Zero-Forcing Precoding and Generalized Inverses

Ami Wiesel, *Student Member, IEEE*, Yonina C. Eldar, *Senior Member, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*—We consider the problem of linear zero-forcing precoding design and discuss its relation to the theory of generalized inverses in linear algebra. Special attention is given to a specific generalized inverse known as the pseudo-inverse. We begin with the standard design under the assumption of a total power constraint and prove that precoders based on the pseudo-inverse are optimal among the generalized inverses in this setting. Then, we proceed to examine individual per-antenna power constraints. In this case, the pseudo-inverse is not necessarily the optimal inverse. In fact, finding the optimal matrix is nontrivial and depends on the specific performance measure. We address two common criteria, fairness and throughput, and show that the optimal generalized inverses may be found using standard convex optimization methods. We demonstrate the improved performance offered by our approach using computer simulations.**

***Index Terms*—Beamforming, generalized inverses, per-antenna constraints, semidefinite relaxation, zero-forcing precoding.**

I. Introduction

TRANSMITTER design for the multiple-input single-output (MISO) multiuser broadcast channel is an important problem in modern wireless communication systems. The main difficulty in this channel is that coordinated receive processing is not possible and that all the signal processing must be employed at the transmitter side. From an information theory perspective, the capacity region of this channel was only recently characterized [1]. From a signal processing point of view, there are still many open questions and there is ongoing search aimed at finding efficient yet simple transmitter design algorithms. In particular, linear precoding schemes which seem to provide a promising tradeoff between performance and complexity were proposed in [2]–[5].

The most common linear precoding scheme is zero-forcing (ZF) beamforming. It is a suboptimal approach that attracted considerable attention since there are computational difficulties even within the class of linear precoding strategies. For example, we are not aware of any efficient techniques for maximizing throughput using linear beamforming. Instead, ZF is a simple method which decouples the multiuser channel into multiple independent subchannels and reduces the design to a power allocation problem. It performs very well in the high signal-to-noise-ratio (SNR) regime or when the number of users is sufficiently large, and is known to provide full degrees of freedom [1]. Moreover, it is easy to generalize this method to incorporate nonlinear dirty paper coding (DPC) mechanisms [1]. There are dozens of papers on ZF precoding focusing on different design criteria [4], [6]–[11]. Among these, two common criteria are maximal fairness and maximum throughput. Due to its simplicity, ZF precoding is also an appealing transmission method in multiple-input multiple-output (MIMO) broadcast channels [12]–[17].

Traditionally, the transmitter is designed under the assumption of a total power constraint [1]–[11]. In practice, there is increasing interest in addressing more complicated scenarios, such as individual per-antenna power constraints. These are more realistic since each transmit antenna has its own power amplifier. Moreover, state-of-the-art communication systems will utilize multiple transmitters, which are geographically separated, but cooperatively send data to the receiving units. In such systems, it is clear that each transmitter will have its own power restrictions. Single-user transmit beamforming in this setting is addressed in [18]. Our work on linear beamforming for multiuser systems [2] was generalized to incorporate per-antenna power constraints in [19]. ZF precoding methods were also extended to deal with individual restrictions [20]–[22].

Interestingly, ZF precoding design is highly related to the concept of generalized inverses in linear algebra [23]. This is easy to understand as the ZF precoder basically inverts the multiuser channel. Previous works using total power constraints [4], [6]–[11] as well as individual per-antenna power constraints [20]–[22] began with the assumption that the precoder has the form of a specific generalized inverse known as the pseudo-inverse. We prove that the pseudo-inverse-based precoder is optimal among the generalized inverses for maximizing any performance measure under a total power constraint. However, when per-antenna power constraints are involved, it is no longer optimal and other inverses may outperform it. Finding the optimal matrix is nontrivial and depends on the specific performance criterion. We consider the two classical criteria, fairness and throughput, and transform the design problems into convex optimization programs which can be solved efficiently using off-the-shelf numerical packages.

The ZF precoding design for maximizing throughput turns out to be a nonconvex optimization problem. One of the methods for handling such problems is to lift it into a higher dimension and then relax the nonconvex constraints. Consequently, there

Manuscript received February 20, 2007; revised January 24, 2008. Published August 13, 2008 (projected). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Timothy N. Davidson. This work was supported by the EU 6/7th framework program, via the NEWCOM/NEWCOM++ network of excellence, by the Israel Science Foundation, and by the Glasberg-Klein Research Fund.

Some of the results in this paper were presented in the Forty-First Conference on Information Sciences and Systems (CISS), The Johns Hopkins University, Baltimore, MD, March 14–16, 2007.

A. Wiesel is with the Technion—Israel Institute of Technology, Haifa 32000, Israel, and also with the Department of Electrical Engineering and Computer Science, The College of Engineering, The University of Michigan, Ann Arbor, MI 48109-2122 USA (e-mail: amiw@tx.technion.ac.il; amiw@umich.edu).

Y. C. Eldar and S. Shamai (Shitz) are with the Technion—Israel Institute of Technology, Haifa, Israel (e-mail: yonina@ee.technion.ac.il; sshlomo@ee.technion.ac.il).

Digital Object Identifier 10.1109/TSP.2008.924638

is an increasing interest in analyzing the tightness of such relaxations [24], [25]. In the context of transmit beamforming, semidefinite relaxation and its tightness have been addressed in [26]–[29]. However, these works do not consider per-antenna power constraints nor the zero-forcing assumption. We apply this method to the problem at hand and use Lagrange duality to prove that the relaxation is always tight in our setting.

The paper is organized as follows. In Section II, we introduce the ZF precoding design problem. A brief review of generalized inverses is provided in Section III. Next, precoding under total power constraint is addressed in Section IV, whereas precoding under individual per-antenna power constraints is considered in Section V. A few numerical results are demonstrated in Section VI.

The following notation is used. Boldface upper case letters denote matrices, boldface lower case letters denote column vectors, and standard lower case letters denote scalars. The superscripts $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^-$, and $(\cdot)^\dagger$ denote the transpose, the conjugate transpose, matrix inverse, generalized inverse and pseudo-inverse, respectively. The operators $\mathrm{Tr}\left\{\cdot\right\}$, $\|\cdot\|$ and $\|\cdot\|_F$ denote the trace, the Euclidean norm, and the Frobenius norm, respectively. The operators $\mathrm{diag}\left\{\mathbf{d}\right\}$ and $\mathrm{diag}\left\{d_k\right\}$ denote a diagonal matrix with the elements $\mathbf{d}$ and d_k, respectively. The matrix $\mathbf{I}$ denotes the identity matrix, $\mathbf{1}$ is the vector of ones, and $\mathbf{e}_k$ is a zeros vector with a one in the kth element. The operators $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary parts, respectively. Finally, $\mathbf{X} \succeq \mathbf{0}$ means that $\mathbf{X}$ is positive semidefinite.

II. Problem Formulation

We consider the standard MISO multiuser broadcast channel

$$y_k = \mathbf{h}_k^H \mathbf{x} + w_k, \quad k = 1, \ldots, K \tag{1}$$

where y_k is the received sample of the k'th user, $\mathbf{h}_k$ is the length N channel to this user, $\mathbf{x}$ is the length N transmitted vector and w_k are zero mean and unit variance complex Gaussian noise samples. For simplicity, we use the following matrix notation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{2}$$

where $\mathbf{y} = [y_1, \ldots, y_K]^T$, $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]^H$ and $\mathbf{w} = [w_1, \ldots, w_K]^T$. Throughout the paper, we will assume that $K \leq N$ and $\mathbf{H}$ is full row-rank.

