Sub-optimal Power Allocation for MIMO Channels

Alex J. Grant
Institute for Telecommunications Research,
University of South Australia, Australia
alex.grant@unisa.edu.au

Leif W. Hanlen
Wireless Signal Processing Program
National ICT Australia
leif.hanlen@nicta.com.au

***Abstract*— We consider t-input r-output Rayleigh fading channels with transmit-sided correlation, where the receiver knows the channel realizations, and the transmitter only knows the channel statistics. Using Lagrange duality, we develop an easily computable, tight upper bound on the loss in information rate due to the use of any given input covariance for this channel. This bound is applied to two simple transmission strategies. The first strategy is a reduced-rank uniform allocation, in which independent, equal power Gaussian symbols are transmitted on the αt strongest eigenvectors of the transmit covariance matrix, where $0 \leq \alpha \leq 1$ is chosen to optimize the resulting information rate. The second strategy is water-filling on the eigenvalues of the transmit covariance matrix. The upper bound on loss shows these strategies are nearly optimal for a wide range of signal to noise ratios and correlation scenarios.**

I. Introduction

Since [1, 2], information theoretic results for t-input r-output fading channels have been extended to channel models of increasing generality. Under various assumptions on the correlation structure of channel gain matrix, several authors have sought expressions for the input-output mutual information. The resulting expressions can be unwieldy red and optimal input distributions may be difficult to find.

This paper is concerned with ergodic zero-mean Rayleigh channels with single-sided correlation at the transmitter. It is assumed throughout that the receiver knows the channel realizations, while the transmitter knows only the channel statistics. Under these assumptions, optimal signaling consists of independent transmission on each eigenvector of the transmit covariance matrix. The problem is to determine the appropriate power allocations. This was addressed in [3, 4], where iterative algorithms for numerical computation of the optimal covariance matrix have been given. The associated computational complexity may be prohibitive for online implementation; motivating simple, near optimal power allocations.

One possibility is uniform power allocation, which is asymptotically optimal for high signal-to-noise ratio (SNR). Uniform power allocation has also been shown to be optimal in a game-theoretic sense [5] when the channel statistics are unknown at the transmitter. At low SNR, it is known that beam-forming (rank-one) transmission is optimal, the transmitter using only the strongest eigenvector of the transmit covariance matrix.

Over a wide range of SNR, however, optimal power allocation is non-trivial. It is interesting to consider the idea of reduced-rank uniform power allocation, in which independent, equal power Gaussian symbols are transmitted on the αt, strongest eigenvectors of the transmit covariance matrix with $0 \leq \alpha \leq 1$ chosen to optimize the resulting rate.

This idea is similar to the strategy proposed in [6]. There, it is assumed the eigenvectors of the channel *realization* (or approximations to them) are known. The transmitter applies a uniform power allocation over a subset of these *orthogonal* channels. In [6] some large systems asymptotic results are given, and they note that the strategy is nearly optimal for a wide range of SNR. Our focus is on the case where only the *statistics* of the channel are known at the transmitter.

Another possibility for an easily computable power allocation is to simply perform water-pouring on the eigenvalues of the transmit covariance matrix. At low and high SNR this strategy is optimal (at low SNR, it reduces to beam-forming, and at high SNR it reduces to uniform allocation).

The remainder of the paper is organized as follows. We give the system model in Section II. In Section III, we derive an easily computable, tight bound on the rate loss that results by using any given power allocation. The bound is based on the Lagrange dual of the mutual information, and avoids the need to compute the channel capacity. In Section IV, examples of reduced rank uniform and water-pouring allocations are considered for finite dimension systems. We apply our bound to show that both strategies perform well. Section V considers large systems asymptotics. We show how to pose a convex optimization problem to compute the asymptotically optimal power allocation. This differs from previous results in that it avoids the requirement of the solution of fixed point equations. We show that for the large dimension, low SNR regime, water-pouring on the transmit covariance eigenvalue spectrum is optimal, and is never more than 1 nat away from the capacity of the channel with perfect channel knowledge at the transmitter.

