On the Optimality of Beamformer Design for Zero-forcing DPC with QR Decomposition

Le-Nam Tran*, Markku Juntti*, Mats Bengtsson†, and Björn Ottersten†‡
*Centre for Wireless Communications and Dept. Commun. Eng., University of Oulu, Oulu, Finland
Email: {le.nam.tran;markku.juntti}@ee.oulu.fi
†ACCESS Linnaeus Center, Royal Institute of Technology (KTH), Stockholm, Sweden
Email: {mats.bengtsson,bjorn.ottersten}@ee.kth.se
‡Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg

***Abstract*—We consider the beamformer design for zero-forcing dirty paper coding (ZF-DPC), a suboptimal transmission technique for MISO broadcast channels (MISO BCs). Beamformers for ZF-DPC are designed to maximize a performance measure, subject to some power constraints and zero-interference constraints. For the sum rate maximization problem under a total power constraint, the existing beamformer designs in the literature are based on the QR decomposition (QRD), which is used to satisfy the ZF constraints. However, the optimality of the QRD-based design is still unknown. First, we prove that the QRD-based design is indeed optimal for ZF-DPC for any performance measure under a sum power constraint. For the per-antenna power constraints, the QRD-based designs become suboptimal, and we propose an optimal design, using a convex optimization framework. Low-complexity suboptimal designs are also presented.**

I. INTRODUCTION

Multiple-input multiple-output (MIMO) transmission techniques exploit spatial dimensions provided by multiple antennas at both ends of a wireless link to increase the channel capacity without the need of additional bandwidth or power [1], compared to single-antenna systems. In this paper, we consider the downlink or broadcast channel (BC) of a single cell, where a multi-antenna base station (BS) wants to send data to multiple receivers simultaneously. Recent information theoretic studies have proved that dirty paper coding (DPC) is the capacity-achieving transmission technique for MIMO BCs [2]. Although this multiuser coding strategy is optimal, finding the resulting optimal transmit covariances faces computational complexity. Thus, there has been a large interest in developing suboptimal solutions to DPC.

Zero-forcing dirty paper coding (ZF-DPC), introduced in [3] for downlink channels with single-antenna receivers, i.e., for multiple-input single-output (MISO) BCs, is a suboptimal alternative to DPC, combining the ZF technique with DPC. Specifically, the data of the kth user in ZF-DPC is multiplied with a beamformer $\boldsymbol{w}_k$, which is designed such that $\boldsymbol{h}_j^H\boldsymbol{w}_k = 0$ for all $j < k$, where $\boldsymbol{h}_j$ is the channel vector of the jth user. These ZF constraints, combined with the use of DPC to cancel the non-causal interference term, decompose a BC into a group of parallel interference-free channels. The goal of the beamformer design is to find $\boldsymbol{w}_k$'s that satisfy the ZF constraints, and optimize a performance measure, subject to some power constraints such as sum power constraint (SPC) or per-antenna power constraints (PAPCs).

For the sum power constraint, a beamformer design was proposed in [3] based on the QR decomposition (QRD). This design is motivated by the fact that the ZF constraints force the product $\boldsymbol{HW}$ to be a lower triangular matrix, where $\boldsymbol{H} = \begin{bmatrix}\boldsymbol{h}_1 & \boldsymbol{h}_2 & \cdots & \boldsymbol{h}_K\end{bmatrix}^H$, $\boldsymbol{W} = [\boldsymbol{w}_1\boldsymbol{w}_2\cdots\boldsymbol{w}_K]$, and K is the number of users. Accordingly, one possibility is to obtain $\boldsymbol{W}$ by applying a QRD to $\boldsymbol{H}$ as proposed in [3]. It is worth noting that this QRD-based design is just one of many feasible designs that meet the ZF constraints. Thus, a natural question to ask is whether the QRD-based method is optimal for ZF-DPC. In fact, the QRD-based design is also used in other related works [4], [5] without investigating its optimality. In this paper, we show that the QRD-based design is indeed optimal for ZF-DPC under a SPC. The work in this paper is motivated by [6], where the pseudo-inverse was proved to be optimal for zero-forcing beamforming under the SPC.