In linear precoding methods, the transmitted vector is a linear transformation of the information symbols (see Fig. 1)

$$\mathbf{x} = \mathbf{T}\mathbf{s} \tag{3}$$

where the length K information vector $\mathbf{s}$ satisfies $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$. The precoding matrix $\mathbf{T}$ is then designed to maximize some performance measure. Typical metrics involve functions of the received signal-to-interference-plus-noise ratios (SINRs):

$$p_k = \frac{\left|[\mathbf{HT}]_{k,k}\right|^2}{\sum_{j \neq k} \left|[\mathbf{HT}]_{k,j}\right|^2 + 1}, \quad k = 1, \ldots, K. \tag{4}$$

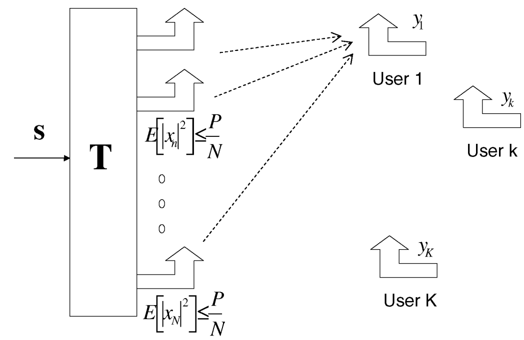

Fig. 1. ZF precoding with per-antenna power constraints.

Direct formulations of design problems incorporating such measures usually lead to intractable optimization problems. ZF precoding is a standard suboptimal approach which is known to provide a promising tradeoff between complexity and performance. Here, $\mathbf{T}$ is designed to achieve zero interference between the users, i.e., $[\mathbf{HT}]_{k,j} = 0$ if $k \neq j$. Moreover, without loss of generality, we assume that $\mathrm{Re}\{[\mathbf{HT}]_{k,k}\} \geq 0$ and $\mathrm{Im}\{[\mathbf{HT}]_{k,k}\} = 0$ for $k = 1, \ldots, K$. Using matrix notation, the ZF condition is equivalent to

$$\mathbf{HT} = \mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\} \tag{5}$$

where $\sqrt{\mathbf{p}} = \left[\sqrt{p_1}, \ldots, \sqrt{p_K}\right]^T$ is a vector with real non-negative elements. These restrictions simplify the design and decouple the broadcast channel into K independent scalar subchannels

$$y_k = \sqrt{p_k} s_k + w_k, \quad k = 1, \ldots, K. \tag{6}$$

Traditionally, precoders are designed subject to a total power constraint of the form

$$E\{\|\mathbf{x}\|^2\} = \mathrm{Tr}\left\{\mathbf{T}\mathbf{T}^H\right\} = \|\mathbf{T}\|_F^2 \leq P \tag{7}$$

where $P > 0$. As we will show in the next sections, the total power constraint simplifies the design problem and leads to simple and efficient precoders. Nonetheless, in practice, many systems are subject to individual per antenna power constraints as illustrated in Fig. 1

$$E\{|x_n|^2\} = \left[\mathbf{T}\mathbf{T}^H\right]_{n,n} \leq \frac{P}{N}, \qquad n = 1, \ldots, N. \tag{8}$$

In order to properly formulate the design problem we need to define its objective. Depending on the application, different criteria may be considered. Two typical performance measures are as follows:

- Fairness: $f(\mathbf{p}) = \min_k p_k$;
- Throughput: $f(\mathbf{p}) = \sum_k \log\left(1 + p_k\right)$.

Therefore, we treat two fundamental design problems. In Section IV, we consider the optimal $\mathbf{T}$ for maximizing $f(\mathbf{p})$ subject to the zero-forcing constraint and a total power constraint. In Section V, we generalize the setting to individual per-antenna power constraints. Both fairness and throughput are addressed in the two problems.

III. Generalized Inverses

The ZF precoding design problem is closely related to the concept of generalized inverses in linear algebra [23], [30]. Therefore, we now briefly review this topic.

Formally, the generalized inverse of a size $K \times N$ matrix $\mathbf{H}$ is any matrix $\mathbf{H}^{-}$ of size $N \times K$ such that $\mathbf{H}\mathbf{H}^{-}\mathbf{H} = \mathbf{H}$. If $\mathbf{H}$ is square and invertible, then $\mathbf{H}^{-} = \mathbf{H}^{-1}$. Otherwise, the generalized inverse is not unique. The pseudo-inverse $\mathbf{H}^{\dagger}$ is a specific generalized inverse that satisfies $\mathbf{H}\mathbf{H}^{\dagger}\mathbf{H} = \mathbf{H}$, $\mathbf{H}^{\dagger}\mathbf{H}\mathbf{H}^{\dagger} = \mathbf{H}^{\dagger}$, $(\mathbf{H}^{\dagger}\mathbf{H})^{H} = \mathbf{H}^{\dagger}\mathbf{H}$ and $(\mathbf{H}\mathbf{H}^{\dagger})^{H} = \mathbf{H}\mathbf{H}^{\dagger}$. It is unique and is known to have minimal Frobenius norm among all the generalized inverses.

In this paper, we assume that $\mathbf{H}$ is a full row-rank matrix. Under this assumption, the generalized inverse is any matrix $\mathbf{H}^{-}$ such that $\mathbf{H}\mathbf{H}^{-} = \mathbf{I}$. The pseudo-inverse is given by $\mathbf{H}^{\dagger} = \mathbf{H}^{H}(\mathbf{H}\mathbf{H}^{H})^{-1}$ and any generalized inverse may be expressed as

$$\mathbf{H}^{-} = \mathbf{H}^{\dagger} + \mathbf{P}_{\perp}\mathbf{U} \tag{9}$$

where $\mathbf{P}_{\perp} = \mathbf{I} - \mathbf{H}^{\dagger}\mathbf{H}$ is the orthogonal projection onto the null space of $\mathbf{H}$ and $\mathbf{U}$ is an arbitrary matrix.

Using the above definitions and properties, it is easy to see the relation between ZF precoding and generalized inverses. Due to (5), the general structure of any ZF precoder is

$$\mathbf{T} = \mathbf{H}^{-}\mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\} = \left[\mathbf{H}^{\dagger} + \mathbf{P}_{\perp}\mathbf{U}\right]\mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}. \tag{10}$$

This reduces the precoder design problem to an optimization with respect to the elements of $\mathbf{p}$ and the specific choice of generalized inverse via $\mathbf{U}$. Roughly speaking, we will show that the optimization of $\mathbf{p}$ depends on the design criteria (fairness versus throughput), whereas the optimization of $\mathbf{U}$ is associated with the power constraints (total versus per-antenna). In fact, the discussion above suggests that the pseudo-inverse ($\mathbf{U} = \mathbf{0}$) is optimal with respect to the total power constraint which is associated with the Frobenius norm. We will show that when more complicated constraints are involved the optimal $\mathbf{U}$ is not necessarily zero.

IV. Total Power Constraint

The problem of ZF precoding design under a total power constraint has already received considerable attention [4], [6]–[10]. To our knowledge, in all of the previous works it was taken for granted that the precoder $\mathbf{T}$ must be based on the pseudo-inverse rather than any other generalized inverse. This simplified the design and reduced it to a power allocation problem. The next theorem proves that the pseudo-inverse is indeed optimal under a total power constraint:

Theorem 1: Let $f(\cdot)$ be an arbitrary function of $\mathbf{p}$. The optimal solution to

$$\begin{aligned} \max_{\mathbf{p}\geq\mathbf{0},\mathbf{T}} \quad & f(\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{H}\mathbf{T} = \mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}; \\ & \mathrm{Tr}\left\{\mathbf{T}\mathbf{T}^{H}\right\} \leq P \end{aligned} \tag{11}$$

is $\mathbf{T}^{\mathrm{opt}} = \mathbf{H}^{\dagger}\mathrm{diag}\left\{\sqrt{\mathbf{p}^{\mathrm{opt}}}\right\}$ where $\mathbf{p}^{\mathrm{opt}}$ is the solution to

$$\begin{aligned} \max_{\mathbf{p}\geq\mathbf{0}} \quad & f(\mathbf{p}) \\ \text{s.t.} \quad & \sum_k p_k \left[\left(\mathbf{H}^{\dagger}\right)^{H}\mathbf{H}^{\dagger}\right]_{k,k} \leq P. \end{aligned} \tag{12}$$

Proof: Due to (10), we can rewrite (11) as

$$\begin{aligned} \max_{\mathbf{p}\geq\mathbf{0},\mathbf{U}} \quad & f(\mathbf{p}) \\ \text{s.t.} \quad & \mathrm{Tr}\left\{\left[\mathbf{H}^{\dagger}+\mathbf{P}_{\perp}\mathbf{U}\right]\mathrm{diag}\{\mathbf{p}\}\left[\mathbf{H}^{\dagger}+\mathbf{P}_{\perp}\mathbf{U}\right]^{H}\right\} \leq P. \end{aligned} \tag{13}$$

