted to IEEE Trans. of Inform Theory*, submitted. *Also available from http://mars.bell-labs.com.*

[5] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE Journal on Selected Areas in Communications*, to be published. *Also available from http://mars.bell-labs.com.*

[6] A. Narula *et al.*, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, October 1998.

[7] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. of Inform Theory*, vol. 47, no. 6, pp. 2632–2639, Sept 2001.

[8] A. J. Syed and A. J. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," *International Symposium on Information Theory (ISIT) 2001.*

[9] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, London, 5 edition, 1994.

[10] L. H. Ryder, *Quantum Field Theory*, Cambridge University Press, 1985.

[11] A. L. Moustakas, A. Sengupta, and S. H. Simon, "MIMO Capacity Distribution in the Presence of Correlated Signal, Interference and Noise : A (not so) Large N analyis" *submitted to IEEE Trans. of Inform Theory. Also Available at http://mars.bell-labs.com, May 2002.*

Lucent Technologies, Murray Hill NJ
E-mail address: shsimon@lucent.com

Lucent Technologies, Whippany NJ
E-mail address: arislm@lucent.com