Derivatives of Mutual Information in Gaussian Vector Channels with Applications

Anke Feiten
Inst. for Theor. Information Technology
RWTH Aachen University
D-52056 Aachen, Germany
feiten@ti.rwth-aachen.de

Stephen Hanly
Dept. of Electrical Engineering
The University of Melbourne
Victoria 3010, Australia
hanly@ee.unimelb.edu

Rudolf Mathar
Inst. for Theor. Information Technology
RWTH Aachen University
D-52056 Aachen, Germany
mathar@ti.rwth-aachen.de

***Abstract*— In this paper, derivatives of mutual information for a general linear Gaussian vector channel are considered. We consider two applications. First, it is shown how the corresponding gradient relates to the minimum mean squared error (MMSE) estimator and its error matrix. Secondly, we determine the directional derivative of mutual information and use this geometrically intuitive concept to characterize the capacity-achieving input distribution of the above channel subject to certain power constraints. The well-known water-filling solution is revisited and obtained as a special case. Also for shaping constraints on the maximum and the Euclidean norm of mean powers explicit solutions are derived. Moreover, uncorrelated sum power constraints are considered. The optimum input can here always be achieved by linear precoding.**

I. Introduction

In this paper, we consider linear vector channels with Gaussian noise and input distribution, in standard notation,

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{n}. \tag{1}$$

The complex $r \times t$ matrix $\boldsymbol{H}$ describes the linear transformation the signal undergoes during transmission. The random noise vector $\boldsymbol{n} \in \mathbb{C}^r$ is circularly symmetric complex Gaussian distributed (see [1]) with expectation $\mathbf{0}$ and covariance matrix $\mathrm{E}(\boldsymbol{n}\boldsymbol{n}^*) = \boldsymbol{I}_r$, denoted by $\boldsymbol{n} \sim \mathrm{SCN}(\mathbf{0}, \boldsymbol{I}_r)$. Finally, $\boldsymbol{X}$ denotes the complex zero mean input vector with covariance matrix

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

where $\boldsymbol{Q}$ may be selected from some set of nonnegative definite feasible matrices $\mathcal{Q}$. We assume complete channel state information in that $\boldsymbol{H}$ is known at the transmitter and the receiver.

An important example are multiple-input multiple-output (MIMO) channels. Seminal work in [1] and [2] has shown that the use of multiple antennas at both ends significantly increases the information-theoretic capacity in rich scattering propagation environments. Other systems can be described by the same model, e.g., CDMA systems, broadcast and multiple-access channels, as well as frequency-selective wideband channels, cf. [3], [4].

The information-theoretic capacity is given by the maximum of the mutual information as

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

over all feasible covariance matrices $\boldsymbol{Q} = \mathrm{E}(\boldsymbol{X}\boldsymbol{X}^*)$ of the input $\boldsymbol{X}$, see [1]. The diagonal elements $(q_{11}, \ldots, q_{tt})$ of $\boldsymbol{Q}$ represent the average power assigned to the transmit antennas.

Gradients and directional derivatives of mutual information are the main theme of the present paper. After introducing the general framework and notation in Section II, we establish an interesting relation between information and estimation theory in the vein of [5] in Section III. Furthermore, we explicitly determine capacity-achieving distributions, or equivalently the mean power assignment to subchannels, for four types of power constraints in Section IV.

In case of sum power constraints at the transmitter, the capacity and the associated optimum power strategy is given by the well known water-filling principle, see [1], [6], [7], [8]. By our methodology, this solution is easily verified as being optimal, and furthermore extended to shaping constraints of the following type.

