MIMO Capacity with Per-Antenna Power Constraint

Mai Vu
Department of Electrical and Computer Engineering, McGill University, Montreal, H3A2A7
Email: mai.h.vu@mcgill.ca

***Abstract*—In this paper, we consider the single-user MIMO channel with per-antenna power constraint. We formulate the capacity optimization problem with per-antenna constraint in the SDP framework and analyze its optimality conditions. We establish in closed-form the optimal input covariance matrix as a function of the dual variable. We then propose a simple algorithm to find this optimal input covariance and the capacity. Results show that the capacity with per-antenna power can be significantly different from that with sum power or with independent multiple access constraint.**

I. Introduction

The capacity of a MIMO wireless channel depends on the constraints on the transmit power and on the availability of the channel state information at the transmitter and receiver. With sum power constraint across all transmit antennas, the capacity and optimal signaling are well established. For channels known at both the transmitter and the receiver, the capacity can be obtained by performing singular value decomposition of the channel and water-filling power allocation on the channel eigenvalues [1].

Under the per-antenna power constraint, the MIMO capacity is less well understood. This per-antenna power constraint, however, is more realistic than sum power in practice because of the constraint on the individual RF chain connected to each antenna. Hence the transmitter may not be able to allocate power arbitrarily among the transmit antennas. Another appealing scenario for the per-antenna constraint is a distributed MIMO system, which has the transmitted antennas located at different physical nodes that cannot share power. Thus understanding the capacity and the optimal signaling scheme under per-antenna power constraint can be useful.

The per-antenna power constraint has been investigated in different problem setups. In [2], the problem of a multiuser downlink channel is considered with per-antenna power. For downlink broadcast, the capacity optimization problem is non-convex. It was argued that linear processing at both the transmitter (by multi-mode beamforming) and the receiver (by MMSE receive beamforming with successive interference cancellation) can achieve the capacity region. Using uplink-downlink duality, the boundary points of the capacity region for the downlink channel with per-antenna constraint can be found by solving a dual uplink problem, which maximizes a weighted sum rate for the uplink channel with sum power constraint across the users and an uncertain noise. The dual uplink problem is convex which facilitates computation. Similar uplink-downlink duality holds with the sum power constraint, in which case, the uplink capacity can be solved using the efficient and distributed iterative water-filling algorithm [3]. This algorithm is based on the closed-form solution for capacity of a *single-user* MIMO channel with sum power constraint. In the case of per-antenna power constraint, however, no closed-form capacity solution so far exists, the lack of which implies full and centralized computation required to solve even the convex uplink capacity problem.

In this paper, we consider the single-user MIMO channel with per-antenna power constraint. The capacity optimization problem with per-antenna power is convex and can be cast in the SDP framework. We analyze its optimality conditions and establish in closed-form the optimal input covariance matrix as a function of the dual variable. We then propose a simple algorithm to find this optimal input covariance and the capacity. Simulation examples show the fast convergence of the proposed algorithm and that capacity with per-antenna power can be significantly different from that with sum power.

For notation, we use bold face lower-case letters for vectors, capital letters for matrices, $(.)^T$ for transpose, $(.)^*$ for conjugate, $(.)^\dagger$ for conjugate transpose, $\succcurlyeq$ and $\preccurlyeq$ for matrix inequalities (positive semi-definite relation), tr$(.)$ for trace, and diag$\{.\}$ for forming a diagonal matrix with the specified elements.

II. Channel Model and Power Constraints

A. Channel model

Consider a multiple-input multiple-output (MIMO) channel with n transmit and m receive antennas. The channel between each transmit-receive pair is a complex, multiplicative factor h_{ij}. Denote the channel coefficient matrix as $\mathbf{H}$ of size $m \times n$, and the transmit signal vector as $\mathbf{x} = [x_1 \dots x_n]^T$. Then the received signal vector of length m can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{1}$$

where $\mathbf{z} \sim \mathcal{CN}(0, \mathbf{I})$ is a vector of additive white circularly complex Gaussian noise. Here we have normalized the noise power at all receivers, which can be done by absorbing the actual noise power into the transmit power constraint.

Assume the channel $\mathbf{H}$ is known at both the transmitter and receiver. The capacity of the MIMO channel depends on the power constraint on the input signal vector $\mathbf{x}$. In all cases, because of the Gaussian noise and known channel at the receiver, the optimal input signal is Gaussian with zero mean [4]. Let $\mathbf{Q} = E[\mathbf{x}\mathbf{x}^\dagger]$ be the covariance of the Gaussian input, then the achievable transmission rate is

$$r = \log\det\left(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right). \tag{2}$$

The remaining question is to establish the optimal $\mathbf{Q}$ that maximizes this rate according to the CSIT condition and a given power constraint.

