Minimax Problems and Directional Derivatives for MIMO Channels

Anke Feiten and Rudolf Mathar
Institute of Theoretical Information Technology
RWTH Aachen University, D-52056 Aachen
phone +49 241 8027700, fax +49 241 22198
E-mail: {feiten,mathar}@ti.rwth-aachen.de

Abstract— **When transmitting over multiple-input-multiple-output (MIMO) channels, in the case of total power constraints and complete channel state information (CSI) the optimum power distribution is obtained by water-filling the squared singular values of the channel matrix $\boldsymbol{H}$. In this paper, we consider the case that nature behaves as an opponent to the optimum transmit strategy in choosing the channel as bad as possible. Interpreting the mutual information as payoff function for two players, the transmitter and a malicious nature, this approach may be seen as a two person zero sum game.**

We first analyze maximum points of the payoff function for a fixed channel matrix under general power restrictions and characterize such points via directional derivatives. Worst channel behavior must be separated from the zero channel where no transmission is possible at all. Loewner semi-ordering of nonnegative Hermitian matrices is employed to ensure minimum channel quality. It is shown that a Nash equilibrium exists for general power constraints. Concrete results are achieved for a limited total power budget and limiting the maximum available power for each subchannel.

I. INTRODUCTION

Multiple-input-multiple-output (MIMO) channels are an important means to improve the performance of wireless systems. Recent seminal work in [1] and [2] has shown that the use of multiple antennas at both ends significantly increases the information-theoretic capacity far beyond that of single-antenna systems in rich scattering propagation environments.

In this paper, we choose two complementary approaches to describe the capacity of MIMO channels. First, we investigate a scenario where the channel state is known. In case of total power constraints at the transmitter, the capacity and the associated optimum power strategy is characterized via directional derivatives of the objective function. Using this characterization the well known water filling solution, see [2], [3], [4], can be easily verified as the optimum point. Analogous results are obtained when the maximum power is bounded. The optimum solution here is also characterized, and explicitly determined, by directional derivatives.

Secondly, we consider the set of all channel matrices which are bounded from below with respect to the Loewner semi-ordering. The approach is embedded into a game theoretic framework. A two-person zero sum game is considered where the two players are the transmitter and a malicious nature. The payoff function is the mutual information.

This setup has also been developed in [5]. The authors [5] obtain a uniform power allocation as the solution of the game under the assumption that channels are isotropically unconstrained. This essentially means that any unitary transformation of some channel matrix $\boldsymbol{H}$, i.e., any other direction of the subchannels, is also an option for nature to choose. We generalize the approach in merely requiring convexity of the set of power distributions and a lower bound for $\boldsymbol{H}^*\boldsymbol{H}$ in the Loewner semi-ordering to avoid the trivial zero solution.

Related minimax problems have been considered in [6] in the framework of worst case analysis, and [7] for the case that possible channels satisfy $\operatorname{tr}(\boldsymbol{H}^*\boldsymbol{H}) \geq \beta$.

This paper is organized as follows. We start by introducing the mathematical techniques used throughout the paper in Section II. In Section III we present the system model and investigate a MIMO channel under the assumption that the channel state is known and fixed. The optimum power strategy is characterized by directional derivatives of a concave function. In the case of total power restrictions the well known water filling solution is confirmed. Analogously, for restrictions on the maximum power an explicit solution and a dual characterization is given. We deal with the case of an unknown channel state in Section IV. The underlying two-person zero sum game is shown to have a saddle point. A short summary and a discussion of future work is given in Section V.

II. PRELIMINARIES

Let f be a real-valued concave function with convex domain $\mathcal{C}$ and $\hat{x}, x \in \mathcal{C}$. The directional derivative of f at $\hat{x}$ in the direction of x is defined as

$$\begin{aligned} Df(\hat{x}, x) &= \lim_{\alpha \to 0+} \frac{1}{\alpha} \Big[f\big((1-\alpha)\hat{x} + \alpha x\big) - f(\hat{x}) \Big] \\ &= \frac{d}{d\alpha} f\big((1-\alpha)\hat{x} + \alpha x\big)\Big|_{\alpha=0+}, \end{aligned} \tag{1}$$

see, e.g., [8]. Since f is concave, $\big(f((1-\alpha)\hat{x}+\alpha x) - f(\hat{x})\big)/\alpha$ is monotone increasing with decreasing $\alpha \geq 0$, and the directional derivative always exists.

