MIMO CAPACITY WITH ANTENNA SELECTION AND INTERFERENCE

Rick S. Blum

Electrical and Computer Engineering Department
Lehigh University

ABSTRACT

System capacity is considered for a group of interfering users employing single user detection and antenna selection of multiple transmit and receive antennas for flat Rayleigh fading channels with independent fading coefficients for each path. The case considered is that where there is very limited channel state information (only the selected antennas) at the transmitter, but channel state information is assumed at the receiver. The focus is on extreme cases with very weak interference or very strong interference. It is shown that the optimum signaling covariance matrix is sometimes different from the standard scaled identity matrix. In fact this is true even for cases without interference if SNR is sufficiently weak. Further the scaled identity matrix is actually that covariance matrix that yields worst performance if the interference is sufficiently strong.

1. INTRODUCTION

Consider a system where cochannel interference is present from $L-1$ other users. Let us focus on the Lth user and assume each user employs n_t transmit antennas and n_r receive antennas. In this case the vector of received complex baseband samples after matched filtering becomes

$$\mathbf{y}_L = \sqrt{\rho_L}\,\mathbf{H}_{L,L}\mathbf{x}_L + \sum_{j=1}^{L-1}\sqrt{\eta_{L,j}}\,\mathbf{H}_{L,j}\mathbf{x}_j + \mathbf{n} \qquad (1)$$

where $\mathbf{H}_{L,j}$ and $\mathbf{x}_j$ represent the normalized channel matrix and the normalized transmitted signal of user j respectively. The signal-to-noise ratio (SNR) of user L is ρ_L and the interference-to-noise ratio (INR) for user L due to interference from user j is $\eta_{L,j}$. For simplicity, we assume all of the interfering signals $\mathbf{x}_j, j=1,\dots,L-1$ are unknown to the receiver and we model each of them as being Gaussian distributed, the usual form of the optimum signal in MIMO problems. Then if we condition on $\mathbf{H}_{L,1},\dots,\mathbf{H}_{L,L}$, the interference-plus-noise from (1), $\sum_{j=1}^{L-1}\sqrt{\eta_{L,j}}\mathbf{H}_{L,j}\mathbf{x}_j+\mathbf{n}$, is Gaussian distributed with the covariance matrix $\mathbf{R}_L = \sum_{j=1}^{L-1}\eta_{L,j}\mathbf{H}_{L,j}\mathbf{S}_j\mathbf{H}_{L,j}^H + \mathbf{I}_{n_r}$ where $\mathbf{S}_j$ denotes the covariance matrix of $\mathbf{x}_j$ and $\mathbf{I}_{n_r}$ is the covariance matrix of $\mathbf{n}$. Under this conditioning, the interference-plus-noise is whitened by multiplying $\mathbf{y}_L$ by $\mathbf{R}_L^{-1/2}$. After performing this multiplication we can use results from [1] to express the ergodic mutual information between the input and output for the user of interest as in $I(\mathbf{x}_L;(\mathbf{y}_L,\mathcal{H}))=$

$$E\left\{log_2\left[\det\left(\mathbf{I}_{n_r}+\rho_L\mathbf{H}_{L,L}\mathbf{S}_L\mathbf{H}_{L,L}^H\mathbf{R}_L^{-1}\right)\right]\right\}. \qquad (2)$$

In (2) the identity $\det(\mathbf{I}+\mathbf{AB})=\det(\mathbf{I}+\mathbf{BA})$ was used. If we wish to compute total system mutual information we should find $\mathbf{S}_1,\dots\mathbf{S}_L$ to maximize $\Psi(\mathbf{S}_1,\dots,\mathbf{S}_L)=\sum_{i=1}^{L}I(\mathbf{x}_i;(\mathbf{y}_i,\mathcal{H}))$.

Now assume that each user in the system selects $n_{st}<n_t$ transmit antennas and $n_{sr}<n_r$ receive antennas[1] using an antenna selection algorithm. Then the observations from the selected antennas follow the model in (1) with n_t and n_r replaced by n_{st} and n_{sr} respectively and $\mathbf{H}_{i,j}$ replaced by $\tilde{\mathbf{H}}_{i,j}$. The matrix $\tilde{\mathbf{H}}_{i,j}$ is obtained by eliminating those columns and rows of $\mathbf{H}_{i,j}$ corresponding to unselected transmit and receive antennas. Thus we can write $\tilde{\mathbf{H}}_{i,j}=f(\mathbf{H}_{i,j})$ where the function f will choose $\tilde{\mathbf{H}}_{i,j}$ to maximize the instantaneous (and thus also the ergodic) mutual information.