Since $\max_{\boldsymbol{x}^*\boldsymbol{x}=1} \boldsymbol{x}^*\boldsymbol{Q}\boldsymbol{x} = \lambda_{\max}(\boldsymbol{Q})$ holds, see [9, p. 510], it follows that

$$\max_{1\le i\le t} q_{ii} \le \max_{1\le i\le t} \lambda_i(\boldsymbol{Q}), \tag{2}$$

where $\lambda_{\max}(\boldsymbol{Q})$ denotes the maximum eigenvalue and $\lambda_1(\boldsymbol{Q}), \cdots, \lambda_t(\boldsymbol{Q})$ the eigenvalues of $\boldsymbol{Q}$. Furthermore,

$$\sum_{i=1}^{t} q_{ii}^2 \le \sum_{i,j=1}^{t} |q_{ij}|^2 = \sum_{i=1}^{t} \lambda_i^2(\boldsymbol{Q}) \tag{3}$$

holds for the average antenna powers q_{ii}. Any upper bound on the maximum eigenvalue, or the Euclidean norm of the eigenvalues hence shapes the maximum power across antennas, or the norm of the power vector, respectively, cf. [10].

Finally, uncorrelated power assignments are investigated in Section IV-D which are appropriate when considering multiple-access channels. Most of the effort found in literature has gone into devising algorithms for numerical solutions. We characterize optimal solutions by directional derivatives and give explicit results in certain special cases, including the case of linear precoding at the transmitter.

II. Derivatives of Mutual Information

We begin by introducing the general framework and notation used throughout the paper. 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

$$\begin{aligned} Df(\hat{x},x) &= \lim_{\alpha\to 0+} \frac{1}{\alpha}\Big[f\big((1-\alpha)\hat{x}+\alpha x\big) - f(\hat{x})\Big] \\ &= \frac{d}{d\alpha} f\big((1-\alpha)\hat{x}+\alpha x\big)\Big|_{\alpha=0+}, \end{aligned} \tag{4}$$

see, e.g., [11]. 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

$$Df(\hat{x},x) = \langle \nabla f(\hat{x}), x-\hat{x}\rangle, \tag{5}$$

whenever ∇f, the derivative of f in the strong sense, exists.

Optimum points are characterized by directional derivatives as follows, for a proof see, e.g., [11].

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 $x \in \mathcal{C}$.

The geometrically intuitive concept of directional derivatives is now used for determining the capacity of vector channel (1) subject to certain power constraints. By the arguments in [1] capacity is obtained as the maximum of the mutual information over all admissible input distributions of $\boldsymbol{X}$ as

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

In the sequel we characterize the covariance matrix $\hat{\boldsymbol{Q}}$ that achieves capacity by computing the directional derivative of mutual information

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

It is well known that f is concave whenever its domain $\mathcal{Q}$ is convex such that the directional derivative exists.

Proposition 2: 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

$$Df(\hat{\boldsymbol{Q}},\boldsymbol{Q}) = \operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,(\boldsymbol{Q}-\hat{\boldsymbol{Q}})\big]. \tag{6}$$

The proof relies on the chain rule for real valued functions of matrix argument $\boldsymbol{X}$, and the fact that $\frac{d}{d\boldsymbol{X}}\det\boldsymbol{X} = (\det\boldsymbol{X})(\boldsymbol{X}^{-1})^*$, cp. [12], where $\boldsymbol{X}^*$ denotes the Hermitian of $\boldsymbol{X}$.

From (5) and Proposition 2 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 $\langle\boldsymbol{A},\boldsymbol{B}\rangle = \operatorname{tr}\boldsymbol{A}\boldsymbol{B}^*$, see [13, p. 286], amounts to

$$\nabla f(\hat{\boldsymbol{Q}}) = \boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}. \tag{7}$$

III. Application I: Gradients and Estimation

In this section we assume that $\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X}+\boldsymbol{n}$ is a linear vector channel with $\boldsymbol{X}$ and $\boldsymbol{n}$ not necessarily Gaussian. Denote by $\boldsymbol{C}$ the covariance of noise $\boldsymbol{n}$. The linear minimum mean squared error (LMMSE) estimator in this non-Gaussian case is equal to the MMSE estimator in the Gaussian case, which is given in [5], [14]. A well known consequence of the orthogonality principle (see [15]) is that the LMMSE estimator $\hat{\boldsymbol{X}}$ is given by

$$\hat{\boldsymbol{X}} = \boldsymbol{Q}\boldsymbol{H}^*\,(\boldsymbol{C}+\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*)^{-1}\,\boldsymbol{Y}$$

with error matrix

$$\mathcal{E} = \big(\boldsymbol{Q}^{-1}+\boldsymbol{H}^*\boldsymbol{C}^{-1}\boldsymbol{H}\big)^{-1}. \tag{8}$$

Leaving the non-Gaussian case and turning back to model (1), we see that matrix $\boldsymbol{Q} - \boldsymbol{Q}\ \nabla f(\boldsymbol{Q})\ \boldsymbol{Q}$ plays an important role in this framework.