In practice, PAPCs are usually more realistic than the SPC, since each antenna is typically equipped with its own power amplifier. The optimal beamformer design for ZF-DPC with PAPCs has not been extensively addressed. Generally, the beamformer design with the PAPCs is more difficult to solve because closed-form or water-filling solutions may not exist. For these cases, numerical algorithms are needed to find the optimal solution. In this paper, we formulate the beamformer design as an optimization problem with rank-1 constraints on the transmit covariance matrices. To solve this problem, we temporarily drop the rank-1 constraints, and consider a relaxed problem, which turns out to be a convex optimization problem. The relaxed problem can be easily solved using general purpose convex optimization packages, e.g., CVX [7]. Particularly, we show that the relaxed problem *always* yields rank-1 solutions, which are also optimal for the original problem.

The rest of the paper is organized as follows. The system

This research has been supported by Tekes, the Finnish Funding Agency for Technology and Innovation, Nokia Siemens Networks, Renesas Mobile Europe, Elektrobit, Xilinx, Academy of Finland, and partly by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 228044.

978-1-4577-2053-6/12/$31.00 ©2012 IEEE

model and ZF-DPC are introduced in Section II. In Section III, we consider the beamformer design for ZF-DPC with a SPC based on QRD, and prove its optimality. Section IV addresses the optimal and suboptimal beamformer designs for ZF-DPC with PAPCs. Numerical results are provided in Section V, and concluding remarks are drawn in Section VI.

Notation: Standard notations are used in this paper. Bold lower and upper case letters represent vectors and matrices, respectively; $\boldsymbol{H}^H$ and $\boldsymbol{H}^T$ are Hermitian and standard transpose of $\boldsymbol{H}$, respectively; $\mathrm{diag}(\boldsymbol{x})$ denotes a diagonal matrix with elements $\boldsymbol{x}$;

II. Zero-forcing dirty paper coding

Consider a single-cell downlink channel with an N-antenna base station and K single-antenna users. Let $\boldsymbol{h}_k \in \mathbb{C}^{N\times 1}$ be the vector channel between the BS and the kth user. The received signal at the kth user is given by

$$y_k = \boldsymbol{h}_k^H \boldsymbol{x}_k + \sum_{j\neq k} \boldsymbol{h}_k^H \boldsymbol{x}_j + n_k, \tag{1}$$

where y_k and $\boldsymbol{x}_k \in \mathbb{C}^{N\times 1}$ denote the received and transmitted signals for the kth user, respectively, and n_k is assumed to be the white complex-Gaussian noise with zero mean and unit variance. We can further write

$$\boldsymbol{x}_k = \boldsymbol{w}_k u_k, \tag{2}$$

where $\boldsymbol{w}_k \in \mathbb{C}^{N\times 1}$ is the beamforming vector and u_k (scalar-valued) is the information-bearing symbol intended for kth user. Substituting (2) into (1) gives

$$y_k = \boldsymbol{h}_k^H \boldsymbol{w}_k u_k + \sum_{j<k} \boldsymbol{h}_k^H \boldsymbol{w}_j u_j + \sum_{j>k} \boldsymbol{h}_k^H \boldsymbol{w}_j u_j + n_k \tag{3}$$

It is well known that DPC is a capacity achieving transmission strategy for MIMO BCs [2]. In fact, DPC is a coding technique that pre-cancels known interference without loss of information. For the kth user, the BS views the interference term $\sum_{j<k} \boldsymbol{h}_k^H \boldsymbol{w}_j u_j$ as known non-causally, and can be perfectly eliminated. As a result, the sum capacity is given by

$$R^{\text{dpc}} = \sum_{k=1}^{K} \log_2 \frac{(1+\sum_{j\geq k}^{K} |\boldsymbol{h}_k^H \boldsymbol{w}_j|^2)}{(1+\sum_{j>k} |\boldsymbol{h}_k^H \boldsymbol{w}_j|^2)}. \tag{4}$$

Several numerical algorithms have been proposed to find optimal beamformers that maximizes R^{dpc} in (4), which are based on a duality between a BC and the resulting multiple access channel (MAC), i.e., BC-MAC duality [8], [9]. However, these iterative algorithms suffer from high computational complexity.