II. System Model and Problem Formulation

Transmission is over a t-input, r-output additive white Gaussian noise channel of the form

$$y[k] = \sqrt{\gamma} H[k] x[k] + z[k] \tag{1}$$

where $y[k] \in \mathbb{C}^{r \times 1}$ is a complex column vector of matched filter outputs at symbol time $k = 1, 2, \ldots, N$ and $H[k] \in \mathbb{C}^{r \times t}$ is the corresponding matrix of complex channel coefficients. The element at row i and column j of $H[k]$ is the complex channel coefficient from transmit element j to receive element

i. The channel matrices $H[k]$ will be selected independently at each symbol according to a correlated Gaussian matrix density. It is assumed that $H[i]$ is known at the receiver, but that only the statistics of H are known at the transmitter.

The vector $x[k] \in \mathbb{C}^{t\times 1}$ is the vector of complex baseband input signals, and $z[k] \in \mathbb{C}^{r\times 1}$ is a complex, circularly symmetric Gaussian vector with $\mathrm{E}\left\{n[k]n[k]^\dagger\right\} = I_r$. The superscript $(\cdot)^\dagger$ means Hermitian adjoint and I_r is the $r \times r$ identity matrix. Let $n = \max(t, r)$ and $m = \min(t, r)$.

A transmitter power constraint $E[\|x[k]\|_2^2] \leq 1$ is enforced and the signal-to-noise ratio is defined as γ. The input covariance matrix is defined as the $t \times t$ matrix

$$Q = \mathrm{E}\left\{x[k]x[k]^\dagger\right\} \tag{2}$$

and hence the power constraint may be written as $\mathrm{tr}(Q) \leq 1$.

The assumptions of additive Gaussian noise and receiver knowledge of $H[k]$ mean that the optimal input density is Gaussian and determination of capacity corresponds to optimizing Q. Defining

$$\Psi(Q) = I(x; y|H) = \mathrm{E}\left\{\log\det\left(I + \gamma HQH^\dagger\right)\right\}, \tag{3}$$

the capacity is the solution of $C = \max_Q \Psi(Q)$ subject to $\mathrm{tr}(Q) \leq 1$ and $Q \succeq 0$. The resulting optimal Q can depend on the channel statistics, but not on the channel realizations.

For $H[k]$ with i.i.d. Gaussian entries, Telatar showed that the optimizing $Q = I_t/t$, meaning that it is optimal to transmit independently with equal power from each antenna. Telatar also gave an expression for computation of $\Psi(I)$, and several other expressions have subsequently been found [4, 7, 8].

Let $\mathcal{N}_{t,r}(M, \Sigma)$ be a multivariate Gaussian density with $r \times t$ mean matrix M and $rt \times rt$ covariance matrix $\Sigma = \mathrm{E}\left\{hh^\dagger\right\}$ where h is formed by stacking the columns of the matrix H into a single column vector. Common special cases include i.i.d. unit variance entries, $\mathcal{N}_{t,r}(0, I)$ (corresponding to independent Rayleigh fading) as considered in [1] and the so-called Kronecker correlation model $\mathcal{N}_{t,r}(M, R \otimes T)$. The latter model corresponds to separable transmit T and receive correlation R, and may be generated via $M + R^{1/2}GT^{1/2}$ where $G \sim \mathcal{N}_{t,r}(0, I)$, where M corresponds to a "line-of-sight" channel component. "One-ring" models of with single-ended correlation structure $H \sim \mathcal{N}_{t,r}(0, I \otimes T)$ have been considered in [9]. In [10] it was shown that for $H \sim \mathcal{N}_{t,r}(0, I \otimes T)$ it is optimal to transmit independently on the eigenvectors of T.