If $\mathcal{C}$ is a subset of a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, it is well known that

$$Df(\hat{x}, x) = \langle \nabla f(\hat{x}), x - \hat{x} \rangle, \tag{2}$$

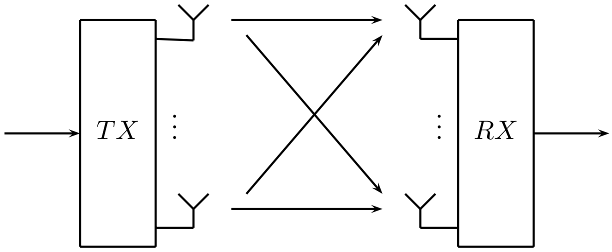

Fig. 1. A MIMO channel

whenever ∇f, the derivative of f in the strong sense, exists. Optimum points are characterized by directional derivatives as follows, for a proof see [8].

Proposition 1: Let $\mathcal{C}$ be a convex set and $f : \mathcal{C} \to \mathbb{R}$ a concave function. Then the maximum of f is attained at $\hat{x}$ if and only if $Df(\hat{x}, x) \leq 0$ for all $x \in \mathcal{C}$.

In the following we make frequent use of the Loewner semi-ordering on the set of nonnegative definite Hermitian matrices, defined as $\boldsymbol{A} \leq \boldsymbol{B}$ if $\boldsymbol{B} - \boldsymbol{A}$ is nonnegative definite. For reference purposes we quote the following result due to Loewner, see [9], A.1.b., p. 510. Let $\lambda_{(1)}(\boldsymbol{A}) \leq \cdots \leq \lambda_{(m)}(\boldsymbol{A})$ denote the eigenvalues of some nonnegative definite Hermitian matrix $\boldsymbol{A}$ in increasing order.

Proposition 2: If $\boldsymbol{A} \leq \boldsymbol{B}$, then $\lambda_{(i)}(\boldsymbol{A}) \leq \lambda_{(i)}(\boldsymbol{B})$ for all $i = 1, \ldots, m$.

III. Equivalence Theorems

We consider a MIMO transmission system with r receive antennas and t transmit antennas. After appropriate renormalization (see [10]) we obtain the channel model

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{Z}$$

with some complex $r \times t$ matrix $\boldsymbol{H}$ and random noise vector $\boldsymbol{Z} \in \mathbb{C}^r$. $\boldsymbol{Z}$ is circularly symmetric complex Gaussian distributed (see [2]) with expectation $\boldsymbol{0}$ and covariance matrix $\mathrm{E}(\boldsymbol{Z}\boldsymbol{Z}^*) = \boldsymbol{I}_r$, denoted by $\boldsymbol{Z} \sim \mathrm{SCN}(\boldsymbol{0}, \boldsymbol{I}_r)$. The complex zero mean input vector $\boldsymbol{X}$ is subject to power constraints described by

$$\mathrm{E}(\boldsymbol{X}\boldsymbol{X}^*) = \boldsymbol{Q} \in \mathcal{Q}$$

for some set of nonnegative definite matrices $\mathcal{Q}$.

By slightly extending the arguments in [2] the capacity of a MIMO channel is derived as the maximum of the mutual information over all admissible input distributions of $\boldsymbol{X}$ as

$$C = \max_{\boldsymbol{Q} \in \mathcal{Q}} I(\boldsymbol{X}, \boldsymbol{Y}) = \max_{\boldsymbol{Q} \in \mathcal{Q}} \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*).$$

We now characterize the covariance matrix $\hat{\boldsymbol{Q}}$ which achieves capacity by using directional derivatives of the function

$$f : \mathcal{Q} \to \mathbb{R} : \boldsymbol{Q} \mapsto \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*).$$

From Ky Fan's inequality it follows immediately that f is concave whenever its domain $\mathcal{Q}$ is convex.