Proposition 3: The MMSE estimator and its error matrix are related to the derivative of the mutual information by

$$\begin{aligned} \hat{\boldsymbol{X}} &= \big(\boldsymbol{Q} - \boldsymbol{Q}\,\nabla f(\boldsymbol{Q})\,\boldsymbol{Q}\big)\boldsymbol{H}^*\boldsymbol{Y}, \\ \mathcal{E} &= \boldsymbol{Q} - \boldsymbol{Q}\,\nabla f(\boldsymbol{Q})\,\boldsymbol{Q}. \end{aligned}$$

Proof: The representation of $\mathcal{E}$ is direct from (8). Using the formula $(\boldsymbol{I}+\boldsymbol{A})^{-1}\boldsymbol{A} = \boldsymbol{I}-(\boldsymbol{I}+\boldsymbol{A})^{-1}$ with $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*$ gives

$$\begin{aligned} &\boldsymbol{Q}\,\nabla f(\boldsymbol{Q})\ \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{Y} \\ &\quad = \boldsymbol{Q}\boldsymbol{H}^*[\boldsymbol{I}+\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*]^{-1}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{Y} \\ &\quad = \boldsymbol{Q}\boldsymbol{H}^*\big(\boldsymbol{I} - [\boldsymbol{I}+\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*]^{-1}\big)\boldsymbol{Y} \\ &\quad = \boldsymbol{Q}\boldsymbol{H}^*\boldsymbol{Y} - \hat{\boldsymbol{X}} \end{aligned}$$

which completes the proof. ■

A related connection between derivatives of mutual information with respect to $\boldsymbol{H}$ and the MMSE estimator is presented in the recent work [5]. Fundamental connections between the derivative of the mutual information w.r.t. the SNR and the MMSE are derived in [16].

IV. Application II: Achieving Capacity

Achieving capacity with an appropriate power distribution means to maximize $f(\boldsymbol{Q})$ over the set of possible power assignments $\mathcal{Q}$. According to Proposition 1 some matrix $\hat{\boldsymbol{Q}}$ maximizes $f(\boldsymbol{Q})$ over some convex set $\mathcal{Q}$ if and only if $Df(\hat{\boldsymbol{Q}},\boldsymbol{Q}) \leq 0$ for all $\boldsymbol{Q} \in \mathcal{Q}$. By (6) this leads to

$$\begin{aligned} &\operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\boldsymbol{Q}\big] \\ &\leq \operatorname{tr}\big[\boldsymbol{H}^*(\boldsymbol{I}_r+\boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{H}\,\hat{\boldsymbol{Q}}\big] \end{aligned} \tag{9}$$

for all $\boldsymbol{Q} \in \mathcal{Q}$. Hence, we obtain the following new characterization of capacity-achieving covariance matrices.

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

$$\max_{\boldsymbol{Q}\in\mathcal{Q}}\ \operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\,\boldsymbol{Q}\big]. \tag{10}$$

The main value of this equivalence is that once a candidate for maximizing $f(\boldsymbol{Q})$ over $\mathcal{Q}$ is claimed it can be verified by merely checking the simple linear inequality (9). Moreover, in concrete cases explicitly evaluable representation of the optimum are obtained, as is developed in the following.

A. Sum Power Constraints

We first investigate total power constraints

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

where $\boldsymbol{Q} \geq 0$ denotes that $\boldsymbol{Q}$ is Hermitian and nonnegative definite. Obviously, the set $\mathcal{Q}_{\mathrm{tot}}$ is convex. Covariance matrices $\boldsymbol{Q}$ which achieve capacity, i.e., solutions of

$$\max_{\operatorname{tr} \boldsymbol{Q} \leq L} \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*) \tag{11}$$

are now characterized as follows.