2. OPTIMUM SIGNALING FOR SYSTEM CAPACITY

We define a general convex combination of two possible solutions, $(\mathbf{S}_1,\dots,\mathbf{S}_L)$ and $(\hat{\mathbf{S}}_1,\dots,\hat{\mathbf{S}}_L)$, as

$$\begin{aligned}(\bar{\mathbf{S}}_1,\dots,\bar{\mathbf{S}}_L) &= (1-t)(\mathbf{S}_1,\dots,\mathbf{S}_L)+t(\hat{\mathbf{S}}_1,\dots,\hat{\mathbf{S}}_L)\\ &= (\mathbf{S}_1,\dots,\mathbf{S}_L)+t(\mathbf{S}'_1,\dots,\mathbf{S}'_L) \qquad (3)\end{aligned}$$

for $0\le t\le 1$ a scalar. Then $\Psi(\mathbf{S}_1,\dots,\mathbf{S}_L)$ is a convex function of $(\mathbf{S}_1,\dots,\mathbf{S}_L)$ if [2]

$$\frac{d^2}{dt^2}\Psi\left(\bar{\mathbf{S}}_1,\dots,\bar{\mathbf{S}}_L\right)\ge 0 \qquad \forall\bar{\mathbf{S}}_1,\dots,\bar{\mathbf{S}}_L. \qquad (4)$$

This material is based on research supported by the Air Force Research Laboratory under agreement No. F49620-01-1-0372 and by the National Science Foundation under Grant No. CCR-0112501.

[1]The case where each user employs a different n_{st} and n_{sr} is also easy to handle.

Similarly $\Psi(\mathbf{S}_1, \ldots, \mathbf{S}_L)$ is a concave function of $(\mathbf{S}_1, \ldots, \mathbf{S}_L)$ if

$$\frac{d^2}{dt^2}\Psi\left(\bar{\mathbf{S}}_1, \ldots, \bar{\mathbf{S}}_L\right) \leq 0 \qquad \forall \bar{\mathbf{S}}_1, \ldots, \bar{\mathbf{S}}_L. \tag{5}$$

There are several useful known relationships for the derivative of a function of a matrix Φ with respect to a scalar parameter t. In particular we note

$$\frac{d}{dt}\ln\left[det(\Phi)\right] = trace\left[\Phi^{-1}\left(\frac{d}{dt}\Phi\right)\right] \tag{6}$$

and

$$\frac{d}{dt}\Phi^{-1} = -\Phi^{-1}\left(\frac{d}{dt}\Phi\right)\Phi^{-1}. \tag{7}$$

Assuming selection is employed we can use (2), (6) and (7) to find (interchanging a derivative and an expected value)

$$\frac{d}{dt}\Psi\left(\bar{\mathbf{S}}_1, \ldots, \bar{\mathbf{S}}_L\right) = \frac{1}{\ln(2)}\sum_{i=1}^{L} E\left\{trace\left[\mathbf{Q}_i^{-1}\frac{d}{dt}\mathbf{Q}_i\right]\right\} \tag{8}$$

where $\mathbf{Q}_i =$

$$\mathbf{I}_{n_{sr}} + \rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\left(\mathbf{I}_{n_{sr}} + \sum_{j=1,j\neq i}^{L}\eta_{i,j}\tilde{\mathbf{H}}_{i,j}\bar{\mathbf{S}}_j\tilde{\mathbf{H}}_{i,j}^H\right)^{-1} = \mathbf{I}_{n_{sr}} + \rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\tilde{\mathbf{Q}}_i^{-1}, \tag{9}$$

$$\frac{d}{dt}\mathbf{Q}_i = \rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H\tilde{\mathbf{Q}}_i^{-1} - \rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\tilde{\mathbf{Q}}_i^{-1}\left(\frac{d}{dt}\tilde{\mathbf{Q}}_i\right)\tilde{\mathbf{Q}}_i^{-1} \tag{10}$$

and

$$\frac{d}{dt}\tilde{\mathbf{Q}}_i = \sum_{j=1,j\neq i}^{L}\eta_{i,j}\tilde{\mathbf{H}}_{i,j}\mathbf{S}_j'\tilde{\mathbf{H}}_{i,j}^H. \tag{11}$$