B. Power constraints

Often the MIMO capacity is studied with sum power constraint across all antennas. In this paper, we consider a more realistic per-antenna power constraint. For comparison, we also include the case of independent multiple-access power constraint. We elaborate on each power constraint below.

1) Sum power constraint: With sum power constraint, the total transmit power from all n antennas is P, but this power can be shared or allocated arbitrarily among the transmit antennas. This constraint translates to a condition on the input covariance as

$$\text{tr}(\mathbf{Q}) \leq P. \tag{3}$$

This constraint allows complete cooperation among the transmit antennas.

2) Independent multiple-access power constraint: In this case, each transmit antenna has its own power budget and acts independently. This constraint can model the case of distributed transmit antennas, such as on different wireless nodes scattered in a field, without explicit cooperation among them. Let P_i be the power constraint on antenna i, then this constraint is equivalent to having a diagonal input covariance $\mathbf{Q} = \text{diag}\{P_i\}$. Denote $\mathbf{P} = \text{diag}\{P_i\}$, where $\text{tr}(\mathbf{P}) = P$ in relation to (3), then the multiple-access power constraint can also be expressed as

$$\mathbf{Q} \preccurlyeq \mathbf{P}. \tag{4}$$

3) Per-antenna power constraint: Here each antenna also has a separate transmit power budget of P_i $(i = 1, \ldots, n)$ but can fully cooperate with each other. Such a channel can model a physically centralized MIMO system, in which the per-antenna power comes from the realistic individual constraint of each transmit RF chain. The channel can also model a distributed (but cooperative) MIMO system, in which each transmit antenna belongs to a sensor or ad hoc node distributed in a network and thus cannot share power. The per-antenna constraint is equivalent to having the input covariance matrix $\mathbf{Q}$ with fixed diagonal values $Q_{ii} = P_i$. Denote $\mathbf{e}_i = [0 \ldots 1 \ldots 0]^T$ as a vector with the i^{th} element equal to 1 and the rest is 0. Then the per-antenna power constraint $Q_{ii} \leq P_i$ can be written as

$$\mathbf{e}_i^T \mathbf{Q} \mathbf{e}_i \leq P_i, \quad i = 1 \ldots n. \tag{5}$$

This constraint is a set of linear constraints on $\mathbf{Q}$. It should be stressed that a constraint on the diagonal values of $\mathbf{Q}$ is not the same as a constraint on the eigenvalues of $\mathbf{Q}$.

III. Capacity optimization problem

For all stated power constraints, this capacity optimization can be cast as follows.

$$\begin{aligned} \max \quad & \log\det\left(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right) \\ \text{s.t.} \quad & g(\mathbf{Q}, \mathbf{P}) \leq 0 \\ & \mathbf{Q} \succcurlyeq 0 \end{aligned} \tag{6}$$

where the input covariance $\mathbf{Q}$ is Hermitian, and $g(\mathbf{Q}, \mathbf{P}) \leq 0$ refers to a power constraint as in (3), (4) or (5).

Since all the power constraints are linear in $\mathbf{Q}$, the above optimization is convex with any power constraint. Thus for each power constraint, the problem admits a unique optimal solution for $\mathbf{Q}$. By Slater's condition [5], this optimal solution always exists because of the strictly feasible value $\mathbf{Q} = \mathbf{P} \succ 0$ which readily satisfies all power constraints.

With sum power constraint (3), the optimal solution is found by the well-known water-filling algorithm [1]. With multiple-access constraint (4), the obvious solution is $\mathbf{Q} = \mathbf{P}$. In this case the optimal signaling is sending independent signals from different antennas, each with the constrained power.