Proposition 5: $\hat{\boldsymbol{Q}}$ is a solution of (11) if and only if

$$\lambda_{\max}\big(\nabla f(\hat{\boldsymbol{Q}})\big) = \frac{r}{L} - \frac{1}{L}\operatorname{tr}(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}. \tag{12}$$

Proof: On one hand

$$\max_{\operatorname{tr} \boldsymbol{Q} \leq L} \operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\boldsymbol{Q}\big] = L\,\lambda_{\max}\big(\nabla f(\hat{\boldsymbol{Q}})\big),$$

and on the other

$$\operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\hat{\boldsymbol{Q}}\big] = r - \operatorname{tr}(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}$$

holds. The assertion thus follows from (10). ■

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

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

with the singular value decomposition $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Gamma}^{1/2}\boldsymbol{V}^*$ actually satisfies condition (12). Here, ν is defined as the water level above the inverse positive eigenvalues γ_i of $\boldsymbol{H}^*\boldsymbol{H}$ defined by $\sum_{i:\gamma_i>0}(\nu - \gamma_i^{-1})^+ = L$.

Using $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Gamma}^{1/2}\boldsymbol{V}^*$ and the optimal solution $\hat{\boldsymbol{Q}}_{\mathrm{wf}}$ it is straightforward to show that $\lambda_{\max}\big(\nabla f(\hat{\boldsymbol{Q}})\big) = \frac{1}{\nu}$. On the other hand, using the same decomposition, after some algebra we obtain $\operatorname{tr}\big[\boldsymbol{I}_r - (\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}_{\mathrm{wf}}\boldsymbol{H}^*)^{-1}\big] = \frac{L}{\nu}$, which verifies (12).

B. Maximum Eigenvalue Constraints

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

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

By inequality (2) this limits the maximum power across antennas, hence forming a peak power constraint.

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

$$\max_{\lambda_{\max}(\boldsymbol{Q}) \leq L} \log\det\big(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*\big). \tag{13}$$

Proposition 6: The maximum in (13) is attained at $\hat{\boldsymbol{Q}} = L\boldsymbol{I}_t$ with value $\sum_{i=1}^r \log(1 + L\gamma_i)$, where γ_i, $i = 1, \ldots, r$, 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 $\boldsymbol{H}$. An upper bound for maximization problem (10) with $\hat{\boldsymbol{Q}} = L\boldsymbol{I}_t$ is derived as

$$\max_{\lambda_{\max}(\boldsymbol{Q}) \leq L} \operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\boldsymbol{Q}\big] \leq \sum_{i=1}^r \frac{L\gamma_i}{1 + L\gamma_i},$$

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 $\boldsymbol{B}$ has been used, see [9, H.1.g, p.248]. Equality holds if $\boldsymbol{Q} = L\boldsymbol{I}_t$, which proves optimality of $\hat{\boldsymbol{Q}} = L\boldsymbol{I}_t$.

The optimum value itself is now easily determined by inspection. ■

C. Sum of Squared Power Constraints

We consider constraints of the form

$$\mathcal{Q}_{\mathrm{squ}} = \big\{\boldsymbol{Q} = (q_{ij})_{i,j=1,\ldots,t} \geq 0 \mid \big(\sum_{i,j=1}^t |q_{ij}|^2\big)^{1/2} \leq L\big\}.$$

By inequality (3) the above L is also an upper bound on the Euclidean norm of mean powers across antennas, which may be interpreted as approximating simultaneous peak and sum power constraints by a single constraint. Furthermore, it holds that

$$\sum_{i,j=1}^t q_{ij}^2 = \operatorname{tr}\big[\boldsymbol{Q}^2\big] = \sum_{i=1}^t \lambda_i^2(\boldsymbol{Q}) = \|\boldsymbol{Q}\|_2^2,$$

where $\|\boldsymbol{Q}\|_2 = \big(\sum_{i,j=1}^t q_{ij}^2\big)^{1/2}$ denotes the ℓ_2-norm in the Hilbert space of all $t \times t$ Hermitian matrices with inner product $\langle \boldsymbol{A}, \boldsymbol{B}\rangle = \operatorname{tr}[\boldsymbol{A}\boldsymbol{B}]$. Hence, $\mathcal{Q}_{\mathrm{squ}}$ may be written as

$$\mathcal{Q}_{\mathrm{squ}} = \big\{\boldsymbol{Q} = (q_{ij})_{i,j=1,\ldots,t} \geq 0 \mid \big(\sum_{i=1}^t \lambda_i^2(\boldsymbol{Q})\big)^{1/2} \leq L\big\}.$$