Closed form solutions have been obtained for the mutual information of single-ended correlated channels [4, 11] and for $H \sim \mathcal{N}_{t,r}(0, R \otimes T)$, [12, 13]. Kuhn-Tucker optimality conditions for $H \sim \mathcal{N}_{t,r}(0, R \otimes T)$ were developed in [4] along with an iterative procedure for numerical optimization of Q. This work was extended to arbitrary (non-Gaussian, non-separable correlation) channel densities in [14]. Majorization results have also been obtained showing that stronger modes should be allocated higher powers [10]. This majorization has motivated water-filling using transmit covariance eigenvalues [15, 16].

Asymptotic large systems ($r, t \to \infty$ with $r/t \to$ a constant) capacity results have been obtained in [17], for the case $H \sim \mathcal{N}_{t,r}(0, R \otimes T)$, but under the assumption $Q = I/t$. Asymptotic results for arbitrary Q were considered in [18], where the asymptotic distribution of the mutual information was found to be normal. Large-systems results have been obtained in [3], concentrating on the case where the eigenvectors of the optimal Q can be identified by inspection.

This paper mainly focuses on the case $H[k] \sim \mathcal{N}_{t,r}(0, I \otimes T)$ and unless stated otherwise we assume $T = \mathrm{diag}(\tau_1, \ldots, \tau_t)$, and $Q = \mathrm{diag}(q_1, \ldots, q_t)$. According to our assumptions, the transmitter has knowledge of T, and hence diagonal matrices can be assumed since the transmitter can work in the eigenspace of T.

We are interested in two particular (sub-optimal) choices for Q.

- The reduced rank uniform strategy has $q_i = 1/k$, $i \leq k$ and $q_i = 0$, $i > k$ for some integer choice of $0 < k \leq t$.
- The water filling strategy has $q_i = \max(0, \nu - 1/\tau_i)$ where the constant ν is chosen so that $\mathrm{tr}(Q) = 1$.

We wish to bound the loss in mutual information due to these choices.

III. Capacity Bounds via Lagrange Duality

We begin with a general result. All proofs are contained in the Appendix.

Theorem 1: Let $I(p)$ be the mutual information of an ergodic channel whose input density is a differentiable convex function of n non-negative transmitter parameters $p_1, p_2, \ldots, p_n$. Suppose these parameters must satisfy $\sum p_i \leq \gamma$ for a given constant γ. Let C be the channel capacity, defined as the maximum of $I(p)$ over all p satisfying the constraints. Then for a specific set of input parameters $p = q$ with $\sum q_i = \gamma$,

$$C - I(q) \leq \max_j \gamma I_j(q) - \sum_j q_j I_j(q), \text{ where}$$
$$I_j(q) = \partial I(p)/\partial p_j\big|_{p=q}$$

Theorem 1 may be of independent interest, but is not much more than a direct application of Lagrange duality. The utility of the theorem comes from the fact that although it may be hard to solve $\partial L(p, \lambda)/\partial p_i = 0$ for p, it is easy to solve for λ. Furthermore, we are able to bound the gap to capacity, without having to find the optimal parameters. The mutual information in Theorem 1 can depend on the parameters $q_1, q_2, \ldots, q_n$ in any convex way. These parameters could be probabilities[1], or in the case of immediate interest, the eigenvalues of the input covariance matrix.

Theorem 2: Let $\Psi(Q)$ be the mutual information (3) of the ergodic channel (1) with $H[k] \sim p_H$ and input covariance matrix $Q = diag(q_1, \ldots, q_t)$. The gap to capacity is bounded

$$C - \Psi(Q) \leq \gamma \max_i M_{ii} - \mathrm{tr}(QM) = \delta$$

[1]In which case we recover something in the spirit of the discussion following [19, Theorem 4.5.1], with the benefit that we get an actual bound on the loss due to the use of suboptimal input distribution q.

where $M = \mathrm{E}_H\{(I + H^\dagger HQ)^{-1} H^\dagger H\}$.