Proposition 3: Let $\mathcal{Q}$ be convex and $\hat{\boldsymbol{Q}}, \boldsymbol{Q} \in \mathcal{Q}$. The directional derivative of f at $\hat{\boldsymbol{Q}}$ in the direction of $\boldsymbol{Q}$ is given by

$$Df(\hat{\boldsymbol{Q}}, \boldsymbol{Q}) = \mathrm{tr}\left[\boldsymbol{H}^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,(\boldsymbol{Q} - \hat{\boldsymbol{Q}})\right]. \tag{3}$$

Proof: We exploit the chain rule for real valued functions $g(\boldsymbol{X})$ where the matrix $\boldsymbol{X}$ is itself a function of a scalar α,

$$\frac{dg}{d\alpha} = \mathrm{tr}\left[\frac{dg}{d\boldsymbol{X}}\Big(\frac{d\boldsymbol{X}}{d\alpha}\Big)^*\right]$$

Furthermore, we use the fact that $\frac{d}{d\alpha}\det\boldsymbol{X} = (\det\boldsymbol{X})(\boldsymbol{X}^{-1})^*$, cp. [11] or [12]. Hence,

$$\begin{aligned}
&\frac{d}{d\alpha} f(\hat{\boldsymbol{Q}} + \alpha(\boldsymbol{Q} - \hat{\boldsymbol{Q}})) \\
&= \frac{d}{d\alpha}\log\det\left(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^* + \alpha\boldsymbol{H}(\boldsymbol{Q} - \hat{\boldsymbol{Q}})\boldsymbol{H}^*\right) \\
&= \mathrm{tr}\left[(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^* + \alpha\boldsymbol{H}(\boldsymbol{Q} - \hat{\boldsymbol{Q}})\boldsymbol{H}^*)^{-1}\boldsymbol{H}(\boldsymbol{Q} - \hat{\boldsymbol{Q}})\boldsymbol{H}^*\right].
\end{aligned}$$

The value $\alpha = 0$ and cyclically exchanging $\boldsymbol{H}^*$ in the trace yields representation (3) ■

From (2) and Proposition 3 we also conclude that the strong derivative of f at $\hat{\boldsymbol{Q}}$ in the Hilbert space of all complex $t \times t$ matrices endowed with the inner product $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{tr}\,\boldsymbol{A}\boldsymbol{B}^*$, see [13], p. 286, amounts to

$$\nabla f(\hat{\boldsymbol{Q}}) = \boldsymbol{H}^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}.$$

According to Proposition 1 the point $\hat{\boldsymbol{Q}}$ maximizes $f(\boldsymbol{Q})$ over some convex set $\mathcal{Q}$ iff $Df(\hat{\boldsymbol{Q}}, \boldsymbol{Q}) \leq 0$ for all $\boldsymbol{Q} \in \mathcal{Q}$. By (3) this leads to

$$\begin{aligned}
&\mathrm{tr}\left[\boldsymbol{H}^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{Q}\right] \\
&\leq \mathrm{tr}\left[\boldsymbol{H}^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\hat{\boldsymbol{Q}}\right]
\end{aligned}$$

for all $\boldsymbol{Q} \in \mathcal{Q}$. Hence, we obtain the following proposition.

Proposition 4: $\max_{\boldsymbol{Q} \in \mathcal{Q}} f(\boldsymbol{Q})$ is attained at $\hat{\boldsymbol{Q}}$ if and only if $\hat{\boldsymbol{Q}}$ is a solution of

$$\max_{\boldsymbol{Q} \in \mathcal{Q}} \mathrm{tr}\left[\boldsymbol{H}^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{Q}\right]. \tag{4}$$

The main value of this equivalence is that once a candidate for maximizing $f(\boldsymbol{Q})$ over $\mathcal{Q}$, hence achieving capacity, is claimed it can be verified by checking a merely linear condition. Moreover, in concrete cases nice characterizations of the optimum are obtained, as is developed in the following.