Applying the Cauchy-Schwarz inequality to (10) gives

$$\begin{aligned}\max_{\boldsymbol{Q}\in\mathcal{Q}_{\mathrm{squ}}} \operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\,\boldsymbol{Q}\big] &\leq \|\nabla f(\hat{\boldsymbol{Q}})\|_2 \max_{\boldsymbol{Q}\in\mathcal{Q}_{\mathrm{squ}}} \|\boldsymbol{Q}\|_2 \\ &= L\,\|\nabla f(\hat{\boldsymbol{Q}})\|_2,\end{aligned}$$

and the maximum is attained iff $\boldsymbol{Q} = \frac{L}{\|\nabla f(\hat{\boldsymbol{Q}})\|_2}\nabla f(\hat{\boldsymbol{Q}})$.

Hence, by (10), $\hat{\boldsymbol{Q}}$ is an optimum argument over $\mathcal{Q}_{\mathrm{squ}}$ if and only if

$$L\,\|\nabla f(\hat{\boldsymbol{Q}})\|_2 = \operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\,\hat{\boldsymbol{Q}}\big]. \tag{14}$$

We seek a solution of (14) in the set of power assignments $\boldsymbol{Q}$ of the form

$$\boldsymbol{Q} = \boldsymbol{V}\operatorname{diag}(q_1, \ldots, q_t)\boldsymbol{V}^*,$$

with $\boldsymbol{V}$ from the singular value decomposition $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Gamma}^{1/2}\boldsymbol{V}$ of $\boldsymbol{H}$. We will show that a solution exists already in this subset of $\mathcal{Q}_{\mathrm{squ}}$, which solves the problem. Let γ_i denote the positive eigenvalues of $\boldsymbol{H}^*\boldsymbol{H}$ and $\boldsymbol{H}\boldsymbol{H}^*$, respectively, augmented by zeros whenever appropriate. After some tedious but elementary matrix algebra equation (14) reads as

$$L\Big(\sum_{i=1}^t \big(\frac{\gamma_i}{1 + \gamma_i\hat{q}_i}\big)^2\Big)^{1/2} = \sum_{i=1}^t \frac{\gamma_i\hat{q}_i}{1 + \gamma_i\hat{q}_i}.$$

Again by the Cauchy-Schwarz inequality, the left hand side of the above is greater than or equal to the right hand side for any $\hat{\boldsymbol{Q}} \in \mathcal{Q}_{\text{squ}}$, with equality if and only if

$$\hat{q}_i = \frac{\alpha\gamma_i}{1+\gamma_i\hat{q}_i},\ i = 1,\ldots,t,$$

for some $\alpha > 0$ such that $\sum_i \hat{q}_i^2 = L^2$. This is a quadratic equation in q_i with the unique solution

$$\begin{aligned}
&\hat{q}_i = 0,\ \text{if } \gamma_i = 0,\\
&\hat{q}_i = \sqrt{\frac{1}{4\gamma_i^2}+\alpha} - \frac{1}{2\gamma_i},\ \text{if } \gamma_i > 0,\\
&\alpha \text{ such that } \sum_{i=1}^t \hat{q}_i^2 = L^2.
\end{aligned} \tag{15}$$

Because of monotonicity in α the solutions $\hat{q}_i$ can be easily determined numerically.

In summary, we have obtained an explicit solution to

$$\max_{\boldsymbol{Q}\in\mathcal{Q}_{\text{squ}}} \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*) \tag{16}$$

as follows.

Proposition 7: $\hat{\boldsymbol{Q}}$ is a solution of (16) if and only if

$$\hat{\boldsymbol{Q}} = \boldsymbol{V}\operatorname{diag}(\hat{q}_1,\ldots,\hat{q}_t)\boldsymbol{V}^*$$

with $\hat{q}_i$ given by (15).