In the case $H[k] \sim \mathcal{N}_{t,r}(0, I \otimes T)$, we can obtain a closed form expression for M, using the determinant representation of $I(Q)$ developed in [4]. Let $a_i = \tau_i q_i$ and $t = r$ (for clarity only, other values can easily be accommodated). Then

$$\Psi(Q) = \frac{1}{t|V|} \sum_{k=1}^{t} |A^{(k)}| \tag{4}$$

where the $t \times t$ matrix V has elements $V_{ij} = a_j^{1-i}$ and the matrix $A^{(k)}$ has elements

$$A_{ij}^{(k)} = \begin{cases} \dfrac{a_j^{i-t}}{\Gamma(k)} \displaystyle\int_0^\infty \log(1 + a_j\lambda)\lambda^{k-1} e^{-\lambda}\, d\lambda & i = k \\ a_j^{i-t} & i \neq k \end{cases}$$

Theorem 3: Let $\Psi(Q)$ be the mutual information (3) of the ergodic channel (1) with $H[k] \sim \mathcal{N}_{t,t}(0, I \otimes T)$ and input covariance matrix $Q = diag(q_1, \dots, q_t)$. The gap to capacity is bounded

$$C - \Psi(Q) \leq \gamma \max_i \psi_i - \sum_i q_i \psi_i.$$

where $\psi_i = d\Psi/dq_i$ is given by

$$\psi_i = \frac{\tau_i}{t|V|} \left(|V| \sum_k |A_i^{(k)}| - |V_i| \sum_k |A^{(k)}| \right)$$

The matrices V_i, $A_i^{(k)}$ are formed from V and $A^{(k)}$ respectively via differentiation of column i with respect to $a_i = \tau_i q_i$ (other columns are unchanged).

Although ψ_i is easy to compute, the solution of $\psi_i = \nu$ for a constant ν (required for optimality) is intractable.

IV. Finite Systems

We give examples to demonstrate application of the duality bound to both the reduced rank uniform and water filling strategies. Selection of the best number of non-zero modes for the reduced rank strategy is easy via (4).

Example 1: Consider the single-ended correlated Rayleigh channel with $t = r = 2$ and $T = \mathrm{diag}(\tau, 2 - \tau)$. We consider the case $1 \leq \tau \leq 2$, since the result is symmetric around $\tau = 1$. Figures 1 (for $\gamma = 0$ dB) and 2 (for $\gamma = 15$ dB) show the bound on the loss δ, normalized by the actual capacity[2] (obtained via the iterative procedure of [4]). The dashed portion of the line marked 2 corresponds to rank-2 uniform power allocation. The dotted line marked 1 is for rank 1 (beam-forming). The solid line marked WF corresponds to water filling on the eigenvalues of T.

Water filling on T is superior to reduced rank uniform allocation, with worst case loss of about 6% at 0 dB and less than 2% for 15 dB. The reduced rank strategy may lose up to 10%. Numerical investigations show that as the SNR increases, the relative loss due to water filling on T continues to decrease, as expected. The loss due to the water filling strategy is never more than 1 nat.