A second derivative yields $\frac{d^2}{dt^2}\Psi\left(\bar{\mathbf{S}}_1, \ldots, \bar{\mathbf{S}}_L\right) =$

$$\frac{1}{\ln(2)}\sum_{i=1}^{L}E\left\{trace\left[\mathbf{Q}_i^{-1}\left(\frac{d^2}{dt^2}\mathbf{Q}_i\right) - \mathbf{Q}_i^{-1}\left(\frac{d}{dt}\mathbf{Q}_i\right)\mathbf{Q}_i^{-1}\left(\frac{d}{dt}\mathbf{Q}_i\right)\right]\right\} \tag{12}$$

with

$$\frac{d^2}{dt^2}\mathbf{Q}_i = -2\rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H\tilde{\mathbf{Q}}_i^{-1}\left(\frac{d}{dt}\tilde{\mathbf{Q}}_i\right)\tilde{\mathbf{Q}}_i^{-1} + 2\rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\tilde{\mathbf{Q}}_i^{-1}\left(\frac{d}{dt}\tilde{\mathbf{Q}}_i\right)\tilde{\mathbf{Q}}_i^{-1}\left(\frac{d}{dt}\tilde{\mathbf{Q}}_i\right)\tilde{\mathbf{Q}}_i^{-1}. \tag{13}$$

3. OPTIMUM SIGNALING FOR WEAK INTERFERENCE

We can use (12) to investigate convexity and concavity for any particular set of SNRs $\rho_i, i = 1, \ldots, L$ and INRs $\eta_{i,j}, i, j = 1, \ldots, L, i \neq j$. We focus on extreme cases, weak or strong interference, to gain insight. For the case of very weak interference we ignore terms which are multiples of $\eta_{i,j}$ (essentially we set $\eta_{i,j} \to 0$ for $i = 1, \ldots, L$, $j = 1, \ldots, L$ and $j \neq i$) and we find $\frac{d}{dt}\tilde{\mathbf{Q}}_i = 0$ so that $\frac{d^2}{dt^2}\mathbf{Q}_i = 0$ which leads to $\frac{d^2}{dt^2}\Psi\left(\bar{\mathbf{S}}_1, \ldots, \bar{\mathbf{S}}_L\right) = -\frac{1}{\ln(2)}\times$

$$\sum_{i=1}^{L}E\left\{trace\left[\left(\mathbf{I}_{n_{sr}} + \rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\right)^{-1}\rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H \left(\mathbf{I}_{n_{sr}} + \rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\right)^{-1}\rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H\right]\right\}. \tag{14}$$

Since $\bar{\mathbf{S}}_i$ is a covariance matrix $(\mathbf{I}_{n_{sr}} + \rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H)^{-1} = (\mathbf{U}^H\mathbf{U}+\mathbf{U}^H\Lambda\mathbf{U})^{-1} = (\mathbf{U}(\mathbf{I}+\Lambda)^{-1}\mathbf{U}^H) = \mathbf{U}(\Omega)^2\mathbf{U}^H = \mathbf{U}(\Omega)\mathbf{U}^H\mathbf{U}(\Omega)\mathbf{U}^H$ where $\mathbf{U}$ is unitary and Λ and Ω are diagonal matrices with non-negative entries. Define $\mathbf{A} = \rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H$ and note that $\mathbf{A}^H = \mathbf{A}$ due to $\mathbf{S}_i'$ being a difference of two covariance matrices (easy to see using $\mathbf{U}\Lambda\mathbf{U}^H$ expansion for each covariance matrix). Thus the $trace$ in (14) is written as $trace[\mathbf{U}(\Omega)^2\mathbf{U}^H\mathbf{A}\mathbf{U}(\Omega)^2\mathbf{U}^H\mathbf{A}] = trace[\mathbf{U}\Omega\mathbf{U}^H\mathbf{A}\mathbf{U}(\Omega)^2\mathbf{U}^H\mathbf{A}\mathbf{U}\Omega\mathbf{U}^H] = trace[\mathbf{B}\mathbf{B}^H]$ since $trace[\mathbf{C}\mathbf{D}] = trace[\mathbf{D}\mathbf{C}]$ [3]. We see $trace[\mathbf{B}\mathbf{B}^H]$ must be non-negative since the matrix inside the trace is non-negative definite so that (14) implies $\Psi(\mathbf{S}_1, \ldots, \mathbf{S}_L)$ is concave. This will be true for sufficiently small $\eta_{i,j}, i, j = 1, \ldots, L, i \neq j$ relative to $\rho_i, i = 1, \ldots, L$. Now by a slight extension of the results in [1], based on concavity one can show [4] the optimum $(\mathbf{S}_1, \ldots, \mathbf{S}_L)$ must be of the form such that it is invariant to transforms by permutation matrices. This implies that the best $(\mathbf{S}_1, \ldots, \mathbf{S}_L)$ must be of the form $(\bar{\mathbf{S}}_1, \ldots, \bar{\mathbf{S}}_L) =$