D. Uncorrelated Sum Power Constraints

We next consider total power constraints that arise in a multiple access channel in which only limited communication is possible between users. We assume that the total transmit power of all users is constrained (which does imply limited coordination, perhaps via the base station) so that the constraining set is

$$\mathcal{Q}_{\text{le}} = \big\{\boldsymbol{Q} = \operatorname{diag}(q_1,\ldots,q_t) \mid q_i \ge 0, \sum_{i=1}^t q_i \le L\big\}.$$

The key difference between the multiple access setup and the general MIMO setup is that here $\boldsymbol{Q}$ is constrained to be diagonal, which precludes coordination between the users at the data symbol rate. We aim at finding the solution of

$$\begin{aligned}
&\max_{\boldsymbol{Q}\in\mathcal{Q}_{\text{le}}} \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*)\\
&\quad = \max_{q_i\ge 0, \sum_i q_i \le L} \log\det\Big(\boldsymbol{I}_r + \sum_{i=1}^t q_i\boldsymbol{h}_i\boldsymbol{h}_i^*\Big),
\end{aligned} \tag{17}$$

where $\boldsymbol{h}_i$ denotes the ith column of $\boldsymbol{H}$.

The same problem is encountered when determining the sum capacity of a MIMO broadcast channel with t antennas at the transmitter and a single antenna at each of r receivers. This capacity coincides with the sum rate capacity of the dual MIMO multiple access channel, a duality theory developed in [3], [4], [17], [18]. In the work [19] algorithmic approaches are discussed and developed for the broadcast channel with scalars q_i substituted by positive definite matrices.

Another very related paper is [20], which considers water-filling in flat fading CDMA multiple access channel. This paper differs from the present one in that it considers an ergodic power constraint (*i.e.,* the average is taken over time, as the fading coefficients change), the model is for CDMA (rather than MIMO) and the numerical solution emerges in the limit as the number of users, and spreading factor, grow to infinity. Nevertheless, our derivation of (19) below provides an alternative approach to the derivation of (23) in [20].

A general class of algorithms for maximizing mutual information subject to power constraints based on interior point methods is designed in [21]. Such general algorithms can be applied to problem (17), but in the present paper we wish to identify structural results for the particular problem at hand, and identify connections to the directional derivatives that are the main topic of the present paper.

The maximum in (17) is attained at the boundary of $\mathcal{Q}_{\text{le}}$, i.e, in the convex set

$$\mathcal{Q}_{\text{eq}} = \big\{\boldsymbol{Q} = \operatorname{diag}(q_1,\ldots,q_t) \mid q_i \ge 0, \sum_{i=1}^t q_i = L\big\}.$$

This follows easily from the fact that $\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}_1\boldsymbol{H}^* \le \boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}_2\boldsymbol{H}^*$ whenever $\boldsymbol{Q}_1 \le \boldsymbol{Q}_2$, and the monotonicity of $\log\det\boldsymbol{A}$ on the set of nonnegative definite Hermitian matrices $\boldsymbol{A}$, see [9, F.2.c., p. 476].

Hence, in the sequel we deal with the problem

$$\max_{q_i\ge 0, \sum_i q_i = L} \log\det\Big(\boldsymbol{I}_r + \sum_{i=1}^t q_i\boldsymbol{h}_i\boldsymbol{h}_i^*\Big). \tag{18}$$

By (10) some power allocation $\hat{\boldsymbol{Q}} = \operatorname{diag}(\hat{q}_1,\ldots,\hat{q}_t)$ solves (18) if and only if it is a solution of

$$\begin{aligned}
&\max_{\boldsymbol{Q}\in\mathcal{Q}_{\text{eq}}} \operatorname{tr}\big[\nabla f(\hat{\boldsymbol{Q}})\boldsymbol{Q}\big]\\
&\quad = \max_{q_i\ge 0, \sum_i q_i = L} \sum_{i=1}^t q_i\boldsymbol{h}_i^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{h}_i,
\end{aligned}$$

where $\boldsymbol{h}_i$ denotes the i-th column of $\boldsymbol{H}$.