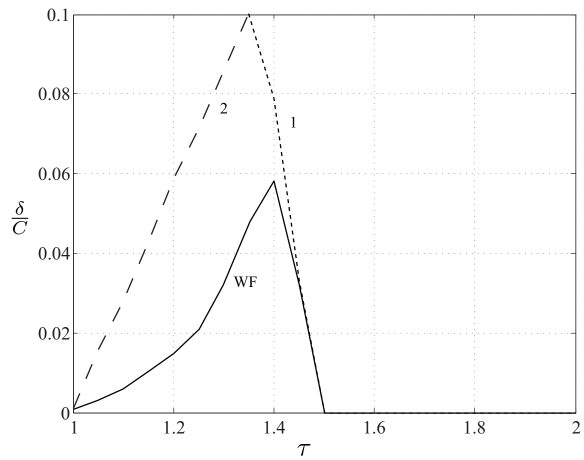

Fig. 1. Relative loss, SNR 0 dB

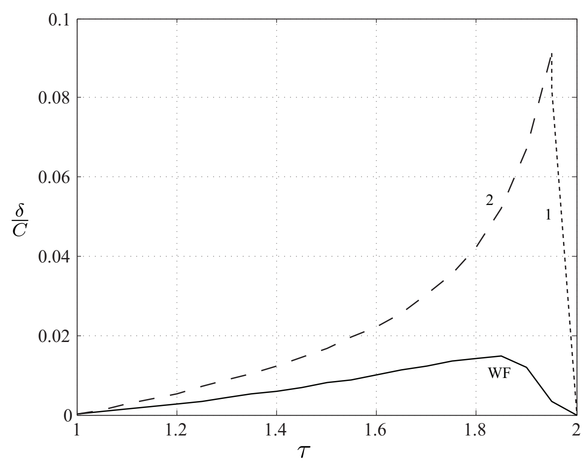

Fig. 2. Relative loss, SNR 15 dB

Example 2: Let $r = t = 10$ and $T = \alpha\,\mathrm{diag}(10, 9, \dots, 1)$ with α chosen so that $\mathrm{tr}(T) = t$. Figure 3 shows the bound on loss (relative to the mutual information Ψ) versus SNR, resulting from the reduced rank uniform strategy.

V. Asymptotic Analysis

Large systems analysis is a powerful technique for analysis of linear matrix channels. The large systems assumption is $r, t \to \infty$ with $t/r \to \beta$, a constant. This asymptotic regime allows convenient application of the theory of large random matrices. Typically convergence to asymptotic values is quite fast, e.g. $r, t > 10$. In this section we consider the optimization of α in the large systems limit.

Suppose in the large systems limit, the empirical eigenvalue distribution of T/t converges to a given density $t(\lambda)$. Without loss of generality, we can assume that $T = \mathrm{diag}(\tau_1, \dots, \tau_t)$ is diagonal, and that in the limit, the τ_i are chosen i.i.d. according to $t(\lambda)$. The corresponding input covariance matrix Q will be diagonal. The matrix Q can be specified by a transmit power allocation function $q(\lambda)$, which specifies how much power is transmitted on an eigenvector whose eigenvalue is λ.

[2]Obviously in this case, we could compute the exact loss. The objective here is to determine the usefulness of the duality bound.

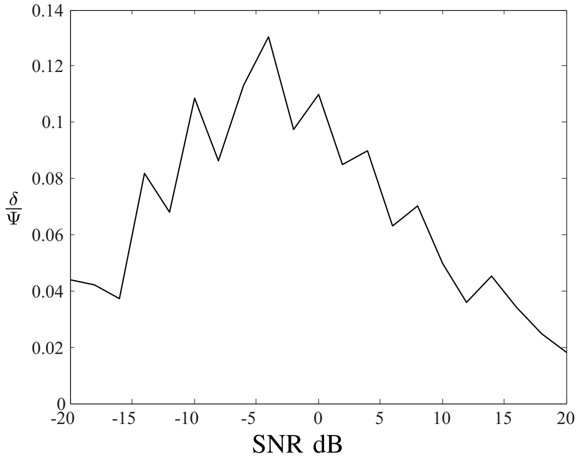

Fig. 3. Relative loss for Example 2.

Under these conditions, the objective function is

$$\frac{\Psi}{t} \to \eta_f(\gamma) = \int \log(1+\gamma\lambda) f(\lambda)\, d\lambda$$

where $f(\lambda)$ is the limiting eigenvalue distribution of $GTQG^\dagger/t$, where $G \sim \mathcal{N}_{t,r}(0, I)$ (this eigenvalue density exists under the assumptions that we have made). Let $\xi = h(\lambda) = \lambda q(\lambda)$. The eigenvalue density of TQ is therefore $g(\xi) = t\left(h^{-1}(\xi)\right) dh^{-1}(\xi)/d\xi$.