A. Capacity optimization using SDP framework

Since problem (6) with per-antenna power is convex, it can be analyzed using the SDP framework. Let $\mathbf{D} = \text{diag}\{d_i\} \succcurlyeq 0$ be a diagonal matrix consisting of Lagrangian multipliers d_i for the per-antenna power constraints in (5). Furthermore, let $\mathbf{M} \succcurlyeq 0$ be the Lagrangian multiplier for the positive semi-definite constraint. Then the Lagrangian for problem (6) can be formed as

$$\begin{aligned} &\mathcal{L}(\mathbf{Q}, \mathbf{D}, \mathbf{M}) \\ &= \log\det\left(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right) - \text{tr}[\mathbf{D}(\mathbf{Q} - \mathbf{P})] + \text{tr}(\mathbf{M}\mathbf{Q}). \end{aligned} \tag{7}$$

Taking the first order derivative of $\mathcal{L}$ in (7) with respect to $\mathbf{Q}$ (see [6] Appendix A.7 for derivatives of a function with respect to a matrix) and equating to zero, we obtain

$$\mathbf{H}^\dagger \left(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)^{-1} \mathbf{H} - \mathbf{D} + \mathbf{M} = 0. \tag{8}$$

Based on the KKT and (8), we obtain a set of optimality conditions as follows.

$$\begin{aligned} \mathbf{H}^\dagger \left(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)^{-1} \mathbf{H} &= \mathbf{D} - \mathbf{M} \\ \mathbf{M}\mathbf{Q} &= 0 \\ \text{diagonal } \mathbf{D} &\succ 0 \\ \text{Hermitian } \mathbf{M}, \mathbf{Q} &\succcurlyeq 0 \end{aligned} \tag{9}$$

The optimal input covariance $\mathbf{Q}$ is the solution to the set of equations in (9).

B. Rank of the optimal input covariance

Combining (8) with the complementary slackness condition $\mathbf{M}\mathbf{Q} = 0$, we have

$$\mathbf{D}\mathbf{Q} = \mathbf{H}^\dagger \left(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)^{-1} \mathbf{H}\mathbf{Q}. \tag{10}$$

Now at optimum, we must have $\mathbf{D} \succ 0$. That is, $\mathbf{D}$ is strictly positive definite and the power constraint is tight. In other words, at optimum, each constraints in (5) must be met with equality, for otherwise we can always increase the input power and get a higher rate; hence the associated dual variables are strictly positive at optimum. Thus at optimum, $\mathbf{D}$ is full-rank, subsequently (10) implies that

$$\text{rank}(\mathbf{Q}) \leq \text{rank}(\mathbf{H}). \tag{11}$$

Therefore, the rank of the optimal input covariance is no more than the channel rank. In other words, the number of independent signal streams (or modes) should be no more than the rank of the channel. This result is similar to that with sum power constraint but is different with multiple access constraint.

Based on (11), since channel $\mathbf{H}$ can support at most $r = \min\{m, n\}$ independent modes (independent signal streams), the rank of $\mathbf{Q}$ is at most r. When the rank of $\mathbf{Q}$ is less than r, that implies mode-dropping (similar to the same concept with sum power constraint). Since $\mathbf{QM} = 0$, $\mathbf{M}$ is a positive semidefinite matrix in the null space of $\mathbf{Q}$. The rank of $\mathbf{M}$ corresponds to the number of modes that has to be dropped for $\mathbf{Q}$ to be positive semidefinite. Suppose that the optimal solution has k modes dropped ($0 \leq k \leq \min\{m, n\}$), then

$$\text{rank}(\mathbf{M}) = k, \quad \text{rank}(\mathbf{Q}) = \min\{m, n\} - k.$$

The difference between the rank of $\mathbf{Q}$ and the size of $\mathbf{Q}$ should be stressed here. The size of $\mathbf{Q}$ is $n \times n$. If $n > m$ (more transmit than receive antennas), the optimal $\mathbf{Q}$ is inherently rank-deficient. In this case, even without any mode-dropping, the maximum rank of $\mathbf{Q}$ is $m < n$. Thus no mode-dropping does not always imply full-rank $\mathbf{Q}$. If $n \leq m$ then the maximum rank of $\mathbf{Q}$ is n without mode-dropping, and only then $\mathbf{Q}$ is full-rank.

C. Optimality conditions

Noting that $\mathbf{D}$ is full-rank and invertible, lets denote $\check{\mathbf{D}} = \mathbf{D}^{-1}$ and define

$$\mathbf{R}_m = \mathbf{HQH}^\dagger \;, \quad \mathbf{F}_m = \mathbf{H}\check{\mathbf{D}}\mathbf{H}^\dagger. \tag{12}$$

Then from the set of optimality conditions in (9), after some manipulation, we can show that

$$(\mathbf{R}_m - \mathbf{F}_m + \mathbf{I}_m)\mathbf{R}_m = 0. \tag{13}$$

The proof is given in Appendix A. Note that both $\mathbf{R}_m$ and $\mathbf{F}_m$ are $m \times m$ Hermitian matrices and the achievable rate for each channel state in (2) can now be expressed as $r(\mathbf{H}) = \log\det(\mathbf{I}_m + \mathbf{R}_m)$, which is a sole function of $\mathbf{R}_m$.