ZF-DPC, derived from DPC, combines DPC and the zero-forcing technique. Specifically, the non-causal interference is canceled by DPC, while the interference $\sum_{j>k} \boldsymbol{h}_k^H \boldsymbol{w}_j u_j$ is eliminated by designing $\boldsymbol{w}_j$ such that

$$\boldsymbol{h}_k^H \boldsymbol{w}_j = 0 \text{ for all } j > k. \tag{5}$$

Consequently, the sum rate of ZF-DPC reduces to

$$R^{\text{zf-dpc}} = \sum_{k=1}^{K} \log_2(1+|\boldsymbol{h}_k^H \boldsymbol{w}_k|^2). \tag{6}$$

Stack the channel vectors of all users in a matrix $\boldsymbol{H}$ defined as

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}_1 & \boldsymbol{h}_2 & \cdots & \boldsymbol{h}_K \end{bmatrix}^H \in \mathbb{C}^{K\times N}, \tag{7}$$

and all beamforming vectors in a matrix $\boldsymbol{W}$ given by

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \cdots & \boldsymbol{w}_K \end{bmatrix} \in \mathbb{C}^{N\times K}. \tag{8}$$

Then, the zero interference constraint is equivalently written by

$$\boldsymbol{H}\boldsymbol{W} = \begin{bmatrix} \sqrt{q_1} & & & & \\ \times & \sqrt{q_2} & & & \\ \times & \times & \ddots & & \\ \times & \times & \times & \ddots & \\ \times & \times & \times & \times & \sqrt{q_K} \end{bmatrix} = \mathsf{L}(\sqrt{\boldsymbol{q}}) \tag{9}$$

where $\sqrt{\boldsymbol{q}} = \begin{bmatrix} \sqrt{q_1} & \sqrt{q_2} & \cdots & \sqrt{q_K} \end{bmatrix}^T$ and $\mathsf{L}(\sqrt{\boldsymbol{q}})$ represents a lower triangular matrix whose arguments are the diagonal elements. The beamformer design problem for ZF-DPC is to find $\boldsymbol{W}$ that maximizes a performance measure under the ZF constraints in (5). In this paper, we mainly place a focus on maximizing $R^{\text{zf-dpc}}$ under the sum power and per-antenna power constraints.

III. Sum power constraint

In this section, we address the sum rate maximization problem for ZF-DPC under a sum power constraint P, which is formulated as

$$\begin{array}{ll} \underset{\boldsymbol{w}_k,\boldsymbol{q}}{\text{maximize}} & \sum_{k=1}^{K} \log_2(1+|\boldsymbol{h}_k^H \boldsymbol{w}_k|^2) \\ \text{subject to} & \boldsymbol{H}\boldsymbol{W} = \mathsf{L}(\sqrt{\boldsymbol{q}}) \\ & \mathrm{tr}(\boldsymbol{W}\boldsymbol{W}^H) \leq P. \end{array} \tag{10}$$

Existing works regarding the design of beamformers for ZF-DPC are based on the QRD proposed in [3]. Specifically, by abuse of notation, let $\boldsymbol{H} = \boldsymbol{G}\boldsymbol{Q}$ be a QRD of $\boldsymbol{H}$, where $\boldsymbol{G}$ is a lower triangular matrix and $\boldsymbol{Q}$ is a unitary matrix. We assume that $N \geq K$ such that all diagonal entries of $\boldsymbol{G}$ are strictly larger than zero. To satisfy the ZF constraints in (9), the beamformer matrix $\boldsymbol{W}$ is designed as

$$\boldsymbol{W} = \boldsymbol{Q}^H \mathrm{diag}(1/\boldsymbol{g})\, \mathrm{diag}(\sqrt{\boldsymbol{q}}) \tag{11}$$

where $g_i = [\boldsymbol{G}]_{i,i}$. To maximize the sum rate, $\boldsymbol{q}$ is found to be the solution of the following problem

$$\begin{array}{ll} \underset{\boldsymbol{q}}{\text{maximize}} & \sum_{k=1}^{K} \log_2(1+q_k) \\ \text{subject to} & \sum_{k=1}^{K} q_k/g_k^2 \leq P \end{array}, \tag{12}$$

which can be easily solved by the water-filling algorithm. Apparently, (11) is just one of many feasible designs for $\boldsymbol{W}$ satisfying (9). Thus, a question that naturally arises is whether (11) and (12) constitute an optimal solution to (10). In the following, we will show that the QRD-based design is indeed optimal for ZF-DPC under a SPC. In fact, we consider a more general optimization problem, which is given by

$$\begin{array}{ll} \text{maximize} & f(\boldsymbol{q}) \\ \text{subject to} & \boldsymbol{H}\boldsymbol{W} = \mathsf{L}(\sqrt{\boldsymbol{q}}) \\ & \operatorname{tr}(\boldsymbol{W}\boldsymbol{W}^H) \leq P \end{array} \tag{13}$$

where $f(\boldsymbol{q})$ is an arbitrary objective function of interest. The following theorem proves the optimality of the QRD-based design for ZF-DPC under the total power constraint.