Now,

$$\begin{aligned} &\mathrm{Tr}\left\{\left[\mathbf{H}^{\dagger} + \mathbf{P}_{\perp}\mathbf{U}\right]\mathrm{diag}\left\{\mathbf{p}\right\}\left[\mathbf{H}^{\dagger} + \mathbf{P}_{\perp}\mathbf{U}\right]^{H}\right\} \\ &\qquad\geq \mathrm{Tr}\left\{\mathbf{H}^{\dagger}\mathrm{diag}\left\{\mathbf{p}\right\}\left[\mathbf{H}^{\dagger}\right]^{H}\right\} \end{aligned} \tag{14}$$

since $\mathbf{P}_{\perp}\mathbf{H}^{\dagger} = \mathbf{0}$ and $\mathbf{P}_{\perp}\mathbf{U}\mathrm{diag}\left\{\mathbf{p}\right\}\mathbf{U}^{H}\mathbf{P}_{\perp} \succeq \mathbf{0}$. Therefore, the following problem:

$$\max_{\mathbf{p}\geq\mathbf{0}} \quad f(\mathbf{p}) \qquad \text{s.t.} \quad \mathrm{Tr}\left\{\mathbf{H}^{\dagger}\mathrm{diag}\left\{\mathbf{p}\right\}\left[\mathbf{H}^{\dagger}\right]^{H}\right\} \leq P \tag{15}$$

is a relaxation of (13) and generates an upper bound on its optimal value. However, this bound can be achieved by choosing $\mathbf{U} = 0$ and is therefore tight. Finally, choosing $\mathbf{U} = \mathbf{0}$ is equivalent to $\mathbf{T} = \mathbf{H}^{\dagger}\mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}$ and results in (12). ■

The importance of this result stems from the fact that (12) is a simple power allocation problem. In particular, assuming that $f(\mathbf{p})$ is concave in $\mathbf{p} \geq \mathbf{0}$, the problem is a concave maximization with one linear constraint. For example, in the throughput problem the problem boils down to [6], [8]

$$\max_{\mathbf{p}\geq 0} \quad \sum_k \log\left(1 + p_k\right) \quad \text{s.t.} \quad \sum_k p_k \left[\left(\mathbf{H}\mathbf{H}^{H}\right)^{-1}\right]_{k,k} \leq P \tag{16}$$

which can be solved using the well known water filling solution.

V. Per-Antenna Power Constraints

We now treat the more difficult case of ZF precoding design under individual per-antenna power constraints. Here, the pseudo-inverse is not necessarily the optimal generalized inverse. In fact, finding the optimal inverse is a nontrivial optimization problem which depends on the specific performance measure. Therefore, we begin by presenting general performance bounds and then address the two standard metrics, fairness, and throughput separately.

The optimal ZF precoder with per-antenna power constraints for maximizing an arbitrary objective function $f(\mathbf{p})$ is the solution to

$$f(\mathbf{p}^{\mathrm{opt}}) = \begin{cases} \max_{\mathbf{p}\geq\mathbf{0},\mathbf{T}} & f(\mathbf{p}) \\ \text{s.t.} & \mathbf{H}\mathbf{T} = \mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}; \\ & \left[\mathbf{T}\mathbf{T}^{H}\right]_{n,n} \leq \frac{P}{N} \ \forall\, n. \end{cases} \tag{17}$$

In general, (17) is a difficult nonconvex optimization problem. However, we can easily bound its optimal value

$$L \leq f(\mathbf{p}^{\rm opt}) \leq U \tag{18}$$

where

$$L = \begin{cases} \max_{\mathbf{p}\geq \mathbf{0}} & f(\mathbf{p}) \\ \text{s.t.} & \sum_k p_k \left|\mathbf{H}^{\dagger}_{n,k}\right|^2 \leq \frac{P}{N} \quad \forall\, n \end{cases} \tag{19}$$

$$U = \begin{cases} \max_{\mathbf{p}\geq \mathbf{0}} & f(\mathbf{p}) \\ \text{s.t.} & \sum_k p_k \left[\left(\mathbf{H}^{\dagger}\right)^H \mathbf{H}^{\dagger}\right]_{k,k} \leq P\,. \end{cases} \tag{20}$$

As proof, just note that the lower bound in (19) can be achieved by using the pseudo-inverse $\mathbf{T} = \mathbf{H}^{\dagger}\mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}$. Indeed, this $\mathbf{T}$ yields $[\mathbf{T}\mathbf{T}^H]_{n,n} = \sum_k p_k|\mathbf{H}^{\dagger}_{n,k}|^2$ as expressed in the constraints of (19). The upper bound is equal to the optimal value of (11) or (12). Clearly, if $\mathbf{T}$ is feasible for (17) then it will also be feasible for (11). Therefore, (11) is a relaxation of (17) and results in an upper bound.

Although simple, these bounds provide some insight into the problem without the need for solving (17) explicitly. Indeed, a sufficient condition for the optimality of the pseudo-inverse is $U = L$. Moreover, when the condition does not hold, we can bound the performance loss due to using the pseudo-inverse by examining the value of $U - L$ or U/L. Depending on the application, if this difference is sufficiently small, then an effective solution can be obtained without the need to solve (17). Otherwise, there may be an advantage in finding the optimal generalized inverse. This optimization is usually more complicated and depends on the specific performance measure. In the following sections, we treat two standard objectives: fairness and throughput.

A. *Fairness*

We begin with the fairness criterion which yields the following optimization problem:

$$\begin{aligned} \max_{\mathbf{p}\geq\mathbf{0},\mathbf{T}} \quad & \min_k p_k \\ \text{s.t.} \quad & \mathbf{H}\mathbf{T} = \mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}; \\ & \left[\mathbf{T}\mathbf{T}^H\right]_{n,n} \leq \frac{P}{N} \quad \forall\, n. \end{aligned} \tag{21}$$

We begin by examining our previous bounds, and provide a simple sufficient condition for the optimality of the pseudo-inverse.

Proposition 1: Consider the fairness optimization problem in (21). The loss in the objective value due to using the (possibly) suboptimal pseudoinverse based precoder is upper bounded by

$$\frac{U}{L} \leq \frac{N\max_n a_n}{\sum\limits_k c_k} = \frac{N\max_n a_n}{\sum\limits_n a_n} \tag{22}$$

where

$$\begin{aligned} a_n &= \left[\mathbf{H}^{\dagger}\left(\mathbf{H}^{\dagger}\right)^H\right]_{n,n}, \quad n = 1,\ldots,N \\ c_k &= \left[\left(\mathbf{H}^{\dagger}\right)^H\mathbf{H}^{\dagger}\right]_{k,k}, \quad k = 1,\ldots,K \end{aligned} \tag{23}$$

and we assume that $\max_n a_n > 0$ and $\sum_k c_k > 0$. In particular, if $a_n = a$ are equal for all n then $U = L$ and the solution to (21) is $\mathbf{T} = (P/Na)\mathbf{H}^{\dagger}$.

Proof: As proof, just note that $L \geq P/(N\max_n a_n)$ since $\mathbf{p} = P/(N\max_n a_n)\mathbf{1}$ is feasible for (19), and that the optimal solution to (20) is simply $\mathbf{p} = P/(\sum_k c_k)\mathbf{1}$. The second equality holds since

$$\sum_k c_k = \mathrm{Tr}\left\{\left(\mathbf{H}^{\dagger}\right)^H\mathbf{H}^{\dagger}\right\} = \mathrm{Tr}\left\{\mathbf{H}^{\dagger}\left(\mathbf{H}^{\dagger}\right)^H\right\} = \sum_n a_n. \tag{24}$$

■

The condition $a_n = a$ in Proposition 1 holds in many practical deterministic channels. For example, it applies whenever the right singular vectors of $\mathbf{H}$ are the Fourier vectors. More details on such matrices and geometrically uniform frames can be found in [31]. Moreover, the condition holds asymptotically in the number of users under different models in which $\mathbf{H}$ is a random matrix. Two typical examples that arise in wireless communication systems are when the elements of $\mathbf{H}$ are zero mean, equal variance and independent complex Gaussian random variables [32], and when $\mathbf{H}$ is modeled using the circular Wyner model [21], [22].