Condition (13) provides the equation for solving for $\mathbf{Q}$. To understand the meaning of this equation better, lets denote

$$\mathbf{S}_m = \mathbf{R}_m - \mathbf{F}_m + \mathbf{I}_m. \tag{14}$$

Then we can show that (see Appendix A)

$$\mathbf{H}\check{\mathbf{D}}\mathbf{M} = \mathbf{S}_m\mathbf{H}. \tag{15}$$

Since $\check{\mathbf{D}}$ is square and full rank, (15) implies $\text{rank}(\mathbf{S}_m) = \text{rank}(\mathbf{M}) = k$. From (12) and (11), we have $\text{rank}(\mathbf{R}_m) = \text{rank}(\mathbf{Q}) = \min\{m, n\} - k$. From (13), we have $\mathbf{S}_m\mathbf{R}_m = 0$. Thus $\mathbf{R}_m$ contains the active transmission modes, and $\mathbf{S}_m$ contains the modes that are dropped. In other words, $\mathbf{S}_m$ is a matrix in the null space of $\mathbf{R}_m$, in the same way that $\mathbf{M}$ is in the null space of $\mathbf{Q}$. In the next section, we will use (13) to solve for the optimal $\mathbf{Q}$.

IV. Optimal input covariance

In this section, we solve for the optimal value of $\mathbf{Q}$ as a function of the dual variable $\mathbf{D}$. Here we focus on the case that $n \leq m$ (solution for the case $n > m$ can be found in [7]). Let the singular value decomposition of the channel matrix be

$$\mathbf{H} = \mathbf{U_H}\mathbf{\Sigma_H}\mathbf{V}_\mathbf{H}^\dagger \tag{16}$$

where

- $\mathbf{U_H}$ is a $m \times m$ unitary matrix of the left singular vectors,
- $\mathbf{V_H}$ is a $n \times n$ unitary matrix of the *right singular vectors*,
- $\mathbf{\Sigma_H}$ is a $m \times n$ diagonal matrix containing the (real) singular values in decreasing order.

For $n \leq m$, $\text{rank}(\mathbf{H}) = n$ and $\text{rank}(\mathbf{Q}) = n - k$. However, $\mathbf{F}_m$ as defined in (12) has size $m \times m$ and is not full-rank. It is necessary to convert this matrix to full-rank as follows. In (16), for $n \leq m$, we have $\mathbf{\Sigma_H} = [\mathbf{\Sigma}_n \; \mathbf{0}_{n,m-n}]^T$, where $\mathbf{\Sigma}_n$ is a $n \times n$ diagonal matrix containing the real singular values of $\mathbf{H}$. Now define

$$\mathbf{K} = \mathbf{V_H}\mathbf{\Sigma}_n\mathbf{V}_\mathbf{H}^\dagger. \tag{17}$$

Next define two $n \times n$ matrices as

$$\mathbf{R}_n = \mathbf{KQK}^\dagger \;, \quad \mathbf{F}_n = \mathbf{K}\check{\mathbf{D}}\mathbf{K}^\dagger. \tag{18}$$

Multiplying (13) on the left with $\mathbf{H}^\dagger$ and on the right with $\mathbf{H}$, and noting that $\mathbf{H}^\dagger\mathbf{H} = \mathbf{KK}^\dagger$, we get

$$\mathbf{K}\left[(\mathbf{R}_n - \mathbf{F}_n + \mathbf{I}_n)\mathbf{R}_n\right]\mathbf{K}^\dagger = 0.$$

Since for $n \leq m$, $\mathbf{K}$ is square and full-rank, the above equation is equivalent to

$$(\mathbf{R}_n - \mathbf{F}_n + \mathbf{I}_n)\mathbf{R}_n = 0. \tag{19}$$

We get an equation in a form similar to (13), but in this case of $n \leq m$, $\mathbf{F}_n$ is full-rank.