Theorem 1. *The optimal solution to the optimization problem in* (13) *is* $\boldsymbol{W}^\star = \boldsymbol{Q}^H \operatorname{diag}(1/\boldsymbol{g}) \operatorname{diag}(\sqrt{\boldsymbol{q}^\star})$*, where* $\boldsymbol{q}^\star$ *is the solution of the following problem*

$$\begin{array}{ll} \text{maximize} & f(\boldsymbol{q}) \\ \text{subject to} & \sum_{k=1}^{K} q_k/g_k^2 \leq P \end{array} \tag{14}$$

Proof: The key idea of the proof is to show that (14) is a relaxation of (13) which yields an upper bound, and this bound is tight. To see this, we need the following lemma, which describes the general structure of $\boldsymbol{W}$ that satisfies the constraint (9) (refer to Appendix A for the proof).

Lemma 1. *The general form of* $\boldsymbol{W}$ *to the zero-forcing constraints in* (9) *is given by*

$$\boldsymbol{W} = \boldsymbol{Q}^H \left[\operatorname{diag}(\sqrt{\boldsymbol{q}}/\boldsymbol{g}) + \boldsymbol{G}_L \operatorname{diag}(\sqrt{\boldsymbol{q}}) + \boldsymbol{G}^{-1}\boldsymbol{L}\right] + \boldsymbol{U}, \tag{15}$$

where $\operatorname{diag}(\sqrt{\boldsymbol{q}}/\boldsymbol{g}) = \operatorname{diag}(\sqrt{q_1}/g_1, \sqrt{q_2}/g_2, \ldots, \sqrt{q_K}/g_K)$*,* $\boldsymbol{L}$ *is strictly lower triangular matrix,* $\boldsymbol{U} \in \mathbb{C}^{N\times K}$ *is an arbitrary matrix that lies in the null space of* $\boldsymbol{Q}$*, i.e.,* $\boldsymbol{Q}\boldsymbol{U} = \boldsymbol{0}$*, and* $\boldsymbol{G}_L$ *is a strictly lower triangular matrix consisting of off-diagonal entries of* $\boldsymbol{G}^{-1}$*, i.e.,*

$$\boldsymbol{G}_L = \boldsymbol{G}^{-1} - \operatorname{diag}(1/\boldsymbol{g}). \tag{16}$$

We now proceed to prove Theorem 1. First, the sum power in (13) becomes

$$\operatorname{tr}(\boldsymbol{W}\boldsymbol{W}^H) = \sum_{k=1}^{K} q_k/g_k^2 + \operatorname{tr}\big\{(\boldsymbol{G}_L \operatorname{diag}(\sqrt{\boldsymbol{q}}) + \tilde{\boldsymbol{L}}) \times (\operatorname{diag}(\sqrt{\boldsymbol{q}})\boldsymbol{G}_L^H + \tilde{\boldsymbol{L}}^H)\big\} + \operatorname{tr}(\boldsymbol{U}\boldsymbol{U}^H) \leq P, \tag{17}$$

where $\tilde{\boldsymbol{L}} = \boldsymbol{G}^{-1}\boldsymbol{L}$. In (17), we use the fact that $\operatorname{tr}\big\{\operatorname{diag}(\sqrt{\boldsymbol{q}_i}/\boldsymbol{g}_i)(\tilde{\boldsymbol{L}}^H + \operatorname{diag}(\sqrt{\boldsymbol{q}})\boldsymbol{G}_L^H)\big\} = 0$, and $\boldsymbol{Q}\boldsymbol{U} = \boldsymbol{0}$. Since $(\boldsymbol{G}_L \operatorname{diag}(\sqrt{\boldsymbol{q}}) + \tilde{\boldsymbol{L}})(\operatorname{diag}(\sqrt{\boldsymbol{q}})\boldsymbol{G}_L^H + \tilde{\boldsymbol{L}}^H) \succeq \boldsymbol{0}$ and $\boldsymbol{U}\boldsymbol{U}^H \succeq \boldsymbol{0}$, (14) is a relaxation of (13), and generates an upper bound on its optimal value. However, this bound can be achieved by setting $\boldsymbol{L} = -\boldsymbol{G}\boldsymbol{G}_L \operatorname{diag}(\sqrt{\boldsymbol{q}})$, and $\boldsymbol{U} = \boldsymbol{0}$, which completes the proof. ■