$$\gamma(\alpha\mathbf{I}_{n_{st}}, \ldots, \alpha\mathbf{I}_{n_{st}}) + (1-\gamma)(\alpha\mathbf{O}_{n_{st}}, \ldots, \alpha\mathbf{O}_{n_{st}}) \tag{15}$$

where $\mathbf{O}_{n_{st}}$ is an n_{st} by n_{st} matrix of all ones, $\alpha = \frac{1}{n_{st}}$, and $0 \leq \gamma \leq 1$. We further note that (15) is in exactly the same form as (3) with $(\mathbf{S}_1, \ldots, \mathbf{S}_L) = \alpha(\mathbf{O}_{n_{st}}, \ldots, \mathbf{O}_{n_{st}})$, $(\hat{\mathbf{S}}_1, \ldots, \hat{\mathbf{S}}_L) = \alpha(\mathbf{I}_{n_{st}}, \ldots, \mathbf{I}_{n_{st}})$ and $t = \gamma$.

Weak SNR

Thus we have determined the best signaling except for an unknown single scalar parameter γ which we now investigate. Generally the best approach will change with SNR. First consider the case of weak SNR (recall we are now already focused on very weak, or no, interference). Using the similarity of (15) to (3) just mentioned and (8) with

$(\mathbf{S}_1,\ldots,\mathbf{S}_L) = \frac{1}{n_{st}}(\mathbf{O}_{n_{st}},\ldots,\mathbf{O}_{n_{st}})$, $(\hat{\mathbf{S}}_1,\ldots,\hat{\mathbf{S}}_L) = \frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}})$, $t=\gamma$, $\eta_{i,j} \to 0 \forall i,j$ and $\rho_i \to 0 \forall i$ we find

$$\mathbf{Q}_i^{-1}\frac{d}{dt}\mathbf{Q}_i \to \frac{d}{dt}\mathbf{Q}_i \to \rho_i \tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H$$

to give $\frac{d}{d\gamma}\Psi\left(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L\right) =$

$$\frac{1}{n_{st}\ln(2)}\sum_{i=1}^{L}\rho_i E\Big\{trace\Big[\tilde{\mathbf{H}}_{i,i}\left(\mathbf{I}_{n_{st}}-\mathbf{O}_{n_{st}}\right)\tilde{\mathbf{H}}_{i,i}^H\Big]\Big\} \tag{16}$$

where the $n_{st}\times n_{st}$ matrix can be explicitly written as

$$\mathbf{I}_{n_{st}}-\mathbf{O}_{n_{st}} = \begin{pmatrix} 0 & -1 & \cdots & -1 & -1 & -1 \\ -1 & 0 & -1 & \cdots & -1 & -1 \\ -1 & -1 & 0 & -1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & -1 & 0 \end{pmatrix}. \tag{17}$$