In the following we assume that none of the $\boldsymbol{h}_i$ equals zero, thereby excluding the case that some of the transmit antennas are irrelevant. Obviously,

$$\begin{aligned}
&\max_{q_i\ge 0, \sum_i q_i = L} \sum_{i=1}^t q_i\boldsymbol{h}_i^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{h}_i\\
&\quad \le L \max_{1\le i\le t} \big\{\boldsymbol{h}_i^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{h}_i\big\}.
\end{aligned}$$

Equality holds if the quadratic forms all have the same value, i.e., if there exists some λ such that

$$\boldsymbol{h}_i^*(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}\boldsymbol{h}_i = \lambda \tag{19}$$

for all $i = 1,\ldots,t$. Since $\boldsymbol{h}_i \ne \boldsymbol{0}$ and since $(\boldsymbol{I}_r + \boldsymbol{H}\hat{\boldsymbol{Q}}\boldsymbol{H}^*)^{-1}$ is positive definite, $\lambda > 0$ follows.

Condition (19) is also necessary for an extreme point of $f\big(\operatorname{diag}(q_1,\ldots,q_t)\big)$ subject to $\sum_i q_i = L$, as may be easily seen from a Lagrangian setup. Additionally making use of convexity we have thus proved the following result.

Proposition 8: If some $\hat{\boldsymbol{Q}} = \operatorname{diag}(\hat{q}_1, \ldots, \hat{q}_t) \in \mathcal{Q}_{\text{eq}}$ satisfies (19), then $\hat{\boldsymbol{Q}}$ is a solution of

$$\max_{\boldsymbol{Q} \in \mathcal{Q}_{\text{eq}}} \log\det(\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^*).$$

On the other hand, any maximizing point $\hat{\boldsymbol{Q}}$ of $\log\det(\boldsymbol{I}_r + \boldsymbol{H}\operatorname{diag}(q_1, \ldots, q_t)\boldsymbol{H}^*)$ subject to $\sum_i q_i = L$ satisfies (19).

Note that in the second part of the above Proposition we have omitted the constraints $q_i \geq 0$ such that physically meaningful solutions are obtained from (19) only if the solution has nonnegative components.

The central condition (19) can be further evaluated. Denote by matrix

$$\boldsymbol{A}_i = \boldsymbol{I}_r + \sum_{j \neq i} q_j \boldsymbol{h}_j \boldsymbol{h}_j^*.$$

Applying the matrix inversion lemma to (19), namely to $\boldsymbol{h}_i^* \big(\boldsymbol{A}_i + q_i \boldsymbol{h}_i \boldsymbol{h}_i^*\big)^{-1} \boldsymbol{h}_i = \lambda$ gives

$$\boldsymbol{h}_i^* \Big(\boldsymbol{A}_i^{-1} - \frac{q_i \boldsymbol{A}_i^{-1} \boldsymbol{h}_i \boldsymbol{h}_i^* \boldsymbol{A}_i^{-1}}{1 + q_i \boldsymbol{h}_i^* \boldsymbol{A}_i^{-1} \boldsymbol{h}_i}\Big) \boldsymbol{h}_i = \lambda$$

such that (19) is equivalent to

$$q_i = \frac{1}{\lambda} - \frac{1}{\boldsymbol{h}_i^* \boldsymbol{A}_i^{-1} \boldsymbol{h}_i}, \quad i = 1, \ldots, t. \tag{20}$$

From the constraints, $\sum_i q_i = L$ is required. Equating the sum over the right hand side of (20) to L yields

$$\frac{1}{\lambda} = \frac{L}{t} + \frac{1}{t} \sum_{i=1}^{t} \frac{1}{\boldsymbol{h}_i^* \boldsymbol{A}_i^{-1} \boldsymbol{h}_i}.$$

This makes (20) a fixed point equation which may be used for deriving iterative schemes. Any fixed point is an optimum solution to (17) provided it has nonnegative components.