Equation (19) can be written as $\mathbf{R}_n^2 + \mathbf{R}_n = \mathbf{F}_n\mathbf{R}_n$. This equality implies that $\mathbf{F}_n\mathbf{R}_n$ is Hermitian and has the same eigenvalue decomposition as $\mathbf{R}_n^2 + \mathbf{R}_n$, which has the same eigenvectors as those of $\mathbf{R}_n$. This is possible only if $\mathbf{R}_n$ and $\mathbf{F}_n$ share the same eigenvectors for the non-zero eigenvalues. Now since for $n \leq m$, $\mathbf{F}_n$ is full rank and Hermitian, it has a unique eigenvalue decomposition (up to any multiplicity of eigenvalues). From (18), $\text{rank}(\mathbf{R}_n) = \text{rank}(\mathbf{Q}) = n - k$. Equation (19) then implies that $\mathbf{R}_n$ must span $n - k$ eigenspaces of $\mathbf{F}_n$. That is, the $n - k$ eigenvectors corresponding to the non-zero eigenvalues of $\mathbf{R}_n$ are the same as $n - k$ eigenvectors of $\mathbf{F}_n$. Let

$$\mathbf{S}_n = \mathbf{R}_n - \mathbf{F}_n + \mathbf{I}_n, \tag{20}$$

then equivalently, (19) implies that $\mathbf{S}_n$ spans the *other* k eigenspaces of $\mathbf{F}_n$ (the k eigenvectors with non-zero eigenvalues of $\mathbf{S}_n$ are the same as the other k eigenvectors of $\mathbf{F}_n$). Here $\mathbf{S}_n$ contains the k modes that are dropped.

Intuitively, this result can be interpreted as follows. Note that we can write $\mathbf{F}_n - \mathbf{I}_n = \mathbf{R}_n - \mathbf{S}_n$. The matrix $\mathbf{F}_n - \mathbf{I}_n$ may contain some positive and some non-positive eigenvalues. Then $\mathbf{R}_n$ is the portion that contains only the positive eigenvalues, and $(-\mathbf{S}_n)$ is the portion that contains only the non-positive eigenvalues. As such, both $\mathbf{R}_n$ and $\mathbf{S}_n$ are positive semidefinite matrices and are orthogonal to each other.

Denote $\check{\mathbf{K}} = \mathbf{K}^{-1}$, then from (18), $\mathbf{Q} = \check{\mathbf{K}}\mathbf{R}_n\check{\mathbf{K}}^\dagger$. Based on (20), we can find the optimal $\mathbf{Q}$ as

$$\mathbf{Q}^\star = \check{\mathbf{D}} - \check{\mathbf{K}}\check{\mathbf{K}}^\dagger + \mathbf{Z}, \tag{21}$$

where $\mathbf{Z} = \check{\mathbf{K}}\mathbf{S}_n\check{\mathbf{K}}^\dagger$. This equation gives the solution for $\mathbf{Q}$ as a function of the dual variable $\mathbf{D}$. Here $\mathbf{K}$ is a function of the channel as defined in (17), while $\mathbf{S}_n$ is determined from $\mathbf{F}_n$ which is a function of $\mathbf{D}$ as in (18). Note that since $\mathbf{R}_n$ contains only the positive eigenmodes of $\mathbf{F}_n - \mathbf{I}_n$, the optimal $\mathbf{Q}^\star$ as formed in (21) is always positive semidefinite. Thus the only step left is to find the optimal dual variable $\mathbf{D}$ such that $\mathbf{Q}^\star$ satisfies the power constraint of $\text{diag}(\mathbf{Q}^\star) = \mathbf{P}$.

To find the optimal $\mathbf{D}$, at this point, we need to use an iterative algorithm which we will discuss next.

V. Algorithm for finding the optimal $\mathbf{Q}$

There appears to be no closed-form analytical solution for a diagonal $\check{\mathbf{D}}$ such that the solution in (21) satisfies $\text{diag}(\mathbf{Q}^\star) = \mathbf{P}$. Fortunately, these equations suggest a way to update $\mathbf{D}$ iteratively. In the follows, we design an iterative algorithm for finding the optimal $\check{\mathbf{D}}$, and hence $\mathbf{Q}^\star$.

The approach is similar to mode-dropping in water filling. We first assume that there is no mode-dropping, so that $\mathbf{S}_n = 0$ and find the corresponding $\mathbf{Q}$ with diagonal $\mathbf{P}$. If that solution for $\mathbf{Q}$ is positive semidefinite, it is the optimal solution. If not, then we need to find a new $\check{\mathbf{D}}$, and equation (21) suggests a way to update $\check{\mathbf{D}}$ at each iteration.