Further simplification of (16) gives

$$\frac{d}{d\gamma}\Psi\left(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L\right) = -\frac{1}{n_{st}\ln(2)}\sum_{p=1}^{L}\rho_p$$
$$E\Big\{\sum_{i=1}^{n_{sr}}\sum_{j=1}^{n_{st}}\sum_{j'=1,j'\neq j}^{n_{st}}\tilde{h}^*(p,p)_{i,j}\tilde{h}(p,p)_{i,j'}\Big\} \tag{18}$$

where $\tilde{h}(p,p)_{i,j}$ denotes the $(i,j)^{th}$ entry of the matrix $\tilde{\mathbf{H}}_{p,p}$. Now notice that without selection ($\tilde{\mathbf{H}}_{p,q}=\mathbf{H}_{p,q}$) the quantity in (18) becomes zero under the assumed model for $\mathbf{H}_{p,q}$ (iid complex Gaussian). Thus selection turns out to be an important aspect in the analysis. In fact, if the selection algorithm turns out to force the term in the second line of (18) to be positive (as we will show next) then we see that (18) is always negative which implies the best solution employs $\gamma=0$ since any increase in γ away from $\gamma=0$ causes a decrease in Ψ.

Consider a matrix $\mathbf{A}$ and let $\lambda_1(\mathbf{A}),\ldots,\lambda_n(\mathbf{A})$ denote the eigenvalues of $\mathbf{A}$. For sufficiently weak SNR ρ_i we can approximate $\ln[\det(\mathbf{I}+\rho_i\mathbf{A})] = \ln[\prod_{i=1}^{n}(1+\rho_i\lambda_i(\mathbf{A}))] = \sum_{i=1}^{n}\ln[1+\rho_i\lambda_i(\mathbf{A})] \approx \rho_i\sum_{i=1}^{n}\lambda_i(\mathbf{A}) = \rho_i trace(\mathbf{A})$. Now consider Ψ itself for the set of covariance matrices in (15) and assume selection is employed. Thus we consider the resulting Ψ as a function of γ and we see $\Psi(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L) =$

$$\frac{1}{n_{st}\ln(2)}\sum_{i=1}^{L}\rho_i E\Big\{trace\Big[\tilde{\mathbf{H}}_{i,i}[\gamma\mathbf{I}_{n_{st}}+(1-\gamma)\mathbf{O}_{n_{st}}]\tilde{\mathbf{H}}_{i,i}^H\Big]\Big\}. \tag{19}$$

Note that the $n_{st}\times n_{st}$ matrix can be explicitly written as $[\gamma\mathbf{I}_{n_{st}}+(1-\gamma)\mathbf{O}_{n_{st}}] =$

$$\begin{pmatrix} 1 & 1-\gamma & \cdots & 1-\gamma & 1-\gamma & 1-\gamma \\ 1-\gamma & 1 & 1-\gamma & \cdots & 1-\gamma & 1-\gamma \\ 1-\gamma & 1-\gamma & 1 & 1-\gamma & \cdots & 1-\gamma \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1-\gamma & 1-\gamma & \cdots & 1-\gamma & 1-\gamma & 1 \end{pmatrix}. \tag{20}$$

Using (20) in (19) gives $\Psi(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L) =$

$$\frac{1}{n_{st}\ln(2)}\sum_{p=1}^{L}\rho_p E\Big\{\sum_{i=1}^{n_{sr}}\sum_{j=1}^{n_{st}}|\tilde{h}(p,p)_{i,j}|^2 +$$
$$(1-\gamma)\sum_{i=1}^{n_{sr}}\sum_{j=1}^{n_{st}}\sum_{j'=1,j'\neq j}^{n_{st}}\tilde{h}^*(p,p)_{i,j}\tilde{h}(p,p)_{i,j'}\Big\}. \tag{21}$$

Thus it is clear from (21) that optimum selection will attempt to make the second term in the expected value as large positive as possible for any $0\leq\gamma<1$. Given this we see (18) is non-positive (note it has the same terms in it) and thus $\gamma=0$ is best for weak SNR cases with small $\rho_i \forall\ i$. Thus the best signaling uses

$$(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L) = \frac{1}{n_{st}}(\mathbf{O}_{n_{st}},\ldots,\mathbf{O}_{n_{st}}). \tag{22}$$

Due to limited space we just outline a justification showing that the optimum approach will use antenna selection to make

$$E\Big\{\sum_{i=1}^{n_{sr}}\sum_{j=1}^{n_{st}}\sum_{j'=1,j'\neq j}^{n_{st}}\tilde{h}^*(p,p)_{i,j}\tilde{h}(p,p)_{i,j'}\Big\}$$

positive. First consider the antenna selection approach which maximizes the ergodic mutual information in (21) when $\gamma=1$. Thus the selection approach will maximize the quantity in (21) with $\gamma=1$ by selecting antennas for each set of instantaneous channel matrices to make the terms inside the expected value in (21) with $\gamma=1$ as large as possible. It is important to note the choice (if $\gamma=1$) depends only on the squared magnitude of elements of the channel matrices.