IV. Per-antenna power constraint

As mentioned earlier, the PAPC is more realistic since each antenna has its own power amplifier [10], [11]. In this section, we are interested in the problem of sum rate maximization under PAPCs, which is expressed as

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^{K} \log_2(1 + q_i) \\ \text{subject to} & \boldsymbol{H}\boldsymbol{W} = \mathsf{L}(\sqrt{\boldsymbol{q}}) \\ & [\boldsymbol{W}\boldsymbol{W}^H]_{n,n} \leq P_n,\ n = 1, 2, \ldots, N \end{array} \tag{18}$$

where P_n is the power constraint for the nth antenna at the BS. Obviously, the beamformer design derived from the QRD becomes suboptimal under PAPCs.

A. Optimal design

Now, we present a numerical algorithm to find an optimal solution to the sum rate maximization in (18). In [6], a convex optimization-based design for zero-forcing beamforming was presented, which can be applied to solve (18). First, let $\boldsymbol{S}_k = \boldsymbol{w}_k\boldsymbol{w}_k^H$ be the covariance matrix of the kth user. Then (18) is equivalent to

$$\begin{array}{ll} \text{maximize} & \sum_{k=1}^{K} \log_2(1 + \boldsymbol{h}_k^H \boldsymbol{S}_k \boldsymbol{h}_k) \\ \text{subject to} & \boldsymbol{h}_i^H \boldsymbol{S}_k \boldsymbol{h}_i = 0\ \forall i < k \\ & \sum_{k=1}^{K} [\boldsymbol{S}_k]_{n,n} \leq P_n\ \forall n \\ & \boldsymbol{S}_k \succeq 0\ \forall k \\ & \operatorname{rank}(\boldsymbol{S}_k) = 1\ \forall k, \end{array} \tag{19}$$

Recall that the rank-1 constraints in (19) are non-convex, and, thus, it is generally NP-hard to solve (19). The method in [6] is based on a relaxation method. Specifically, the relaxed problem is formed by dropping the rank-1 constraints in (19), resulting in

$$\begin{array}{ll} \underset{\boldsymbol{S}_k}{\text{maximize}} & \sum_{k=1}^{K} \log_2(1 + \boldsymbol{h}_k^H \boldsymbol{S}_k \boldsymbol{h}_k) \\ \text{subject to} & \boldsymbol{h}_i^H \boldsymbol{S}_k \boldsymbol{h}_i = 0\ \forall i < k \\ & \sum_{k=1}^{K} [\boldsymbol{S}_k]_{n,n} \leq P_n\ \forall n \\ & \boldsymbol{S}_k \succeq 0\ \forall k. \end{array} \tag{20}$$

Problem (20) is a convex optimization problem, and thus can be solved efficiently using standard optimization packages, e.g., CVX [7] or YALMIP [12]. If the optimal solution $\boldsymbol{S}_k^\star$ of (20) is a rank-1 matrix, then it is also optimal for (19). If $\boldsymbol{S}_k^\star$ has a rank larger than 1, then consider the following optimization problem

$$\begin{array}{ll} \underset{\boldsymbol{t}}{\text{maximize}} & \Re(\boldsymbol{h}_k^H \boldsymbol{t}) \\ \text{subject to} & \boldsymbol{h}_i^H \boldsymbol{t} = 0\ \forall i < k \\ & |[\boldsymbol{t}]_n|^2 \leq [\boldsymbol{S}_k^\star]_{n,n}, \end{array} \tag{21}$$

where $\Re(x)$ is the real part of x. Let t_k be the optimal solution to the above problem. Then, it is proved that $\boldsymbol{S}_k = \boldsymbol{t}_k\boldsymbol{t}_k^H$ is the rank-1 optimal solution to (20), which is also optimal to (19). The proof can be found in [6].