We now continue with the general solution to (21). As can be expected, the fairness criterion implies that

$$\mathbf{p} = p\mathbf{1} \tag{25}$$

for some $p \geq 0$ is optimal. As proof, assume that the optimal solution is $\overline{\mathbf{T}}$ and $\overline{\mathbf{p}} \geq \mathbf{0}$. If $\overline{p}_k = 0$ for some k then $\mathbf{T} = \mathbf{0}$ and $\mathbf{p} = \mathbf{0}$ are also optimal. Otherwise, define $p = \min_k \overline{p}_k$, $\mathbf{T} = \sqrt{p}\,\overline{\mathbf{T}}\left[\mathrm{diag}\left\{\sqrt{\overline{\mathbf{p}}}\right\}\right]^{-1}$ and $\mathbf{p} = p\mathbf{1}$. Then, $\mathbf{T}$ and $\mathbf{p}$ are also feasible (since $p/\overline{p}_k \leq 1$ for all k) and provide the same objective value as $\overline{\mathbf{T}}$.

Due to (10) and (25), we obtain

$$\mathbf{T} = \sqrt{p}\left[\mathbf{H}^{\dagger} + \mathbf{P}_{\perp}\mathbf{U}\right] \tag{26}$$

for some $\mathbf{U}$. This reduces the problem to

$$\begin{aligned} \max_{\mathbf{U},p\geq 0} \quad & p \\ \text{s.t.} \quad & p\left\|\left[\mathbf{H}^{\dagger}+\mathbf{P}_{\perp}\mathbf{U}\right]^{H}\mathbf{e}_n\right\|^2 \leq \frac{P}{N} \quad \forall\, n. \end{aligned} \tag{27}$$

Now, it is clear that

$$p=\frac{P}{N\max_n\left\|\left[\mathbf{H}^{\dagger}+\mathbf{P}_{\perp}\mathbf{U}\right]^{H}\mathbf{e}_n\right\|^2} \tag{28}$$

where $\mathbf{U}$ is the solution to

$$\min_{\mathbf{U},t} \quad t \qquad \text{s.t.} \quad \left\|\left(\mathbf{H}^{\dagger}+\mathbf{P}_{\perp}\mathbf{U}\right)^{H}\mathbf{e}_n\right\| \leq t \quad \forall\, n. \tag{29}$$

Problem (29) is a convex second order cone program (SOCP). It can be solved efficiently using standard optimization packages [33], [34].

B. Throughput

Next, we consider the throughput objective function

$$\begin{aligned} \max_{\mathbf{p}\geq\mathbf{0},\mathbf{T}} \quad & \sum_k \log\left(1+p_k\right) \\ \text{s.t.} \quad & \mathbf{HT}=\mathrm{diag}\left\{\sqrt{\mathbf{p}}\right\}; \\ & \left[\mathbf{TT}^H\right]_{n,n} \leq \frac{P}{N} \quad \forall\, n. \end{aligned} \tag{30}$$

This is a difficult nonconcave maximization problem due to the square root of $\mathbf{p}$. In this section, we will show how it can be solved using modern convex optimization tools. But before that, we can examine the optimality of the pseudo-inverse using our general bounds.

Proposition 2: Consider the fairness optimization problem in (30). The loss in the objective value due to using the (possibly) suboptimal pseudoinverse based precoder is upper bounded by

$$\begin{aligned} U-L &\leq N\frac{P}{N\max_n a_n+P}\left[\frac{\max_n a_n}{\min_k c_k}-\frac{K}{N}\right] \\ &\xrightarrow{P\to\infty} N\left[\frac{\max_n a_n}{\min_k c_k}-\frac{K}{N}\right] \end{aligned} \tag{31}$$

where a_n and c_k are defined in (23) and we assume that $\max_n a_n > 0$ and $\min_k c_k > 0$. This loss is clearly a power loss and does not effect the multiplexing gain, i.e., the number of degrees of freedom. In particular, at high SNR, the loss is bounded by a constant which does not depend on the SNR.

Proof: The proof is a straightforward consequence of the well known result that uniform power allocation tends to maximize the throughput in high SNR [35]. For completeness, the details are provided in Appendix I. ■

In the remainder of this section, we provide an exact solution to (30) which finds the optimal generalized inverse. For this purpose, it is convenient to rewrite the problem using the notation in (1), i.e., $\mathbf{h}_k=\mathbf{H}^H\mathbf{e}_k$ and $\mathbf{t}_k=\mathbf{Te}_k$ for $k=1,\ldots,K$. Thus, $p_k=\left|\mathbf{h}_k^H\mathbf{t}_k\right|^2$ and (30) is equivalent to

$$\begin{aligned} \max_{\mathbf{t}_k} \quad & \log\left|\mathbf{I}+\mathrm{diag}\left\{\left|\mathbf{h}_k^H\mathbf{t}_k\right|^2\right\}\right| \\ \text{s.t.} \quad & \left|\mathbf{h}_j^H\mathbf{t}_k\right|^2=0,\ \forall\, k\neq j; \\ & \sum_k\left[\mathbf{t}_k\mathbf{t}_k^H\right]_{n,n}\leq\frac{P}{N}\ \forall\, n. \end{aligned} \tag{32}$$

Next, we linearize the quadratic terms by defining $\mathbf{Q}_k=\mathbf{t}_k\mathbf{t}_k^H\succeq\mathbf{0}$ for $k=1,\ldots,K$, which results in

$$\begin{aligned} \max_{\mathbf{Q}_k} \quad & \log\left|\mathbf{I}+\mathrm{diag}\left\{\mathbf{h}_k^H\mathbf{Q}_k\mathbf{h}_k\right\}\right| \\ \text{s.t.} \quad & \mathbf{h}_j^H\mathbf{Q}_k\mathbf{h}_j=0,\ \forall\, k\neq j; \\ & \sum_k\left[\mathbf{Q}_k\right]_{n,n}\leq\frac{P}{N},\ \forall\, n; \\ & \mathbf{Q}_k\succeq\mathbf{0}\ \forall\, k; \\ & \mathrm{rank}\left\{\mathbf{Q}_k\right\}=1\ \forall\, k. \end{aligned} \tag{33}$$

The only nonconvex constraints in (33) are the rank-one restrictions. Therefore, we now relax the problem and omit these problematic constraints to obtain

$$\begin{aligned} \max_{\mathbf{Q}_k} \quad & \log\left|\mathbf{I}+\mathrm{diag}\left\{\mathbf{h}_k^H\mathbf{Q}_k\mathbf{h}_k\right\}\right| \\ \text{s.t.} \quad & \mathbf{h}_j^H\mathbf{Q}_k\mathbf{h}_j=0, \quad \forall\, k\neq j; \\ & \sum_k\left[\mathbf{Q}_k\right]_{n,n}\leq\frac{P}{N}, \quad \forall\, n; \\ & \mathbf{Q}_k\succeq\mathbf{0} \quad \forall\, k. \end{aligned} \tag{34}$$

Problem (34) is a standard determinant maximization (MAXDET) program subject to linear matrix inequalities [36]. It is a convex optimization problem and there are off-the-shelf numerical optimization packages which can solve it efficiently [34]. If the optimal $\mathbf{Q}_k$ are all of rank-one, then we can easily recover $\mathbf{t}_k$ from them and find the optimal solution to (30). Fortunately, the following theorem proves that the relaxation is always tight.