The initial step of this algorithm is straightforward. Assume no mode-dropping, then $\mathbf{S}_n = \mathbf{0}$ and (21) implies $\mathbf{Q} = \check{\mathbf{D}} - \check{\mathbf{K}}\check{\mathbf{K}}^\dagger$. Thus we can just simply choose diagonal matrix $\check{\mathbf{D}}$ as

$$\check{\mathbf{D}}_0 = \mathbf{P} + \text{diag}(\check{\mathbf{K}}\check{\mathbf{K}}^\dagger). \quad (22)$$

This solution always satisfies $\mathbf{D} \succ 0$ since $\check{\mathbf{K}}\check{\mathbf{K}}^\dagger$ as a positive semidefinite matrix has non-negative diagonal values. Then we can form

$$\mathbf{Q}_0^\star = \check{\mathbf{D}}_0 - \check{\mathbf{K}}\check{\mathbf{K}}^\dagger \quad (23)$$

and check if this $\mathbf{Q}_0^\star$ is positive semidefinite. If it is, (23) is the optimal input covariance.

If $\mathbf{Q}_0^\star$ in (23) is non-positive semidefinite, then we adjust $\check{\mathbf{D}}$ using an iterative procedure as follows. Say at iteration i ($i \geq 0$), we have obtained $\check{\mathbf{D}}_i$. Then we can form $\mathbf{F}_{n,i}$, $\mathbf{S}_{n,i}$ and $\mathbf{Q}_i$ as

$$\begin{aligned} \mathbf{F}_{n,i} &= \mathbf{K}\check{\mathbf{D}}_i\mathbf{K} \\ -\mathbf{S}_{n,i} &= \text{ non-positive eigenmodes of } (\mathbf{F}_{n,i} - \mathbf{I}_n) \\ \mathbf{Z}_i &= \check{\mathbf{K}}\mathbf{S}_{n,i}\check{\mathbf{K}}^\dagger \\ \mathbf{Q}_i &= \check{\mathbf{D}}_i - \check{\mathbf{K}}\check{\mathbf{K}}^\dagger + \mathbf{Z}_i. \end{aligned} \quad (24)$$

The $\mathbf{Q}_i$ as computed in (24) is always positive semidefinite (as a consequence of (21)). Thus the term $\mathbf{Z}_i \not\succcurlyeq 0$ is the adjustment at step i to make $\mathbf{Q}_i \succcurlyeq 0$. But it does not guarantee that the diagonal of $\mathbf{Q}_i$ will be $\mathbf{P}$. From (24), noting that $\check{\mathbf{D}}$ is diagonal, we update $\check{\mathbf{D}}_{i+1}$ by the difference between the diagonal of $\mathbf{Q}_i$ and $\mathbf{P}$ as

$$\check{\mathbf{D}}_{i+1} = \check{\mathbf{D}}_i + \mathbf{P} - \text{diag}(\mathbf{Q}_i) \quad (25)$$

Iteration stops when the diagonal values of $\mathbf{Q}_i$ is close to $\mathbf{P}$ within an acceptable tolerance. In implementation, we choose to stop when the duality gap satisfies $|\text{tr}[\mathbf{D}(\mathbf{Q}-\mathbf{P})]| < \epsilon$. Since problem (6) is convex and satisfies Slater's condition, this stopping criterion always guarantees the optimal solution. The iterative procedure for finding $\mathbf{Q}^\star$ when $n \leq m$ is summarized in Algorithm 1, *drop-rank-n*$(\cdot)$.

Algorithm 1 *drop-rank-n*$(n, \check{\mathbf{D}}_0, \mathbf{K}, \check{\mathbf{K}}, \mathbf{P}, \epsilon)$: Iterative search for $\mathbf{Q}^\star$ when $n \leq m$.

Require: $\check{\mathbf{D}}_0$ diagonal $\succ 0$, $\mathbf{P}$ diagonal $\succ 0$, $\epsilon > 0$. Here $\mathbf{K}$ is defined as in (17) and $\check{\mathbf{K}} = \mathbf{K}^{-1}$.