A. Total Power Constraint

We first investigate total power constraints

$$\mathcal{Q}_{\mathrm{tot}} = \{\boldsymbol{Q} \geq \boldsymbol{0} \mid \mathrm{tr}\,\boldsymbol{Q} \leq L\}.$$

Obviously, the set $\mathcal{Q}_{\mathrm{tot}}$ is convex. Covariance matrices $\boldsymbol{Q}$ which achieve capacity, i.e., solutions of

$$\max_{\mathrm{tr}\,\boldsymbol{Q} \leq L} \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*) \tag{5}$$

are now characterized as follows.

Proposition 5: $\hat{\boldsymbol{Q}}$ is a solution of (5) iff

$$\begin{aligned}&\lambda_{\max}\big(\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big)\\&\quad=\frac{r}{L}-\frac{1}{L}\operatorname{tr}(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\end{aligned}\tag{6}$$

Proof: We start from equation (4) and first note that $\{\boldsymbol{x}\boldsymbol{x}^*\geq 0\mid \boldsymbol{x}^*\boldsymbol{x}=L\}$ is the set of extreme points of $\mathcal{Q}_{\text{tot}}$. As at least one optimal solution is attained at an extreme point it follows that

$$\begin{aligned}&\max_{\operatorname{tr}\boldsymbol{Q}\leq L}\ \operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{Q}\big]\\&\quad=\max_{\boldsymbol{x}^*\boldsymbol{x}=L}\ \operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{x}\boldsymbol{x}^*\big]\\&\quad=\max_{\boldsymbol{x}^*\boldsymbol{x}=L}\ \boldsymbol{x}^*\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{x}\\&\quad=L\,\lambda_{\max}\big(\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big),\end{aligned}\tag{7}$$

where we used Fischer's minmax representation, $\lambda_{\max}(\boldsymbol{A})=\max_{\boldsymbol{x}^*\boldsymbol{x}=1}\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}$, in the last equation and where $\lambda_{\max}(\boldsymbol{A})$ denotes the maximum eigenvalue of the matrix $\boldsymbol{A}$.

Further, using the identity $(\boldsymbol{I}+\boldsymbol{A})^{-1}\boldsymbol{A}=\boldsymbol{I}-(\boldsymbol{I}+\boldsymbol{A})^{-1}$ we conclude

$$\begin{aligned}&\operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\hat{\boldsymbol{Q}}\big]\\&=\operatorname{tr}\big[(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*\big]\\&=\operatorname{tr}\big[\boldsymbol{I}_r-(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\big]\\&=r-\operatorname{tr}(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}.\end{aligned}\tag{8}$$

Dividing the right hand side of (7) and (8) by L completes the proof. ■

It is easy to see that the well known water filling solution

$$\hat{\boldsymbol{Q}}_{\text{wf}}=\boldsymbol{V}\operatorname{diag}(\nu-\gamma_i^{-1})^+\boldsymbol{V}^*$$

actually satisfies condition (6). Here ν is defined by water filling the inverse positive eigenvalues γ_i of $\boldsymbol{H}^*\boldsymbol{H}$ as $\sum_{i:\gamma_i>0}\big(\nu-\gamma_i^{-1}\big)^+=L$.

Using the singular value decomposition $\boldsymbol{H}=\boldsymbol{U}\boldsymbol{\Gamma}^{1/2}\boldsymbol{V}^*$ and the optimal solution $\hat{\boldsymbol{Q}}_{\text{wf}}$ it is straightforward to show that

$$\lambda_{\max}\big(\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big)=\frac{1}{\nu}.$$

On the other hand, using the same decomposition some algebra gives

$$\operatorname{tr}\big(\boldsymbol{I}_r-(\boldsymbol{I}_r+\boldsymbol{h}\hat{\boldsymbol{Q}}_{\text{wf}}\boldsymbol{H}^*)^{-1}\big)=\frac{L}{\nu},$$

which verifies (6).