Theorem 2: Problem (34) always has a solution with rank-one matrices. This solution can be found as follows: Let $\mathbf{Q}_k^{\mathrm{opt}}$ for $k=1,\ldots,K$ be a (possibly high rank) optimal solution to (34). For each k define $\mathbf{t}_k$ as the optimal solution to

$$\begin{aligned} \max_{\mathbf{t}} \quad & \mathrm{Re}\left\{\mathbf{h}_k^H\mathbf{t}\right\} \\ \text{s.t.} \quad & \mathbf{h}_j^H\mathbf{t}=0,\ \forall\, k\neq j; \\ & \left|\left[\mathbf{t}\right]_n\right|^2\leq\left[\mathbf{Q}_k^{\mathrm{opt}}\right]_{n,n}\ \forall\, n. \end{aligned} \tag{35}$$

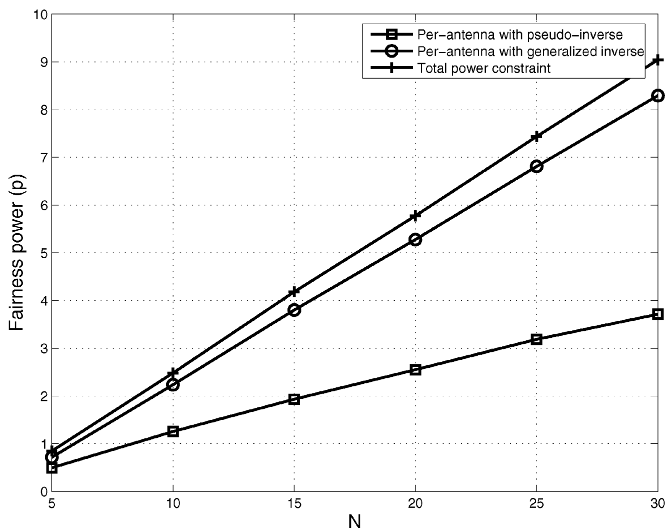

Fig. 2. Maximal fairness ZF precoding as a function of N.

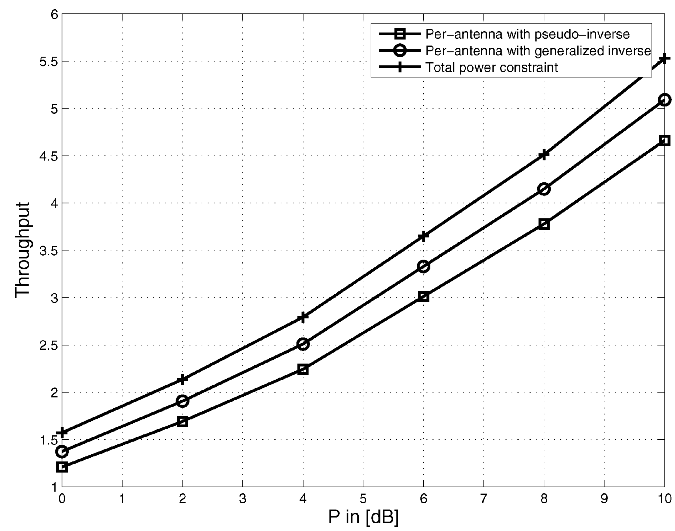

Fig. 3. Maximal throughput ZF precoding as a function of P.

Then, $\overline{\mathbf{Q}}_k^{\rm opt} = \mathbf{t}_k\mathbf{t}_k^H$ for $k=1,\ldots,K$ is a rank-one solution to (34).

Proof: See Appendix II. ∎

In practice, our experience shows that the MAXDET software [34] usually provides a rank-one solution automatically. If it does not, then the theorem provides a constructive method for finding a rank-one solution by solving K simple convex programs of the form (35).

VI. Numerical Results

We now demonstrate our results using two numerical examples. In the first example, we consider the fairness ZF precoding design under individual per-antenna power constraints. We simulate a system with $K=3$ users and $P=1$ (in the fairness case, the value of P is not important as it just scales the resulting power). The elements of the matrix $\mathbf{H}$ are randomly generated as independent, zero mean and unit variance complex Gaussian random variables. We estimate the average received power p in (26). For comparison, we also estimate this mean power when we assume $\mathbf{U}=\mathbf{0}$, i.e., restrict the precoder to be a standard pseudo-inverse, and when we replace the per-antenna power constraints with a total power constraint. The results are presented in Fig. 2 as a function of the number of transmit antennas N. As expected, the stricter per-antenna constraints result in a lower received power. However, the graph shows that part of this loss can be recovered by optimizing $\mathbf{U}$ and finding the appropriate generalized inverse.

In the second example, we consider the maximization of the throughput under the same setting as before except that now $N=4$, and we simulate different Ps. The estimated sum-rates are provided in Fig. 3. Again, it is easy to see the degradation in performance due to the individual per-antenna power constraints, as well as the advantage of optimizing the generalized inverse.

VII. Conclusion

In this paper, we consider ZF precoding design in MISO broadcast channels. We discussed the intimate relation between ZF precoding and the theory of generalized inverses. Our results show that designing the precoders based on the standard pseudo-inverse is optimal under the assumption of a total power constraint. However, when more complex power constraints are involved, e.g., individual total per-antenna power constraints, the pseudo-inverse is no longer sufficient and other generalized inverses may provide better performance. In general, finding the optimal inverse is a difficult optimization problem which is highly dependent on the specific design criterion. We consider two classical criteria, fairness and throughput and demonstrate how to transform these problems into standard convex optimization programs.

Using the methods that we developed it is straightforward to generalize the setting to a variety of applications. More practical criteria may be addressed using the semidefinite relaxation approach as long as these are concave in the received powers, e.g., weighted sum-rate. In addition, other power constraints may be implemented, e.g., the expected value of the squared norm of subblocks of $\mathbf{x}$. Such constraints may be important in modern systems where multiple base stations, each with multiple antennas, cooperatively transmit data to the same users.

Precoding with generalized power constraints is an important problem in modern communication systems and there are still many open questions. More advanced linear precoding schemes should be addressed. For example, it is well known that in low SNR conditions, and under channel uncertainty, regularizing the pseudo-inverse can considerably improve the performance. It is interesting to examine this property in the context of generalized inverses. Future work should also address the implications of our results on nonlinear schemes such as ZF DPC precoding.

Another extension of our work is to consider the well known duality between receive and transmit processing. It has already been shown in [19] that precoding with per-antenna power constraints is the dual of decoding under noise uncertainty conditions. ZF decoding using the pseudo-inverse (the decorrelator) is probably the most common decoding algorithm. Our results suggest that other generalized inverses may outperform it under uncertainty conditions.

APPENDIX I
PROOF OF PROPOSITION 2

We want to bound $U-L$ in (19)–(20) from above when $f(\mathbf{p})=\sum_k \log(1+p_k)$. First, note that

$$L \geq \sum_k \log\left(1+\frac{P}{N\max_n a_n}\right) \tag{36}$$

where $a_n>0$ are defined in (23), since $\mathbf{p}=(P/N\max_n a_n)\mathbf{1}$ is a feasible but (possibly) suboptimal solution to (19). On the other hand, we will upper bound U using its Lagrange dual function as proposed in [35]

$$\begin{aligned} U &\leq \max_{\mathbf{p}}\left\{\sum_k \log(1+p_k)+\sum_k \mu_k p_k+\lambda\left[P-\sum_k p_k c_k\right]\right\} \\ &= \max_{\mathbf{p}}\left\{\sum_k \log(1+p_k)+\sum_k p_k\left[\mu_k-\lambda c_k\right]+\lambda P\right\} \end{aligned} \tag{37}$$

where c_k are defined in (23) and the inequality holds for any nonnegative $\boldsymbol{\mu}=[\mu_1,\ldots,\mu_K]^T\geq\mathbf{0}$ and $\lambda\geq 0$. The bound is finite only if $\mu_k-\lambda c_k<0$ for all k. Under this condition, the maximum with respect to $\mathbf{p}$ is achieved when

$$\frac{1}{1+p_k}+\mu_k-\lambda c_k=0, \quad k=1,\ldots,K \tag{38}$$

and yields

$$U\leq\sum_k\left[\log\left(\frac{1}{\lambda c_k-\mu_k}\right)-1+\lambda c_k-\mu_k\right]+\lambda P. \tag{39}$$

This last bound holds for any $\boldsymbol{\mu}\geq\mathbf{0}$ and $\lambda\geq 0$ which satisfy $\mu_k-\lambda c_k<0$ for all k. Now, let

$$\begin{aligned} \mu_k &= \lambda c_k-\frac{1}{1+\frac{P}{N\max_n a_n}}; \\ \lambda &= \frac{1}{\left(1+\frac{P}{N\max_n a_n}\right)\min_k c_k} \end{aligned} \tag{40}$$

then all the conditions are satisfied and

$$\begin{aligned} U \leq \sum_k \log\left(1+\frac{P}{N\max_n a_n}\right) \\ +N\frac{\frac{P}{N\max_n a_n}}{1+\frac{P}{N\max_n a_n}}\left[\frac{\max_n a_n}{\min_k c_k}-\frac{K}{N}\right]. \end{aligned} \tag{41}$$