Initialize $i = 0$. (iteration count)
Initialize $\Delta = 1 + \epsilon$. (loop terminating variable)
while $(\Delta > \epsilon)$ **do**
  Form $\mathbf{F}_i = \mathbf{K}\check{\mathbf{D}}_i\mathbf{K}^\dagger - \mathbf{I}_n$. (note that $-\mathbf{I}_n$ is included here)
  Perform the eigenvalue decomposition $\mathbf{F}_i = \mathbf{U_F}\mathbf{\Lambda_F}\mathbf{U_F}^\dagger$. Let k be the number of non-positive eigenvalues of $\mathbf{F}$.
  Form $\mathbf{S}_i = -\mathbf{U}_\mathbf{F}^k\mathbf{\Lambda}_\mathbf{F}^k{\mathbf{U}_\mathbf{F}^k}^\dagger$ where
    $\mathbf{\Lambda}_\mathbf{F}^k$ is the $k \times k$ diagonal matrix of all k non-positive eigenvalues of $\mathbf{F}$
    $\mathbf{U}_\mathbf{F}^k$ consists of the corresponding k eigenvectors.
  Form $\mathbf{Z}_i = \check{\mathbf{K}}\mathbf{S}_i\check{\mathbf{K}}^\dagger$.
  Form $\mathbf{Q}_i = \check{\mathbf{D}}_i - \check{\mathbf{K}}\check{\mathbf{K}}^\dagger + \mathbf{Z}_i$.
  Form $\check{\mathbf{D}}_{i+1} = \check{\mathbf{D}}_i + \mathbf{P} - \text{diag}(\mathbf{Q}_i)$.
  Compute $\Delta = |\text{tr}[\mathbf{D}_i(\mathbf{Q}_i - \mathbf{P})]|$.
  $i \leftarrow i + 1$.
end while
return $\check{\mathbf{D}}_i$ and $\mathbf{Q}_i$.

P_1/P	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$P = 1$	47	33	28	24	22	22	23	26	36
$P = 10$	31	20	0	0	0	0	0	16	24

TABLE I
Number of iterations for two different total power P.

VI. Numerical Examples

In this section, we perform numerical examples for a 2×2 channel. The test channel is generated randomly according to the circularly complex Gaussian with zero mean and unit variance and is equal to

$$\mathbf{H} = \begin{bmatrix} -0.2581 + 0.6535i & -0.2623 + 0.9434i \\ 0.4385 + 0.3081i & 0.4090 - 0.2288i \end{bmatrix}. \quad (26)$$

First, let the total power $P = 1$ and let the power constraint of the first antenna P_1 varies. Figure 1 shows the simulation results comparing the different capacities. We see that the capacity with per-antenna power can be significantly different from either with sum power or multiple access constraints. Figure 2 shows the convergence of the proposed algorithm in terms of the objective function and the duality gap for 3 different values of P_1 (tolerance $\epsilon = 10^{-6}$). We observe that when P_1 is closer to $P/2$, convergence is faster.

Second, if we increase the total power, we also observe faster convergence. Table I shows the number of iterations for each value of P_1/P at two different total power P for the same test channel in (26). In some cases, we see 0 iteration, implying that the initial value in (23) is optimal.

Third, the channel right singular vectors, which are the optimal input covariance eigenvectors under sum power, are no longer optimal under per-antenna power. Here the optimal

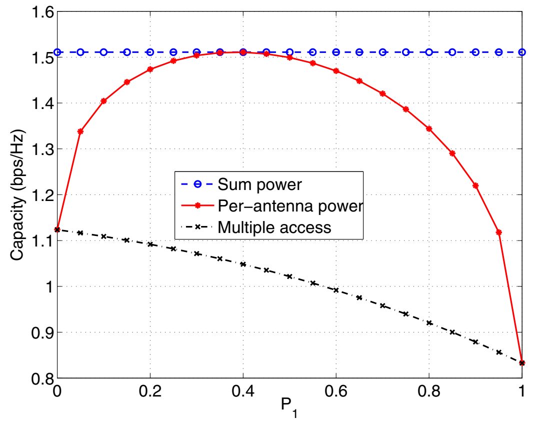

Fig. 1. Capacities for a 2×2 channel under different power constraints.

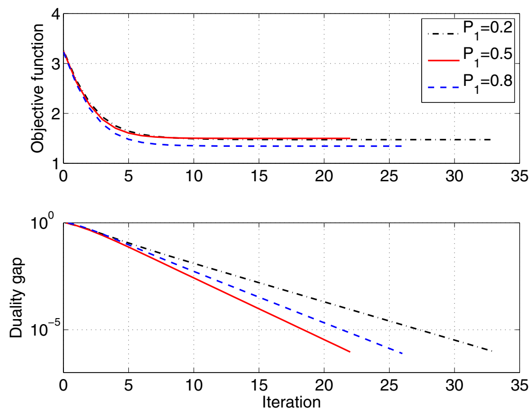

Fig. 2. Convergence behavior for the test channel with per-antenna power (tolerance $\epsilon = 10^{-6}$).