In fact, our experimental results with numerical optimization packages to solve (20) *always* yield rank-1 solutions. Actually, we can prove that the optimal solutions of (20) are always rank-1. To prove this, we first reformulate (20) as

$$\begin{array}{ll} \underset{\tilde{\boldsymbol{S}}_k}{\text{maximize}} & \log_2 \left| \boldsymbol{I} + \operatorname{diag}(\tilde{\boldsymbol{h}}_1^H \tilde{\boldsymbol{S}}_1 \tilde{\boldsymbol{h}}_1, \ldots, \tilde{\boldsymbol{h}}_K^H \tilde{\boldsymbol{S}}_K \tilde{\boldsymbol{h}}_K) \right| \\ \text{subject to} & \sum_{k=1}^{K} [\boldsymbol{V}_k \tilde{\boldsymbol{S}}_k \boldsymbol{V}_k^H]_{n,n} \leq P_n\ \forall n \\ & \tilde{\boldsymbol{S}}_k \succeq 0\ \forall i. \end{array} \tag{22}$$

where $\boldsymbol{S}_k = \boldsymbol{V}_k \tilde{\boldsymbol{S}}_k \boldsymbol{V}_k^H$, and $\boldsymbol{V}_k \in \mathbb{C}^{N\times(N-k+1)}$ is a basis of the null space of a matrix defined as $\boldsymbol{H}_k = \begin{bmatrix}\boldsymbol{h}_1 & \boldsymbol{h}_2 & \cdots & \boldsymbol{h}_{k-1}\end{bmatrix}^H \in \mathbb{C}^{(k-1)\times N}$, and $\tilde{\boldsymbol{h}}_i = \boldsymbol{V}_i^H \boldsymbol{h}_i$.[1]

[1] $\boldsymbol{V}_k$ can be computed efficiently using QRD or singular value decomposition of $\boldsymbol{H}_k$.

Note that solving (22) is more computationally efficient than solving (20) since $\tilde{\boldsymbol{S}}_k$ in (22) has a lower dimension than $\boldsymbol{S}_k$ in (20). Now, it is sufficient to show the following lemma.

Lemma 2. *The optimal solutions* $\tilde{\boldsymbol{S}}_k^\star$ *to* (22) *satisfy* $\operatorname{rank}(\tilde{\boldsymbol{S}}_k^\star) \le 1$ for all $1 \le k \le K$.

Proof: See Appendix B. ■

B. Suboptimal designs

In practice, it is of particular interest to find a suboptimal design for ZF-DPC that performs close to the optimal solution, but requires lower complexity. Herein, we present a two suboptimal designs which are easier to solve. The first one is derived from the QRD-based design for the total power constraint. The second one is based on an approximation of the sum rate in the high SNR regime.

1) Suboptimal design I: The first suboptimal design is obtained by first setting $\boldsymbol{L} = -\boldsymbol{G}\boldsymbol{G}_L \operatorname{diag}(\sqrt{\boldsymbol{q}})$ and $\boldsymbol{U} = \boldsymbol{0}$ in (14), i.e., $\boldsymbol{W} = \boldsymbol{Q}^H \operatorname{diag}(1/\boldsymbol{g}) \operatorname{diag}(\sqrt{\boldsymbol{q}})$, and then find the power allocation vector $\boldsymbol{q}$ to meet the PAPCs. With $\boldsymbol{W} = \boldsymbol{Q}^H \operatorname{diag}(1/\boldsymbol{g}) \operatorname{diag}(\sqrt{\boldsymbol{q}})$, the constraint $[\boldsymbol{W}\boldsymbol{W}^H]_{n,n} \le P_n$ is equivalent to

$$[\boldsymbol{W}\boldsymbol{W}^H]_{n,n} = \sum_k |\boldsymbol{Q}_{k,n}|^2 q_k / g_k^2 \le P_n \tag{23}$$

Define a matrix $\boldsymbol{A} \in R^{N\times K}$ as $[\boldsymbol{A}]_{i,j} = |\boldsymbol{Q}_{j,i}|^2/g_j^2$, and $\boldsymbol{p} = [\ P_1 \quad P_2 \quad \ldots \quad P_N\ ]^T$, then (23) can be rewritten in a more compact form as

$$\boldsymbol{A}\boldsymbol{q} \le \boldsymbol{p}. \tag{24}$$

The power allocation vector is found to be the solution of the following problem

$$\begin{array}{ll} \underset{\boldsymbol{q}\ge 0}{\text{maximize}} & \sum_{i=1}^K \log_2(1+q_i) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{q} \le \boldsymbol{p}. \end{array} \tag{25}$$