If we use this selection approach when $\gamma\neq 1$ then the terms multiplying $1-\gamma$ in (21) will be averaged to zero due to the symmetry in the selection criterion. Due to the cross terms in (21) in the term multiplying $(1-\gamma)$ we can use selection to do better by modifying the selection approach. The basic idea is to select antennas to get an additional net positive contribution from the cross terms to increase (21). A more detailed justification is given in [5].

Strong SNR

Now consider the case of strong SNR. To simplify matters, here we assume $n_{st}=n_{rt}$. Then using (8) we find

$$\mathbf{Q}_i^{-1}\frac{d}{dt}\mathbf{Q}_i \to \left(\rho_i\tilde{\mathbf{H}}_{i,i}\bar{\mathbf{S}}_i\tilde{\mathbf{H}}_{i,i}^H\right)^{-1}\rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H,$$

so that $\frac{d}{d\gamma}\Psi\left(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L\right) = \frac{L}{\ln(2)}\times$

$$E\left\{trace\left[\left(\gamma\mathbf{I}_{n_{st}} + (1-\gamma)\mathbf{O}_{n_{st}}\right)^{-1}\left(\mathbf{I}_{n_{st}} - \mathbf{O}_{n_{st}}\right)\right]\right\} = \frac{n_{st}L(n_{st}-1)(\gamma-1)}{\gamma((n_{st}-1)\gamma - n_{st})\ln(2)} \geq 0 \quad (23)$$

which is positive for $0 \leq \gamma < 1$ since $(n_{st}-1)\gamma < n_{st}$ and zero if $\gamma = 1$. In (23) we used $trace[\mathbf{CD}] = trace[\mathbf{DC}]$ [3]. Thus for the strong SNR case $\rho_i \to \infty\ \forall\ i$ when the interference is very weak (and $n_{st} = n_{rt}$) the best signaling uses (15) with $\gamma = 1$. Thus

$$(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L) = \frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}}) \quad (24)$$

is the best signaling in this case.

4. OPTIMUM SIGNALING FOR STRONG INTERFERENCE

Now consider the other extreme of dominating interference where $\eta_{i,j}, i = 1,\ldots,L,\ j = 1,\ldots,L$ is large (compared to $\rho_1,\ldots,\rho_L$). Provided $\eta_{i,j} >> 1$ we can approximate (9) as $\mathbf{Q}_i = \mathbf{I}_{n_{sr}}$. After applying this to (12) and using (13) for large $\eta_{i,j}$ so that $\tilde{\mathbf{Q}}_i^{-1} \approx \left(\sum_{j=1,j\neq i}^{L}\eta_{i,j}\tilde{\mathbf{H}}_{i,j}\bar{\mathbf{S}}_j\tilde{\mathbf{H}}_{i,j}^H\right)^{-1}$ we find the first term inside the trace in (12) depends inversely on $\eta_{i,j}$ while the second term inside the trace in (12) depends inversely on $\eta_{i,j}^2$ so that the first term dominates for large $\eta_{i,j}$. Further, we can interchange the expected value and the trace in (12) so we are concerned with the expected value of (13). Now note that the first term in (13) consists of the product of a term $\mathbf{A} = \tilde{\mathbf{H}}_{i,i}\mathbf{S}_i'\tilde{\mathbf{H}}_{i,i}^H$ and another term depending on $\tilde{\mathbf{H}}_{i,j}$ for $j \neq i$. Now consider the expected value of (13) computed first as an expected value conditioned on $\tilde{\mathbf{H}}_{i,j}$ and then this expected value is averaged over $\tilde{\mathbf{H}}_{i,j}$. Now note that the conditional expected value of $\mathbf{A}$ becomes the zero matrix. Thus the contribution from the first term in (13) averages to zero and the simplified result $\frac{d^2}{dt^2}\Psi\left(\bar{\mathbf{S}}_1,\ldots,\bar{\mathbf{S}}_L\right)$ can be shown to be nonnegative using similar manipulations as for (14). Now using the same permutation argument as used for the weak interference case we can argue that the worst performance is obtained by $(\mathbf{S}_1,\ldots,\mathbf{S}_L) = \gamma\frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}}) + (1-\gamma)\frac{1}{n_{st}}(\mathbf{O}_{n_{st}},\ldots,\mathbf{O}_{n_{st}})$.