The following parametric relation shows how to compute η_f in terms of the eigenvalue density g (which is easy to find), rather than f, which is hard to find.

$$\eta_f(\gamma) = \frac{1}{\beta}\eta_g(s) - \ln\frac{s}{\gamma} + \frac{s}{\gamma} - 1 \tag{5}$$

$$\gamma(s) = s\left(1 - \frac{s}{\beta}\eta_g'(s)\right)^{-1}. \tag{6}$$

This avoids working with Stieltjes transforms. It is easily verified that (5)-(6) is the same as

$$\eta_f(\gamma) = \frac{1}{\beta}\zeta_q(s) - \ln\frac{s}{\gamma} + \frac{s}{\gamma} - 1 \tag{7}$$

$$\gamma(s) = s\left(1 - \frac{s}{\beta}\zeta_q'(s)\right)^{-1}, \tag{8}$$

$$\zeta_q(s) = \int \log\left(1 + s\lambda q(\lambda)\right) t(\lambda)\, d\lambda. \tag{9}$$

The large-systems optimization problem is therefore

$$\max_{s,q} \frac{1}{\beta}\zeta_q(s) - \ln\frac{s}{\gamma} + \frac{s}{\gamma} - 1 \tag{10}$$

$$\text{subject to } q(\lambda) \geq 0 \tag{11}$$

$$s \geq 0 \tag{12}$$

$$\gamma - s \geq 0 \tag{13}$$

$$\int q(\lambda)t(\lambda)\, d\lambda - 1 = 0 \tag{14}$$

$$s\left(1/\gamma + \zeta_q'(s)/\beta\right) - 1 = 0 \tag{15}$$

Note that (8) is enforced in the constraint (15), and that (14) enforces the transmit power constraint $\operatorname{tr}(Q) \leq 1$. It can be verified that the objective function is convex and that the constraints define a convex set, and standard numerical optimization procedures can be applied.

Retaining the diagonal assumption on T, let

$$z[k] = \sqrt{\gamma}\, Tx[k] + w[k], \tag{16}$$

where w is white Gaussian noise with $\mathrm{E}\{ww^\dagger\} = I_t$. This channel consists of t parallel channels, and

$$C_\perp = \max_Q I(x; z \mid T)$$

is achieved by water-pouring on the diagonal entries of T. In the large systems limit $C_\perp = \zeta_p(\gamma)$, where

$$p(\lambda) = \max\left(0, \xi - \frac{1}{\lambda}\right) \tag{17}$$

$$\int_{1/\xi}^{\infty} p(\lambda)t(\lambda)\, d\lambda = 1. \tag{18}$$

It is interesting to consider the loss $\Delta = C_\perp - C$ due to the non-orthogonality caused by H. Now $I(x; y \mid T, H) \leq I(x; z \mid T) \leq C_\perp$ and hence for any particular input covariance Q, $\Psi(Q) \leq C \leq C_\perp$ (note that (16) is the channel that results from perfect transmitter knowledge of H). Thus the loss Δ can be upper bounded by $\Delta(Q) = C_\perp - \Psi(Q)$ for any Q. In the large systems limit, we will write $\Delta(q)$, where $q(\lambda, \gamma)$ defines Q.

Theorem 4: In the large systems regime with $\beta = 1$, $0 \leq C_\perp - C \leq \Delta(p, \gamma) \leq 1$ nat, where $\Delta(p, \gamma)$ is monotonic in γ, $\Delta(p, 0) = 0$ and $\Delta(p, \infty) = 1$.

Thus the loss in capacity due to the non-orthogonality introduced by H is bounded by 1 nat, and tends to zero at low SNR. Furthermore, since the slope of Δ is zero at $\gamma = 0$, this optimality is also stable at low SNR.

Corollary 1: With the same conditions of Theorem 4, $0 \leq C - \Psi(p) \leq \Delta(p, \gamma) \leq 1$ nat.

This corollary shows that the water pouring power allocation is not too bad. The duality bounds of Section III can be readily extended to the parametric formulation of (7)-(8).