The complexity of solving (25) is greatly lower than that of solving (22) since the number of optimization variables in (25) is only K, while that in (22) is $\sum_{k=1}^K \frac{1}{2}(N-k+1)\times(N-k+2)$. Let R_{sub_I}, R_{PAPC} and R_{sum} denote the optimal value of (25), (19), and (14), respectively. Assume the total power constraint in (14) is the sum of all PAPCs in (19), i.e., $P = \sum_{n=1}^N P_n$. Then we immediately have

$$R_{\text{sub}_\text{I}} \le R_{\text{PAPC}} \le R_{\text{sum}}. \tag{26}$$

Assuming equal power constraints at each antenna, i.e., $P_n = P/N$ for all $n = 1, 2, \ldots, N$, we can bound the gap $R_{\text{sum}} - R_{\text{subI}}$ by

$$\begin{aligned} R_{\text{sum}} - R_{\text{sub}_\text{I}} &\le N\Big(\frac{P}{\max\gamma_n + P}\Big)\left[\max_k g_k^2 \max_n \gamma_n - \frac{K}{N}\right] \\ &\overset{P\to\infty}{\longrightarrow} N\left[\max_k g_k^2 \max_n \gamma_n - \frac{K}{N}\right], \end{aligned} \tag{27}$$

where $\gamma_n = \sum_k [\boldsymbol{A}]_{n,k}$. The proof follows similar arguments as those in [6], and is omitted here for brevity.

2) Suboptimal design II: The second suboptimal design is derived from maximizing an objective function which is proportional to the sum rate of ZF-DPC. Interestingly, the second suboptimal design can be solved efficiently using second-order cone programing (SOCP). In the high SNR regime, we can approximate $\log_2(1+q_i) \approx \log_2 q_i$, and consider the following optimization

$$\begin{array}{ll} \text{maximize} & \prod_{k=1}^K |\boldsymbol{h}_k^H \boldsymbol{w}_k| \\ \text{subject to} & \boldsymbol{h}_i^H \boldsymbol{w}_k = 0,\ i < k \\ & \sum_{k=1}^K |[\boldsymbol{w}_k]_n|^2 \le P_n,\ n = 1, \ldots, N. \end{array} \tag{28}$$

Removing the ZF constraints in (28) generates an equivalent problem as

$$\begin{array}{ll} \underset{\tilde{\boldsymbol{w}}_k}{\text{maximize}} & \prod_{k=1}^K |\tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{w}}_k| \\ \text{subject to} & \sum_{k=1}^K |[\boldsymbol{V}_k \tilde{\boldsymbol{w}}_k]_n|^2 \le P_n,\ \forall n \end{array} \tag{29}$$

where $\tilde{\boldsymbol{h}}_k = \boldsymbol{V}_k^H \boldsymbol{h}_k$, and $\boldsymbol{w}_k = \boldsymbol{V}_k \tilde{\boldsymbol{w}}_k$. Without loss of optimality we can assume that $\tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{w}}_k$ is real[2], and reformulate (28) as

$$\begin{array}{ll} \text{maximize} & (\prod_{k=1}^K p_k) \\ \text{subject to} & \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{w}}_k \ge p_k \\ & \sum_{k=1}^K |[\boldsymbol{V}_k \tilde{\boldsymbol{w}}_k]_n|^2 \le P_n,\ \forall n \end{array} \tag{30}$$

An efficient way to solve (30) is to recast it as a SOCP problem, which can be efficiently solved via specialized interior-point methods. For example, considering the case $K = 4$, and using hyperbolic constraints [13], we obtain

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & \left\| \begin{bmatrix} 2t_1 \\ \tilde{\boldsymbol{h}}_1^H \tilde{\boldsymbol{w}}_1 - \tilde{\boldsymbol{h}}_2^H \tilde{\boldsymbol{w}}_2 \end{bmatrix} \right\| \le \tilde{\boldsymbol{h}}_1^H \tilde{\boldsymbol{w}}_1 + \tilde{\boldsymbol{h}}_2^H \tilde{\boldsymbol{w}}_2, \\ & \left\| \begin{bmatrix} 2t_2 \\ \tilde{\boldsymbol{h}}_3^H \tilde{\boldsymbol{w}}_3 - \tilde{\boldsymbol{h}}_4^H \tilde{\boldsymbol{w}}_4 \end{bmatrix} \right\| \le \tilde{\boldsymbol{h}}_3^H \tilde{\boldsymbol{w}}_3 + \tilde{\boldsymbol{h}}_4^H \tilde{\boldsymbol{w}}_4, \\ & \left\| \begin{bmatrix} 2t \\ t_1 - t_2 \end{bmatrix} \right\| \le t_1 + t_2, \\ & \sum_{k=1}^K |[\boldsymbol{V}_k \tilde{\boldsymbol{w}}_k]_n|^2 \le P_n,\ n = 1, 2, \ldots, N. \end{array} \tag{31}$$

The extension to other values of K is straightforward using the concept of second-order cone representable functions [13]. Solving (31) is more computationally efficient than solving (22), since the number of variables in (31) is much smaller than that of (22), especially when N is large. Recall that $\tilde{\boldsymbol{S}}_k$ in (22) is an $(N-k+1)\times(N-k+1)$ Hermitian matrix, while $\tilde{\boldsymbol{w}}_k$ in (31) is an $N\times 1$ vector.