B. Maximum Eigenvalue Constraint

An analogous characterization can be derived if the maximum eigenvalue is bounded by some constant L as

$$\mathcal{Q}_{\max}=\big\{\boldsymbol{Q}\geq\boldsymbol{0}\mid\lambda_{\max}(\boldsymbol{Q})\leq L\big\}.$$

Physically this means to constrain the maximum transmit power across antennas.

From Fischer's minmax representation $\lambda_{\max}(\boldsymbol{A})=\max_{\boldsymbol{x}^*\boldsymbol{x}=1}\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}$, cf. [9], p. 510, it follows that the set $\mathcal{Q}_{\max}$ is convex. We aim at determining the solution of

$$\max_{\lambda_{\max}(\boldsymbol{Q})\leq L}\ \log\det\big(\boldsymbol{I}_r+\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*\big).\tag{9}$$

First an explicit solution of (9) is determined, subsequently we give a nice characterization of the optimum point, dual to Propositon 5.

Proposition 6: The maximum in (9) is attained at $\hat{\boldsymbol{Q}}=L\boldsymbol{I}_t$ with value $\sum_{i=1}^r\log(1+L\gamma_i)$, where γ_i, $i=1,\dots,r$, denote the eigenvalues of $\boldsymbol{H}^*\boldsymbol{H}$.

Proof: Let $\boldsymbol{H}=\boldsymbol{U}\boldsymbol{\Gamma}^{1/2}\boldsymbol{V}^*$ denote the singular value decomposition of $\boldsymbol{H}$. Setting $\hat{\boldsymbol{Q}}=L\boldsymbol{I}_t$ maximization problem (4) may be rewritten as

$$\begin{aligned}&\max_{\lambda_{\max}(\boldsymbol{Q})\leq L}\ \operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{Q}\big]\\&=\max_{\lambda_{\max}(\boldsymbol{Q})\leq L}\ \operatorname{tr}\big[\boldsymbol{V}\operatorname{diag}\big(\gamma_i/(1+L\gamma_i)\big)\boldsymbol{V}^*\boldsymbol{Q}\big]\\&=\max_{\lambda_{\max}(\boldsymbol{Q})\leq L}\ \operatorname{tr}\big[\operatorname{diag}\big(\gamma_i/(1+L\gamma_i)\big)\boldsymbol{V}^*\boldsymbol{Q}\boldsymbol{V}\big]\\&=\max_{\lambda_{\max}(\boldsymbol{Q})\leq L}\ \operatorname{tr}\big[\operatorname{diag}\big(\gamma_i/(1+L\gamma_i)\big)\boldsymbol{Q}\big]\\&\leq\sum_{i=1}^r\frac{L\gamma_i}{1+L\gamma_i},\end{aligned}$$

where $\operatorname{tr}\boldsymbol{A}\boldsymbol{B}\leq\sum\lambda_{(i)}(\boldsymbol{A})\lambda_{(i)}(\boldsymbol{B})$ for the ordered eigenvalues of nonnegative definite Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is exploited in the last line, see [9], H.1.g, p.248. Equality holds if $\boldsymbol{Q}=L\boldsymbol{I}_t$, which proves optimality of $\hat{\boldsymbol{Q}}=L\boldsymbol{I}_t$. Using again the singular value decomposition of $\boldsymbol{H}$ we get

$$\begin{aligned}\log\det(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)&=\log\det(\boldsymbol{I}_r+L\boldsymbol{\Gamma})\\&=\sum_{i=1}^r\log(1+L\gamma_i)\end{aligned}$$

which completes the proof. ■

In a nice duality to Proposition 5, exchanging maximum eigenvalue and trace and keeping the right hand side unaltered, the following holds.