In the following special case we get an explicit result. Assume that the columns of $\boldsymbol{H} = (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_t)$ are pairwise orthogonal, i.e., $\boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{I}_t$. Then applying the generalized matrix inversion lemma, cf. [13, p.124],

$$(\boldsymbol{A} + \boldsymbol{C}\boldsymbol{D}^*)^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{C}(\boldsymbol{I} + \boldsymbol{D}^*\boldsymbol{A}^{-1}\boldsymbol{C})^{-1}\boldsymbol{D}^*\boldsymbol{A}^{-1},$$

and setting $\boldsymbol{Q}_i = \operatorname{diag}(q_1, \ldots, q_{i-1}, 0, q_{i+1}, \ldots, q_t)$ yields

$$\begin{aligned} &\boldsymbol{h}_i^* \boldsymbol{A}_i^{-1} \boldsymbol{h}_i = \boldsymbol{h}_i^* (\boldsymbol{I}_r + \boldsymbol{H}\boldsymbol{Q}_i\boldsymbol{H}^*)^{-1} \boldsymbol{h}_i \\ &= \boldsymbol{h}_i^* \big(\boldsymbol{I}_r - \boldsymbol{H}\boldsymbol{Q}_i^{1/2}(\boldsymbol{I} + \boldsymbol{Q}_i^{1/2}\boldsymbol{H}^*\boldsymbol{H}\boldsymbol{Q}_i^{1/2})^{-1}\boldsymbol{Q}_i^{1/2}\boldsymbol{H}^*\big)\boldsymbol{h}_i \\ &= \boldsymbol{h}_i^* \boldsymbol{h}_i - 0 = 1 \end{aligned}$$

In summary, it follows that

$$q_i = \frac{L}{t} + 1 - 1 = \frac{L}{t} \quad \text{for all } i = 1, \ldots, t,$$

i.e., all the powers q_i of an optimal solution are equal in this case.

Orthogonality of the columns of $\boldsymbol{H}$ can be achieved by linear precoding, see [22]. Write $\boldsymbol{H} = \tilde{\boldsymbol{H}}\boldsymbol{B}$, where the $r \times t$ matrix $\tilde{\boldsymbol{H}}$ of rank t describes the physical channel and $\boldsymbol{B}$ a $t \times t$ precoding matrix. It holds that

$$\boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{B}^*\tilde{\boldsymbol{H}}^*\tilde{\boldsymbol{H}}\boldsymbol{B} = \boldsymbol{B}^*\boldsymbol{T}\operatorname{diag}(\gamma_i)\boldsymbol{T}^*\boldsymbol{B}$$

for some unitary matrix $\boldsymbol{T}$ and diagonal matrix of positive eigenvalues γ_i, $i = 1, \ldots, t$. Choosing $\boldsymbol{B} = \boldsymbol{T}\operatorname{diag}(\gamma_i^{-1/2})$ yields $\boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{I}_t$, as desired. While such precoding is itself suboptimal, in terms of the criteria expressed in (17), it has the advantage that the signal processing at the receiver is much simplified, and this signal processing needs to occur at the symbol rate. Given such precoding, an equal power allocation is then optimal, as shown above.

References

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

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

[3] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[4] W. Yu and J. Cioffi, "Sum capacity of gaussian vector broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, September 2004.

[5] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 141–154, January 2006.

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

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

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

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

[10] D. P. Palomar, "Unified framework for linear MIMO transceivers with shaping constraints," *IEEE Communications Letters*, vol. 8, no. 12, pp. 697–699, December 2004.

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

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

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

[14] D. Tse and S. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, March 1999.

[15] H. V. Poor, *An introduction to signal detection and estimation*. New York: Springer Verlag, 1994.

[16] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[17] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2002.

[18] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, October 2003.

[19] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1570–1580, April 2005.

[20] P. Viswanath, D. Tse, and V. Anantharam, "Asymptotically optimal water-filling in vector multiple-access channels," *IEEE Transactions on Information Theory*, vol. 47, no. 1, pp. 241–267, January 2001.

[21] L. Vandenberghe, S. Boyd, and S. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 19, no. 2, pp. 499–533, April 1998.

[22] K. Lee and D. Peterson, "Optimal linear coding for vector channels," *IEEE Transactions on Communications*, vol. 24, no. 12, pp. 1283–1290, December 1976.