$\mathbf{D}$ varies with the per-antenna constraint $\mathbf{P}$ and the signal-to-noise ratio (SNR). Consequently, often both the eigenvectors and eigenvalues of the optimal covariance in (21) also vary with $\mathbf{P}$ and SNR. The proposed algorithm updates both simultaneously.

VII. Conclusion

We have established the MIMO capacity with per-antenna power constraint for channel known at both transmitter and receiver. The optimal input covariance matrix shows that its eigenvectors are not the same as the channel right singular vectors as in the case of sum power constraint. We design an efficient algorithm to find this optimal input covariance. Simulation results show that the per-antenna constraint can affect the capacity significantly.

Appendix

A. Proofs of MIMO optimal conditions in (13) and (15)

In the following derivation, we make repeated use of the following identity expansion:

$$\mathbf{I}_m = \left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1}. \qquad (27)$$

Note that the order of the two factors in this expansion is interchangeable.

To prove (13), multiplying the first equation in (9) on the right with $\mathbf{QH}^\dagger$ and on the left with $\mathbf{H\check{D}}$, and noting that $\mathbf{MQ} = 0$, we get

$$\mathbf{H\check{D}H}^\dagger\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1}\mathbf{HQH}^\dagger = \mathbf{HQH}^\dagger.$$

Now subtracting both sides by $\mathbf{H\check{D}H}^\dagger$, then applying the identity expansion (27), this equation simplifies to

$$-\mathbf{H\check{D}H}^\dagger\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1} = \mathbf{HQH}^\dagger - \mathbf{H\check{D}H}^\dagger.$$

Next adding both sides of this equation with $\mathbf{I}_m$ and again using the identity expansion (27) on the left side, we get

$$\begin{aligned}&\left(\mathbf{I}_m + \mathbf{HQH}^\dagger - \mathbf{H\check{D}H}^\dagger\right)\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1}\\ &= \mathbf{I}_m + \mathbf{HQH}^\dagger - \mathbf{H\check{D}H}^\dagger.\end{aligned}$$

Denote $\mathbf{S}_m = \mathbf{I}_m + \mathbf{HQH}^\dagger - \mathbf{H\check{D}H}^\dagger$, the above equation becomes

$$\mathbf{S}_m\left(\mathbf{I}_m - \left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1}\right) = 0.$$

Then applying the identity expansion (27) once more, we obtain

$$\mathbf{S}_m\left(\mathbf{HQH}^\dagger\right)\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1} = 0.$$

Since $\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1}$ is a full-rank square matrix, the above equation is equivalent to

$$\mathbf{S}_m\left(\mathbf{HQH}^\dagger\right) = 0.$$

This equality gives (13).

To show (15), from the first equation in (9), subtracting $\mathbf{H}^\dagger\mathbf{H}$ from both sides, we have

$$\mathbf{H}^\dagger\left[\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1} - \mathbf{I}_m\right]\mathbf{H} = \mathbf{D} - \mathbf{M} - \mathbf{H}^\dagger\mathbf{H}.$$

Using the identity expansion (27), we obtain

$$\mathbf{H}^\dagger\left(\mathbf{I}_m + \mathbf{HQH}^\dagger\right)^{-1}\mathbf{HQH}^\dagger\mathbf{H} = \mathbf{H}^\dagger\mathbf{H} - \mathbf{D} + \mathbf{M}.$$

Now replacing part of the expression on the left with the first equation in (9), we have

$$(\mathbf{D} - \mathbf{M})\mathbf{QH}^\dagger\mathbf{H} = \mathbf{H}^\dagger\mathbf{H} - \mathbf{D} + \mathbf{M}.$$

But $\mathbf{MQ} = 0$, hence we have

$$\mathbf{M} = \mathbf{DQH}^\dagger\mathbf{H} + \mathbf{D} - \mathbf{H}^\dagger\mathbf{H}.$$

Multiplying both sides on the left with $\mathbf{H\check{D}}$, we obtain (15).

References

[1] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., 2006.

[2] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, Jun 2007.

[3] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *Information Theory, IEEE Transactions on*, vol. 50, no. 1, pp. 145–152, 2004.

[4] I. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov 1999.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2003. [Online]. Available: http://www.stanford.edu/∼boyd/cvxbook.html

[6] H. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. New York: Wiley, 2002.

[7] M. Vu, "The capacity of mimo channels with per-antenna power constraint," *Arxiv preprint arXiv:1106.5039*, 2011.