Proposition 7: $\hat{\boldsymbol{Q}}$ is a solution of (9) iff

$$\begin{aligned}&\operatorname{tr}\big(\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big)\\&\quad=\frac{r}{L}-\frac{1}{L}\operatorname{tr}(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}.\end{aligned}\tag{10}$$

Proof: Again we use the inequality $\operatorname{tr}\boldsymbol{A}\boldsymbol{B}\leq\sum\lambda_{(i)}(\boldsymbol{A})\lambda_{(i)}(\boldsymbol{B})$. for nonnegative definite Hermitian matrices $\boldsymbol{A}$, $\boldsymbol{B}$. Applying this inequality to (4) gives

$$\begin{aligned}&\max_{\lambda_{\max}(\boldsymbol{Q})\leq L}\ \operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{Q}\big]\\&\qquad=L\sum_{i=1}^r\lambda_i\big(\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big)\\&\qquad=L\operatorname{tr}\big(\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big).\end{aligned}$$

According to (8), the maximum value of (4) is attained at $\hat{\boldsymbol{Q}}$ with value

$$r - \operatorname{tr}(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}.$$

Dividing both sides by L concludes the proof. ∎

Optimality of $\hat{\boldsymbol{Q}} = L\boldsymbol{I}_t$ can now also be verified from (10). We start from the left hand side and conclude

$$\begin{aligned}
&\frac{1}{L}\operatorname{tr}\big(L\boldsymbol{H}^*(\boldsymbol{I}_r + L\boldsymbol{H}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\big)\\
&= \frac{1}{L}\operatorname{tr}\big(\boldsymbol{I}_r - (\boldsymbol{I}_r + L\boldsymbol{H}\boldsymbol{H}^*)^{-1}\big)\\
&= \frac{1}{L}\big(r - \operatorname{tr}(\boldsymbol{I}_r + L\boldsymbol{H}\boldsymbol{H}^*)^{-1}\big),
\end{aligned}$$

which is the right hand side of (10).

IV. Minimax and Maximin: The Worst Case Channel

In Section III the optimum power strategy $\hat{\boldsymbol{Q}}$ of the transmitter is characterized by directional derivatives of the function $\log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*)$ when the channel state is known and fixed. If total power restrictions $\operatorname{tr}\boldsymbol{Q} \le L$ apply the well known water filling solution is approved, while $\boldsymbol{Q} = L\boldsymbol{I}_t$ is the optimum strategy when maximum power is limited.

We now discuss that the channel behaves as an opponent against the optimal strategy of the transmitter. This approach has a game theoretic interpretation as a two person zero sum game where nature is playing against the transmitter as is thoroughly investigated in [5]. Bounds on the malicious behaviour are necessary because otherwise the channel would choose transmission matrix $\boldsymbol{H} = \boldsymbol{0}$ resulting in zero capacity with no transmission possible at all. Lower channel bounds on the channel matrix $\boldsymbol{H}$ are formalized by using the Loewner semi-ordering as

$$\boldsymbol{H}^*\boldsymbol{H} \ge \boldsymbol{B} \tag{11}$$

with a given nonnegative definite matrix $\boldsymbol{B}$. This is a very general approach covering a number of semi-orderings. It particularly implies minimum channel eigenvalue and general individual channel eigenvalue constraints as considered in [5] by choosing $\boldsymbol{B}$ a diagonal matrix with decreasingly ordered elements

$$\boldsymbol{B} = \operatorname{diag}\big(b_{(1)}, \dots, b_{(r)}\big),$$

since from (11) and Proposotion 2 it follows that

$$\lambda_{(i)}(\boldsymbol{H}\boldsymbol{H}^*) \ge b_{(i)}.$$

A. Transmitter moves first

The worst case move of nature against the optimal transmitter policy is described by

$$\min_{\boldsymbol{H}^*\boldsymbol{H}\ge\boldsymbol{B}} \max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*\big) \tag{12}$$

It turns out that the worst case channel is attained at $\boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{B}$ against any optimal strategy $\hat{\boldsymbol{Q}} \in \mathcal{Q}$ of the inner maximization problem. This is intuitively clear, however, a formal proof needs intriguingly deep results.