VI. Conclusion

We have considered two simple transmission strategies for the MIMO Rayleigh fading channel with single-ended correlation and statistical transmitter channel state information. Using a simple upper bound on capacity, obtained from the Lagrange dual function, we have shown that near-optimal performance can be obtained over a wide range of transmit power and channel correlations.

Appendix

Proof: [Theorem 1] Let $J(q) = -I(q)$. Capacity is the solution to the following optimization problem.

$$\min_x J(p) \quad \text{subject to}$$

$$f_0(p) = \sum_i p_i - \gamma \leq 0$$

$$f_i(p) = -p_i \leq 0, \quad i = 1, 2, \ldots, n.$$

The corresponding Lagrangian is

$$L(p,\lambda) = J(x) + \sum_{i=0}^{n} \lambda_i f_i(p)$$

and the Lagrange dual is $g(\lambda) = \min_p L(p,\lambda)$ The Lagrange dual provides an lower bound to the objective function $g(\lambda) \leq J(p)$ for all p, λ. The bound is tightened by further maximizing over the Lagrange multiplier λ. In cases where the duality gap is zero $-C = \max_\lambda \min_p L(p,\lambda)$. Our goal however is to bound the loss due to the use of a sub-optimal choice of p.

We determine a particular feasible value μ of the Lagrange multiplier that results in $g(\mu) = L(q,\mu)$, i.e. with $\lambda = \mu$, the Lagrangian is minimized by $p = q$. With these choices,

$$C - I(q) \leq \delta = -\mu_0 \left(\sum_i q_i - \gamma \right) + \sum_{i=1}^{n} \mu_i q_i \tag{19}$$

$$= \sum_{i=1}^{n} \mu_i q_i. \tag{20}$$

It remains to find μ. The Lagrangian is sufficiently well behaved so that μ satisfies

$$0 = \partial L(p,\mu)/\partial p_i|_{p=q} = -I_i(q) + \mu_0 - \mu_i.$$

Hence we require $\mu_i = \mu_0 - I_i(p)$. Furthermore, for μ to be dual feasible, $\mu_i \geq 0$, $i = 0, 1, \ldots, n$. This is achieved with $\mu_0 \geq \max_i I_i(p)$, and choosing the largest such μ_0 minimizes δ. Substituting into (20) completes the proof. Note that whenever the duality gap is zero, the bound is tight i.e. $\delta = 0 \iff q$ is optimal. ■

Proof: [Theorem 2] Application of Theorem 1, using [4] $\partial\Psi(Q)/\partial q_i = \mathrm{E}\left\{ \left((I + H^\dagger H Q)^{-1} H^\dagger H \right)_{ii} \right\}$. ■

Proof: [Theorem 3] Differentiation of (4) with respect to q_m is straightforward, if tedious. The useful property is that both $A^{(k)}$ and V depend on q_m only in column m, which means that differentiation of the corresponding determinants becomes differentiation of column m (via the multi-linearity of determinant). ■

Proof: [Theorem 4] Consider the large systems limit, and compute $\Delta(p)$ for $p(\lambda)$ defined by (17), (18).

$$\Delta(p,\gamma) = \zeta_p(\gamma) - \zeta_p(s) + \log\frac{s}{\gamma} - \frac{s}{\gamma} + 1$$
$$= \int t(\lambda) \log \frac{1+\gamma\lambda p(\lambda)}{1+s\lambda p(\lambda)}\, d\lambda + \log\frac{s}{\gamma} - \frac{s}{\gamma} + 1.$$

We shall first evaluate the high and low SNR limits of $\Delta(p)$. The proof will be complete by showing that $\frac{d\Delta}{d\gamma} \geq 0$ for $\gamma \geq 0$. This will ensure that $\Delta(p)$ is bounded by these limiting values.