Therefore

$$\begin{aligned} U-L &\leq N\frac{\frac{P}{N\max_n a_n}}{1+\frac{P}{N\max_n a_n}}\left[\frac{\max_n a_n}{\min_k c_k}-\frac{K}{N}\right] \\ &\xrightarrow{P\to\infty} N\left[\frac{\max_n a_n}{\min_k c_k}-\frac{K}{N}\right]. \end{aligned} \tag{42}$$

APPENDIX II
PROOF OF THEOREM 2

First, we rewrite (34) using additional slack variables:

$$\begin{aligned} \max_{\boldsymbol{\pi}_k\geq\mathbf{0}} \quad & \sum_k \log\left(1+\gamma_k\left(\boldsymbol{\pi}_k\right)\right) \\ \text{s.t.} \quad & \sum_k \left[\boldsymbol{\pi}_k\right]_n \leq \frac{P}{N} \quad \forall\, n \end{aligned} \tag{43}$$

where

$$\gamma_k\left(\boldsymbol{\pi}_k\right)=\begin{cases}\max_{\mathbf{Q}\succeq\mathbf{0}} & \mathbf{h}_k^H\mathbf{Q}\mathbf{h}_k \\ \text{s.t.} & \mathbf{h}_j^H\mathbf{Q}\mathbf{h}_j=0, \ \forall k\neq j; \\ & [\mathbf{Q}]_{n,n}\leq[\boldsymbol{\pi}_k]_n \quad \forall\, n\end{cases} \tag{44}$$

where $k=1,\ldots,K$. Using this new formulation, all we need to show is that (44) always has an optimal solution of rank-one. In fact, we will prove a more general result.

Lemma 1: Consider the following optimization problem:

$$\mathcal{S}=\begin{cases}\max_{\mathbf{Q}\succeq\mathbf{0}} & \mathbf{c}^H\mathbf{Q}\mathbf{c} \\ \text{s.t.} & \operatorname{Tr}\{\mathbf{Q}\mathbf{A}_i\}\leq b_i, \quad i=1,\ldots,I\end{cases} \tag{45}$$

where $b_i\geq 0$ and $\mathbf{A}_i\succeq\mathbf{0}$. If $\mathcal{S}$ is bounded, then it always has a rank-one solution.

Proof: See Appendix III. ■

Problem (44) is a special case of Lemma 1. Due to the constraints $\mathbf{Q}\succeq\mathbf{0}$ and $[\mathbf{Q}]_{n,n}\leq[\boldsymbol{\pi}_k]_n$ its optimal value is bounded, and it must have an optimal solution of rank-one. The fact that (44) has an optimal rank-one solution also provides a simple way for finding it. Let $\mathbf{Q}_k^{\rm opt}(\boldsymbol{\pi}_k)$ be the optimal solution to (44) for some k. Then, it is clearly also the solution to

$$\gamma_k\left(\boldsymbol{\pi}_k\right)=\begin{cases}\max_{\mathbf{Q}\succeq\mathbf{0}} & \mathbf{h}_k^H\mathbf{Q}\mathbf{h}_k \\ \text{s.t.} & \mathbf{h}_j^H\mathbf{Q}\mathbf{h}_j=0, \quad \forall k\neq j\,; \\ & [\mathbf{Q}]_{n,n}\leq\left[\mathbf{Q}_k^{\rm opt}\left(\boldsymbol{\pi}_k\right)\right]_{n,n} \quad \forall\, n.\end{cases} \tag{46}$$

Due to Lemma 1, we can restrict the attention to rank-one matrices $\mathbf{Q}=\mathbf{t}\mathbf{t}^H$ and solve

$$\gamma_k\left(\boldsymbol{\pi}_k\right)=\begin{cases}\max_{\mathbf{t}} & \left|\mathbf{t}^H\mathbf{h}_k\right|^2 \\ \text{s.t.} & \mathbf{t}^H\mathbf{h}_j=0, \ \forall k\neq j; \\ & \left|\mathbf{t}^H\mathbf{e}_n\right|^2\leq\left[\mathbf{Q}_k^{\rm opt}(\boldsymbol{\pi}_k)\right]_{n,n} \ \forall n\,.\end{cases} \tag{47}$$

This last problem is nonconvex due to the quadratic objective. However, it can be solved by assuming that $\mathbf{t}^H\mathbf{h}_k$ is real and nonnegative, taking its square root value and noting that the optimal solution does not change (up to a phase rotation)

$$\sqrt{\gamma_k(\boldsymbol{\pi}_k)} = \begin{cases} \max_{\mathbf{t}} & \mathrm{Re}\left\{\mathbf{t}^H\mathbf{h}_k\right\} \\ \text{s.t.} & \mathbf{t}^H\mathbf{h}_j = 0, \quad \forall\, k \neq j; \\ & \left|\mathbf{t}^H\mathbf{e}_n\right|^2 \leq \left[\mathbf{Q}_k^{\mathrm{opt}}(\boldsymbol{\pi}_k)\right]_{n,n} \quad \forall\, n\,. \end{cases} \tag{48}$$

As proof, assume that $\overline{\mathbf{t}}$ is optimal for (48) then it is clearly feasible for (47) and results in the required objective value. On the other hand, assume that $\overline{\mathbf{t}}$ is optimal for (47). Let ϕ be the angle of $\overline{\mathbf{t}}^H\mathbf{h}_k$, i.e., $\overline{\mathbf{t}}^H\mathbf{h}_k = e^{j\phi}|\overline{\mathbf{t}}^H\mathbf{h}_k|$. Then, $\widehat{\mathbf{t}} = e^{-j\phi}\overline{\mathbf{t}}$ is feasible for (48) and results in the required objective value.

APPENDIX III
PROOF OF LEMMA 1

We begin by eliminating all the constraints for which $b_i = 0$. Assume that $b_j = 0$ for all j in $J = \{j_1, \ldots, j_J\}$, and positive for all other indices. Define $\overline{\mathbf{A}} = [\mathbf{A}_{j_1}, \ldots, \mathbf{A}_{j_J}]$, and let $\mathbf{P} = \mathbf{I} - \overline{\mathbf{A}}\overline{\mathbf{A}}^\dagger$ be the orthogonal projection onto the null space of $\overline{\mathbf{A}}$. Now, $\mathrm{Tr}\{\mathbf{Q}\mathbf{A}_j\} = 0$ for all $j \in J$ if and only if $\mathbf{Q} = \mathbf{P}\mathbf{Q}\mathbf{P}$. Thus, $\mathcal{S}$ in (45) is equivalent to

$$\begin{aligned} \max_{\mathbf{Q}\succeq\mathbf{0}} \quad & \mathbf{c}^H\mathbf{P}\mathbf{Q}\mathbf{P}\mathbf{c} \\ \text{s.t.} \quad & \mathrm{Tr}\{\mathbf{P}\mathbf{Q}\mathbf{P}\mathbf{A}_i\} \leq b_i, \quad i \notin J, \\ & \mathbf{Q} = \mathbf{P}\mathbf{Q}\mathbf{P}. \end{aligned} \tag{49}$$

Next, we omit the $\mathbf{Q} = \mathbf{P}\mathbf{Q}\mathbf{P}$ constraint and obtain

$$\begin{aligned} \max_{\mathbf{Q}\succeq\mathbf{0}} \quad & \overline{\mathbf{c}}^H\mathbf{Q}\overline{\mathbf{c}} \\ \text{s.t.} \quad & \mathrm{Tr}\left\{\mathbf{Q}\overline{\mathbf{A}_i}\right\} \leq b_i, \quad i \notin J \end{aligned} \tag{50}$$

where $\overline{\mathbf{c}} = \mathbf{P}\mathbf{c}$, $\overline{\mathbf{A}_i} = \mathbf{P}\mathbf{A}_i\mathbf{P}$ and $b_i > 0$ for $i \notin J$ are all strictly positive. If $\mathbf{Q}^{\mathrm{opt}}$ is a rank-one optimal solution to (50) then $\mathbf{P}\mathbf{Q}^{\mathrm{opt}}\mathbf{P}$ is a rank-one optimal solution to (49). Therefore, we can prove the lemma for (50) instead of (45). For simplicity, we continue with the notation in (45) but assume that $b_i > 0$ for all i.