Next we attempt to find the exact γ giving worst performance. Towards this goal we first consider the form of $\Psi(\mathbf{S}_1,\ldots,\mathbf{S}_L)$ for $(\mathbf{S}_1,\ldots,\mathbf{S}_L)$ on the line $(\mathbf{S}_1,\ldots,\mathbf{S}_L) = \gamma\frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}}) + (1-\gamma)\frac{1}{n_{st}}(\mathbf{O}_{n_{st}},\ldots,\mathbf{O}_{n_{st}})$ which is (from (2)) $\Psi(\mathbf{S}_1,\ldots,\mathbf{S}_L) =$

$$\begin{aligned}
&\sum_{i=1}^{L} E\left\{log_2\left[\det\left(\mathbf{I}_{n_{sr}} + \rho_i\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i\tilde{\mathbf{H}}_{i,i}^H\mathbf{R}_i^{-1}\right)\right]\right\} \\
&= \sum_{i=1}^{L}\rho_i E\left\{trace\left[\tilde{\mathbf{H}}_{i,i}\mathbf{S}_i\tilde{\mathbf{H}}_{i,i}^H\mathbf{R}_i^{-1}\right]\right\} \\
&= \frac{1}{n_{st}\ln(2)}\sum_{p=1}^{L}\rho_p E\left\{\sum_{i=1}^{n_{sr}}\sum_{j=1}^{n_{st}}|\hat{h}(p,p)_{i,j}|^2\right. \\
&\left. +(1-\gamma)\sum_{i=1}^{n_{sr}}\sum_{j=1}^{n_{st}}\sum_{j'=1,j'\neq j}^{n_{st}}\hat{h}^*(p,p)_{i,j}\hat{h}(p,p)_{i,j'}\right\} \quad (25)
\end{aligned}$$

where the first simplification follows from large $\eta_{i,j}$ and the same simplifications used in (19). The second simplification follows from those in (21) but now $\hat{h}(p,p)_{i,j}$ denotes the $(i,j)^{th}$ entry of the matrix $\mathbf{R}_p^{-\frac{1}{2}}\tilde{\mathbf{H}}_{p,p}$. Now note that antenna selection will attempt to make the second term, which multiplies the positive constant $1-\gamma$, as large positive as it possibly can. Thus we see that best performance on the line $(\mathbf{S}_1,\ldots,\mathbf{S}_L) = \gamma\frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}}) + (1-\gamma)\frac{1}{n_{st}}(\mathbf{O}_{n_{st}},\ldots,\mathbf{O}_{n_{st}})$ must be obtained for $\gamma = 0$ and the the worst performance must occur at $\gamma = 1$ or $(\mathbf{S}_1,\ldots,\mathbf{S}_L) = \frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}})$. Thus the overall worst signaling in this case uses $(\mathbf{S}_1,\ldots,\mathbf{S}_L) = \frac{1}{n_{st}}(\mathbf{I}_{n_{st}},\ldots,\mathbf{I}_{n_{st}})$. So the best signaling for cases without interference and selection is the worst for strong interference and selection.

5. REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian Channels," European Transactions on Telecommunications, 10(6): 585-595, Nov.-Dec. 1999.

[2] Stephen Boyd and Lieven Vandenberghe, Convex Optimization, (currently a set of notes for classes at Stanford and UCLA but will be published in book form in 2003).

[3] Harry L. Van Trees, Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory, John Wiley & Sons, New York, 2002 (especially Appendix A, pp. 1342, 1345, 1349, 1351, 1359, 1401).

[4] P. J. Voltz, "Characterization of the optimum correlation matrix for MIMO with antenna subset selection", submitted to IEEE Trans. on Comm.

[5] R. S. Blum, "MIMO Capacity with Antenna Selection and Interference", in review.