By Taylor series expansion,

$$\gamma(s) = s + s^2 \int \lambda p(\lambda) t(\lambda) d\lambda + O(s^3).$$

Therefore, at low SNR, both $\gamma \to 0$ and $s \to 0$ such that $s/\gamma \to 1$ from below, and $\lim_{\gamma\to 0} \Delta(p,\gamma) = 0$. At high SNR, both $\gamma \to \infty$ and $s \to \infty$, with the ratio $s/\gamma \to 0$. Hence $\lim_{\gamma\to\infty} \Delta(p,\gamma) = 1$. It remains to show that $d\Delta/d\gamma \geq 0$. Now

$$\frac{d\Delta}{d\gamma} = \int \lambda t(\lambda) p(\lambda) \frac{\gamma - s}{\gamma(1+\gamma\lambda p(\lambda)(1+s\lambda p(\lambda))} d\lambda$$
$$\geq 0$$

since s, γ, λ, $t(\lambda)$, $p(\lambda)$ are all non-negative and $\gamma \geq s$. It can also be seen that $\frac{d\Delta}{d\gamma} \to 0$ for both low and high SNR. ■

References

[1] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Euro. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[2] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[3] A. M. Tulino, A. Lozano, and S. Verdu, “Impact of antenna correlation on the capacity of multiantenna channels,” *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2491–2509, July 2005.

[4] L. W. Hanlen and A. J. Grant, “Capacity analysis of correlated MIMO channels,” Sept. 2004, submitted to IEEE Trans. Inform. Theory.

[5] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, “Uniform power allocation in MIMO channels: a game-theoretic approach,” *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1707–1727, July 2003.

[6] W. Dai, Y. Liu, B. Rider, and V. K. N. Lau, “On the information rate of MIMO systems with finite rate channel state feedback and power on/off strategy,” in *IEEE Intl. Symp. Inform. Theory, ISIT*, vol. 1, 2005, pp. 1549–1553.

[7] A. J. Grant, “Rayleigh fading multiple-antenna channels,” *EURASIP J. App. Sig. Proc., Special Issue on Space-Time Coding, Part I*, no. 3, pp. 316 – 329, Mar. 2002.

[8] E. Biglieri and G. Taricco, “Transmission and reception with multiple antennas: Theoretical foundations,” *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 2, pp. 183–332, 2004.

[9] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, “Fading correlation and its effect on the capacity of multielement antenna systems,” *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[10] S. A. Jafar, S. Vishwanath, and A. Goldsmith, “Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback,” in *IEEE Intl. Conf. Commun., ICC*, vol. 7, June 11–14 2001, pp. 2266–2271.

[11] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdú, “Capacity of MIMO channels with one-sided correlation,” in *IEEE Intl. Symp. Spread Spectrum Techniques and Applications ISSSTA*, 2004, pp. 515–519.

[12] M. Kiessling and J. Speidel, “Exact ergodic capacity of MIMO channels in correlated Rayleigh fading environments,” *Int. Zurich Seminar*, Feb. 2004.

[13] S. H. Simon and A. L. Moustakas, “Eigenvalue density of correlated complex random Wishart matrices,” *Phys E. Review*, vol. 69, pp. 065 101–1 – 065 101–4, June 11 2004.

[14] L. W. Hanlen and A. J. Grant, “Optimal transmit covariance for MIMO channels with statistical transmitter side information,” in *IEEE Intl. Symp. Inform. Theory, ISIT*, Adelaide, Australia, Sept. 4–9 2005.

[15] S. H. Simon and A. L. Moustakas, “Optimizing MIMO antenna systems with channel covariance feedback,” *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 406–417, Apr. 2003.

[16] S. K. Jayaweera and H. V. Poor, “On the capacity of multiple-antenna systems in Rician fading,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, May 2005.

[17] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, “Capacity scaling in MIMO systems under correlated fading,” *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[18] C. Martin and B. Ottersten, “Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 13 500–1359, jul 2004.

[19] R. G. Gallager, *Information Theory and Reliable Communication*. New York, USA: John Wiley & Sons, 1968.

National ICT Australia is funded through the Australian Government's *Backing Australia's Ability initiative*, in part through the Australian Research Council.