Consider the following problem:

$$\mathcal{Q} = \begin{cases} \max_{\mathbf{q}} & \left|\mathbf{q}^H\mathbf{c}\right|^2 \\ \text{s.t.} & \mathbf{q}^H\mathbf{A}_i\mathbf{q} \leq b_i\,. \end{cases} \tag{51}$$

Program $\mathcal{S}$ in (45) is the SDP relaxation of $\mathcal{Q}$. That is $\mathrm{val}\{\mathcal{S}\} \geq \mathrm{val}\{\mathcal{Q}\}$, and if $\mathbf{q}$ is optimal for $\mathcal{Q}$ then $\mathbf{Q} = \mathbf{q}\mathbf{q}^H$ is feasible for $\mathcal{S}$. Thus, all we need to prove is that $\mathrm{val}\{\mathcal{S}\} \leq \mathrm{val}\{\mathcal{Q}\}$. We will do this by considering their corresponding dual programs.

We begin with $\mathcal{S}$ which is a convex optimization problem. Its Lagrange dual is

$$d\mathcal{S} = \begin{cases} \min_{\lambda\succeq\mathbf{0}} & \sum_i \lambda_i b_i \\ \text{s.t.} & \sum_i \lambda_i\mathbf{A}_i - \mathbf{c}\mathbf{c}^H \succeq \mathbf{0}\,. \end{cases} \tag{52}$$

The primal problem is strictly feasible since b_i are all positive. Therefore, Slater's condition for strong duality holds and $\mathrm{val}\{\mathcal{S}\} = \mathrm{val}\{d\mathcal{S}\}$.

We now move on to $\mathcal{Q}$ in (51). Unfortunately, this is a nonconvex problem due to the quadratic objective. However, we can find its optimal value by defining

$$\mathcal{L} = \begin{cases} \max_{\mathbf{q}} & \mathrm{Re}\left\{\mathbf{q}^H\mathbf{c}\right\} \\ \text{s.t.} & \mathbf{q}^H\mathbf{A}_i\mathbf{q} \leq b_i \end{cases} \tag{53}$$

and noting that $(\mathrm{val}\{\mathcal{L}\})^2 = \mathrm{val}\{\mathcal{Q}\}$. As proof, assume that $\overline{\mathbf{q}}$ is optimal for $\mathcal{L}$ then it is clearly feasible for $\mathcal{Q}$ and results in the required objective value. On the other hand, assume that $\overline{\mathbf{q}}$ is optimal for $\mathcal{Q}$. Let ϕ be the angle of $\overline{\mathbf{q}}^H\mathbf{c}$, i.e., $\overline{\mathbf{q}}^H\mathbf{c} = e^{j\phi}|\overline{\mathbf{q}}^H\mathbf{c}|$. Then, $\widehat{\mathbf{q}} = e^{-j\phi}\overline{\mathbf{q}}$ is feasible for $\mathcal{L}$ and results in the required objective value.

The main advantage of this linearization is that $\mathcal{L}$ is a convex optimization problem which can be solved using its Lagrange dual program

$$d\mathcal{L} = \min_{\lambda\geq 0}\max_{\mathbf{q}} \mathrm{Re}\left\{\mathbf{q}^H\mathbf{c}\right\} - \sum_i \lambda_i\mathbf{q}^H\mathbf{A}_i\mathbf{q} + \sum_i \lambda_i b_i. \tag{54}$$

Adding an auxiliary variable $t \geq 0$ yields

$$d\mathcal{L} = \begin{cases} \min_{\lambda, t\geq 0} & t + \sum_i \lambda_i b_i \\ \text{s.t.} & \sum_i \lambda_i\mathbf{q}^H\mathbf{A}_i\mathbf{q} - \mathrm{Re}\left\{\mathbf{q}^H\mathbf{c}\right\} + t \geq 0\;\forall\,\mathbf{q}. \end{cases} \tag{55}$$

We now simplify the constraint using the following lemma.

Lemma 2: [37, p. 135]:[1] Let $\mathbf{S}$ be an Hermitian matrix. The condition $\mathbf{x}^H\mathbf{S}\mathbf{x} + 2\mathrm{Re}\{\mathbf{s}^H\mathbf{x}\} + s \geq 0$ holds for all $\mathbf{x}$ if and only if

$$\begin{bmatrix} \mathbf{S} & \mathbf{s} \\ \mathbf{s}^H & s \end{bmatrix} \succeq \mathbf{0}.$$

Applying the lemma to our problem yields

$$d\mathcal{L} = \begin{cases} \min_{\lambda, t\geq 0} & t + \sum \lambda_i b_i \\ \text{s.t.} & \begin{bmatrix} \sum_i \lambda_i\mathbf{A}_i & -\frac{1}{2}\mathbf{c} \\ -\frac{1}{2}\mathbf{c}^H & t \end{bmatrix} \succeq \mathbf{0}\,. \end{cases} \tag{56}$$

If $t = 0$ then $\mathbf{c} = \mathbf{0}$ and the proof is completed since $\mathrm{val}\{\mathcal{S}\} = \mathrm{val}\{\mathcal{Q}\} = 0$. Otherwise, $t > 0$ and we can utilize Schur's complement lemma below.

Lemma 3 (Schur's Complement) [30]: Let $s > 0$. The following two conditions are equivalent:

$$\begin{bmatrix} \mathbf{S} & \mathbf{s} \\ \mathbf{s}^H & s \end{bmatrix} \succeq \mathbf{0} \quad \Leftrightarrow \quad \mathbf{S} - \frac{1}{s}\mathbf{s}\mathbf{s}^H \succeq \mathbf{0}. \tag{57}$$

to obtain

$$d\mathcal{L} = \begin{cases} \min_{\lambda\geq 0, t>0} & \sum_i \lambda_i b_i + t \\ \text{s.t.} & \sum_i \lambda_i\mathbf{A}_i - \frac{1}{4t}\mathbf{c}\mathbf{c}^H \succeq \mathbf{0}\,. \end{cases} \tag{58}$$

As before, Slater's condition holds due to the strict feasibility. Thus, strong duality assures that $\mathrm{val}\{\mathcal{L}\} = \mathrm{val}\{d\mathcal{L}\}$ and if we square the objective again and use the monotonicity of x^2

[1]This reference deals with the real case version of the lemma. The extension to the complex case is straightforward.

in $x \geq 0$, we obtain the following dual of $\mathcal{Q}$ (this is not the Lagrange dual but just the squared value of $d\mathcal{L}$)

$$d\mathcal{Q}=(d\mathcal{L})^2=\begin{cases} \min_{\lambda\geq 0, t>0} & \left(\sum_i \lambda_i b_i + t\right)^2 \\ \text{s.t.} & \sum_i \lambda_i \mathbf{A}_i - \frac{1}{4t}\mathbf{c}\mathbf{c}^H \succeq \mathbf{0} \end{cases} \quad (59)$$

which satisfies $\mathrm{val}\{\mathcal{Q}\} = \mathrm{val}\{d\mathcal{Q}\}$. Next, we exchange variables and optimize over $\overline{\lambda}_i = 4t\lambda_i \geq 0$ instead of λ_i

$$d\mathcal{Q}=\begin{cases} \min_{\overline{\lambda}\geq 0, t>0} & \left(\frac{\sum_i \overline{\lambda}_i b_i}{4t} + t\right)^2 \\ \text{s.t.} & \sum_i \overline{\lambda}_i \mathbf{A}_i - \mathbf{c}\mathbf{c}^H \succeq \mathbf{0}. \end{cases} \quad (60)$$

Now examining (52) and (60) we see that their feasible sets are identical, and in order to prove that $\mathrm{val}\{d\mathcal{Q}\} \geq \mathrm{val}\{d\mathcal{S}\}$ all we need to show is that

$$\frac{\mu}{4t} + t \geq \sqrt{\mu} \quad \text{for all} \quad t > 0 \quad (61)$$

where $\mu = \sum_i \lambda_i b_i \geq 0$. But this is easily proved by noting that the left-hand side of (61) is convex in $t > 0$ and attains its minimum when

$$\frac{\partial}{\partial t}\left[\frac{\mu}{4t} + t\right] = -\frac{\mu}{4t^2} + 1 = 0 \quad (62)$$

and

$$t_{\min} = \frac{1}{2}\sqrt{\mu} \quad (63)$$

which yields

$$\frac{\mu}{4t_{\min}} + t_{\min} = \sqrt{\mu} \quad (64)$$

as required.