For any two complex nonnegative matrices $\boldsymbol{A} \le \boldsymbol{B}$ it holds that

$$\det\boldsymbol{A} \le \det\boldsymbol{B},$$

see [9], p. 463. This monotonicity easily carries over to

$$\begin{aligned}
\log\det\big(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*\big) &= \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{H}\big)\\
&\ge \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{B}\big)
\end{aligned}$$

for all $\boldsymbol{Q} \in \mathcal{Q}$ and $\boldsymbol{H}^*\boldsymbol{H} \ge \boldsymbol{B}$. Hence,

$$\max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{H}\big) \ge \max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{B}\big)$$

for all $\boldsymbol{H}^*\boldsymbol{H} \ge \boldsymbol{B}$, and finally

$$\begin{aligned}
&\min_{\boldsymbol{H}^*\boldsymbol{H}\ge\boldsymbol{B}} \max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{H}\big)\\
&\qquad \ge \max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{B}\big)
\end{aligned} \tag{13}$$

such that the minimum is attained at $\boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{B}$, the worst nature can do. We hence have proved the following.

Proposition 8: If $\boldsymbol{H}^*\boldsymbol{H} \ge \boldsymbol{B}$, for any set of transmitter strategies $\mathcal{Q}$ the worst channel behavior against the optimum transmitter strategy is attained at the lower bound $\boldsymbol{B}$, i.e., (13) holds.

B. Nature moves first

If nature always chooses the worst channel against any power distribution $\boldsymbol{Q} \in \mathcal{Q}$ we are faced with the reversed problem

$$\max_{\boldsymbol{Q}\in\mathcal{Q}} \min_{\boldsymbol{H}^*\boldsymbol{H}\ge\boldsymbol{B}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{H}\big). \tag{14}$$

Using the same arguments as above yields

$$\min_{\boldsymbol{H}^*\boldsymbol{H}\ge\boldsymbol{B}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{H}\big) = \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{B}\big)$$

such that (14) simplifies to

$$\begin{aligned}
&\max_{\boldsymbol{Q}\in\mathcal{Q}} \min_{\boldsymbol{H}^*\boldsymbol{H}\ge\boldsymbol{B}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{H}\big)\\
&\qquad = \max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{B}\big).
\end{aligned} \tag{15}$$

Hence, both the minimax problem (13) and maximin problem (15) have the same value. From a game theoretic point of view there exists a Nash equilibrium, a saddlepoint of the function $\log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*)$, and the value of the game is given by

$$\max_{\boldsymbol{Q}\in\mathcal{Q}} \log\det\big(\boldsymbol{I}_r + \boldsymbol{Q}\boldsymbol{B}\big). \tag{16}$$

If $\mathcal{Q} = \mathcal{Q}_{\text{tot}}$, i.e., $\operatorname{tr}\boldsymbol{Q} \le L$, the value of the game (16) is attained at $\hat{\boldsymbol{Q}}$ where the eigenvectors of $\hat{\boldsymbol{Q}}$ are aligned to the eigenvectors of $\boldsymbol{B}$ and the eigenvalues water-fill those of $\boldsymbol{B}$.

In the case of maximum power constraints by $\lambda_{\max}(\boldsymbol{Q}) \le L$ the maximum of (16) is attained at $\hat{\boldsymbol{Q}} = L\boldsymbol{I}_t$. Propositions 5 and 7 provide characterizations of both solutions, respectively.

We conclude with a numerical example. In Figure 2, $\log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*)$ is plotted as a function of the signal-to-noise ratio $\mathrm{SNR} = tL$ for different channel matrices $\boldsymbol{H}$ and different covariance matrices $\boldsymbol{Q}$ and for $r = t = 4$.