Acknowledgment

The authors would like to thank the Associate Editor, Prof. T. N. Davisdon, for his constructive suggestions which significantly improved the quality of this paper.

References

[1] G. Caire and S. Shamai (Shitz), "On the achievable throughput of multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[2] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[3] M. Stojnic, H. Vikalo, and B. Hassibi, "Maximizing the sum-rate of multi-antenna broadcast channels using linear preprocessing," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2338–2342, Sep. 2006.

[4] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[5] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization," *IEEE Trans. Signal Process.*, vol. 55, pp. 5436–5446, Nov. 2007.

[6] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: Performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 3857–3868, Oct. 2005.

[7] B. R. Vojcic and W. M. Jang, "Transmitter precoding in synchronous multiuser communication," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1346–1355, Oct. 1998.

[8] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun. (Special Issue on 4G Wireless Systems)*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[9] T. Yoo and A. J. Goldsmith, "Optimality of zero-forcing beamforming with multiuser diversity," in *Proc. IEEE Int. Conf. Communications (ICC)*, 2005, vol. 1, pp. 542–546.

[10] F. Boccardi, F. Tosato, and G. Caire, "Precoding schemes for the MIMO-GBC," in *Proc. IEEE Int. Zurich Seminar (IZS-2006)*, Feb. 2006.

[11] D. Bartolome and A. I. Perez-Neira, "Spatial scheduling in multiuser wireless systems: From power allocation to admission control," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2082–2091, Aug. 2006.

[12] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[13] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, Sep. 2006.

[14] L. U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.

[15] K. K. Wong, "Maximizing the sum-rate and minimizing the sum-power of a broadcast 2-user 2-input multiple-output antenna system using a generalized zeroforcing approach," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3406–3412, Dec. 2006.

[16] Z. Pan, K. K. Wong, and T. S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969–1973, Nov. 2004.

[17] P. S. Udupa and J. S. Lehnert, "Optimizing zero-forcing precoders for MIMO broadcast systems," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1516–1524, Aug. 2007.

[18] X. Zheng, Y. Xie, J. Li, and P. Stoica, "MIMO transmit beamforming under uniform elemental power constraint," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5395–5406, Nov. 2007.

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

[20] F. Boccardi and H. Huang, "Optimum power allocation for the MIMO-BC zero-forcing precoder with per-antenna power constraints," in *Proc. Conf. Information Sciences Systems (CISS)*, Mar. 2006.

[21] O. Somekh, O. Simeone, Y. Bar-Ness, and A. Haimovich, "Distributed multi-cell zero-forcing beamforming in cellular downlink channels," in *Proc. IEEE GLOBECOM 2006*, 2006.

[22] O. Somekh, O. Simeone, Y. Bar-Ness, A. Haimovich, and S. Shamai (Shitz), "Distributed multi-cell zero-forcing beamforming in cellular downlink channels," *IEEE Trans. Inf. Theory*, 2008, to be published.

[23] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications*. New York: Wiley, 1971.

[24] W. K. Ma, T. N. Davidson, K. M. Wong, Z. Q. Luo, and P. C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 912–922, Apr. 2002.

[25] N. D. Sidiropoulos, T. N. Davidson, and Z. Q. Luo, "Transmit beamforming for physical layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[26] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," presented at the 37th Annu. Allerton, Control, Computing, Monticello, Sep. 1999.

[27] M. Bengtsson and B. Ottersten, "Handbook of antenna in wireless communications," in *Optimal and Suboptimal Transmit Beamforming*. Boca Raton, FL: CRC Press, 2001, vol. 18.

[28] D. Hammarwall, M. Bengtsson, and B. Ottersten, "On downlink beamforming with indefinite shaping constraints," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3566–3580, Sep. 2006.

[29] D. Samuelsson, M. Bengtsson, and B. Ottersten, "An efficient algorithm for solving the downlink beamforming problem with indefinite constraints," presented at the IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), Philadelphia, PA, Mar. .

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[31] Y. C. Eldar and H. Bolcskei, "Geometrically uniform frames," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 993–1006, Apr. 2003.

[32] Y. C. Eldar and A. M. Chan, "On the asymptotic performance of the decorrelator," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 2309–2313, Sep. 2003.

[33] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimizations over symmetric cones," *Optimizat. Methods Softw.*, vol. 11–12, pp. 625–653, 1999.

[34] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *IEEE Int. Symp. Computer-Aided Control Systems Design (CACSD)*, Taipei, Taiwan, R.O.C., Sep. 2004, pp. 284–289.

[35] W. Yu and J. M. Cioffi, "Constant power water-filling: Performance bound and low-complexity implementation," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 23–28, Jan. 2006.

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

[37] A. Nemirovski, Lectures on Modern Convex Optimization [Online]. Available: http://www2.isye.gatech.edu/ nemirovs

Ami Wiesel (S'02) received the B.Sc. and M.Sc. degrees, both in electrical engineering, from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 2000 and 2002, respectively, and the Ph.D. degree in electrical engineering rom the Technion—Israel Institute of Technology, Haifa, Israel, in 2007.

He is currently a Postdoc Fellow with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor.

Dr. Wiesel received the Young Author Best Paper Award for an IEEE TRANSACTIONS IN SIGNAL PROCESSING paper in 2008 and a Student Paper Award for an IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC) 2005 paper. He was awarded the Weinstein Study Prize in 2002, the Intel Award in 2005, the Viterbi Fellowship in 2005 and 2007 and the Marie Curie Fellowship in 2007.

Yonina C. Eldar (S'98–M'02–SM'07) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering both from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2001.

From January 2002 to July 2002, she was a Postdoctoral Fellow at the Digital Signal Processing Group at MIT. She is currently an Associate Professor in the Department of Electrical Engineering at the Technion—Israel Institute of Technology, Haifa, Israel. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT. Her research interests are in the general areas of signal processing, statistical signal processing, and computational biology.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in electrical engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002 to 2005, she was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005 the Andre and Bella Meyer Lectureship, and in 2007 the Henry Taub Prize for Excellence in Research, and in 2008 the Hershel Rich Innovation Award, the Award for Women with Distinguished Contributions, and the Muriel & David Jacknow Award for Excellence in Teaching. She is a member of the IEEE Signal Processing Theory and Methods technical committee, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the *EURASIP Journal of Signal Processing*, and the *SIAM Journal on Matrix Analysis and Applications*, and she is on the Editorial Board of *Foundations and Trends in Signal Processing*.

Shlomo Shamai (Shitz) (S'80–M'82–SM'89–F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 1975, 1981 and 1986, respectively.

From 1975 to 1985, he was a Senior Research Engineer with the Communications Research Laboratories. Since 1986, he has been with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompass a wide spectrum of topics in information theory and statistical communications. He is especially interested in theoretical limits in communication with practical constraints, multiuser information theory and spread spectrum systems, multiple-input-multiple-output communications systems, information theoretic models for wireless networks and systems, information theoretic aspects of magnetic recording, channel coding, combined modulation and coding, turbo codes and LDPC, in channel, source, and combined source-channel applications, iterative detection and decoding algorithms, coherent and noncoherent detection, and information theoretic aspects of digital communication in optical channels.

Dr. Shamai (Shitz) is a member of the Union Radio Scientifique Internationale (URSI). He is the recipient of the 1999 van der Pol Gold Medal of URSI, and a corecipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003 and the 2004 Joint IT/COM Societies Paper Award, and the 2007 Information Theory Society Paper Award. He is also the recipient of the 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and also serves on the Board of Governors of the Information Theory Society.