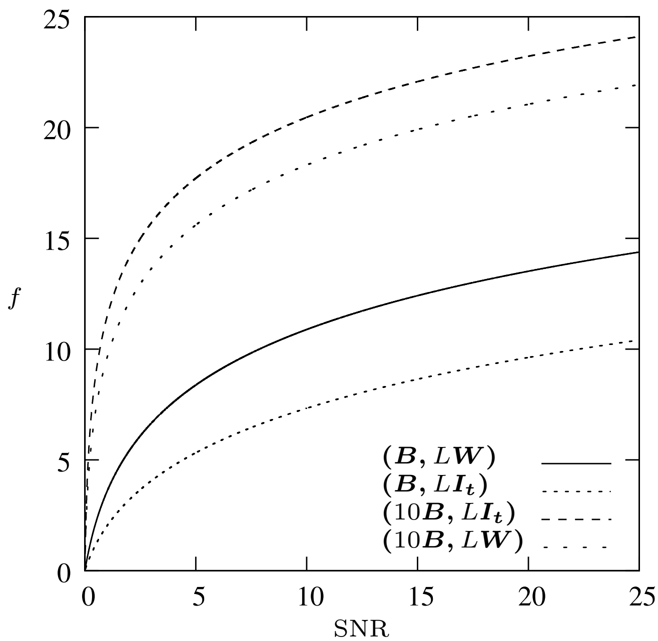

Fig. 2.
$\log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*)$ for different pairs $(\boldsymbol{H}\boldsymbol{H}^*, \boldsymbol{Q})$ as a function of $\mathrm{SNR} = tL$.

The lower two curves correspond to the worst channel, i.e. to $\boldsymbol{H}\boldsymbol{H}^* = \boldsymbol{B} = \mathrm{diag}(10, 8, 6, 4)$, where the solid line represents the solution to the maximin-problem treated in this section. We consider a second, arbitrary channel defined by $\boldsymbol{H}\boldsymbol{H}^* = 10 \cdot \boldsymbol{B}$ and the two covariance matrices $\boldsymbol{Q}_1 = L \cdot \boldsymbol{I}$ and $\boldsymbol{Q}_2 = L \cdot \mathrm{diag}(0.5, 0.2, 0.2, 0.1)$. As expected, the curve belonging to the covariance matrix $L \cdot \boldsymbol{I}$ is always the best.

V. Conclusion

The central methodology in this paper are directional derivatives of the mutual information in MIMO channels. Hereby, optimum power distributions for complete channel state information are generally characterized. This result is used to explicitly derive the capacity in the case of total and maximum power constraints, the first of which recovers the well known water filling solution. Furthermore, capacity is investigated when nature behaves as an opponent to the optimum strategy in always deteriorating the channel to a given lower bound. This approach can be interpreted as a two person zero sum game. It turns out that an equilibrium point exists whenever the worst case channel is bounded from below.

Future work will be devoted to determining optimum power distributions for general power constraints, e.g., by ℓ_p-Norms, $p \geq 1$. The extreme cases $p = 1$ and $p = \infty$ are treated in the present paper. We will moreover investigate other lower bounds on the minimum guaranteed channel quality and the uniqueness of corresponding solutions.

Acknowledgment

This work was supported by DFG grant Ma 1184/11-3.

References

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

[2] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommuncations*, vol. 10, no. 6, pp. 585–595, 1999.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Whiley, 1991.

[4] R. G. Gallager, *Information Theory and Reliable Communications*. New York: John Wiley & Sons Ltd., 1968.

[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, 2003.

[6] S. Verdu and V. Poor, "On minimax robustness: a general approach and applications," *IEEE Transactions on Information Theory*, vol. 30, pp. 328–340, 1984.

[7] N. Chiurtu, B. Rimoldi, and E. Telatar, "On the capacity of multi-antenna Gaussian channels," in *Proceedings IEEE International Symposium on Information Theory, ISIT 2001*, Washington DC, June 2001, p. 53.

[8] P. Whittle, "Some general points in the theory of optimal experimental design," *J. Roy. Statist. Soc. B*, vol. 35, pp. 123–130, 1973.

[9] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press, 1979.

[10] E. Biglieri and G. Taricco, *Transmission and Reception with Multiple Antennas: Theoretical Foundations*. Delft: now Publishers, 2004.

[11] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: John Wiley & Sons Ltd., 1988.

[12] K. Petersen and M. Pedersen, *The Matrix Cookbook*. http://2302.dk/uni/matrixcookbook.html, 2005.

[13] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia: SIAM, 2000.