V. Numerical results

In this section, we provide numerical examples to demonstrate the results in this paper. In Figs. 1 and 2, we draw the average sum rate of optimal and suboptimal beamformer

[2]If $\tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{w}}_i$ is a complex number, we can write $\tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{w}}_i = |\tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{w}}_i| e^{j\theta_i}$, and let $\bar{\boldsymbol{w}}_i = \tilde{\boldsymbol{w}}_i e^{-j\theta_i}$. Obviously, $\{\bar{\boldsymbol{w}}_i\}$ satisfy all the constraints and achieve the same objective value as that in (28).

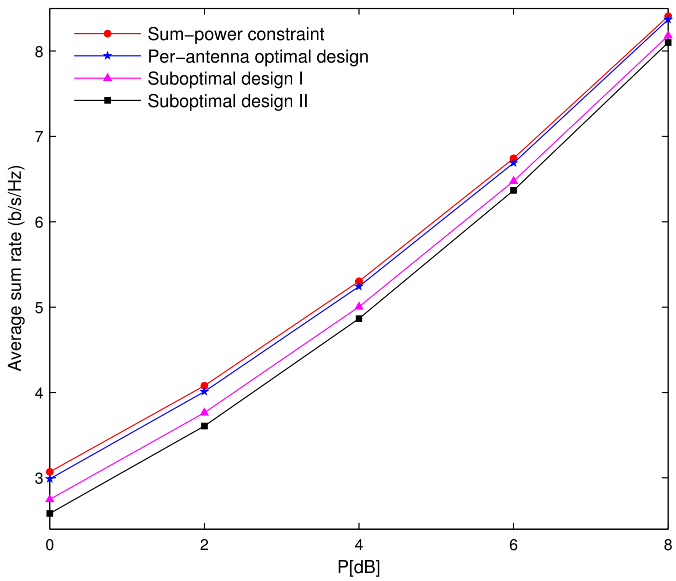

Fig. 1. Sum rate comparison of optimal and suboptimal designs for ZF-DPC with equal power constraints, $N = K = 4$.

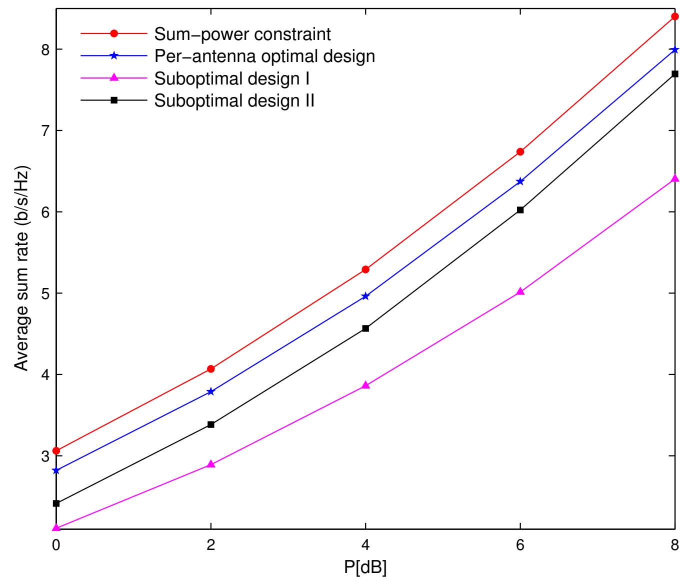

Fig. 2. Sum rate comparison of optimal and suboptimal designs for ZF-DPC with unequal power constraints, $K = 4$.

design methods for ZF-DPC schemes as a function of P, the total transmit power. A quasi-static fading model is used in our simulation, where independent realizations of $\boldsymbol{h}_k$ are generated as zero mean and unit variance complex Gaussian random variables for each snapshot. We consider a system model with $N = K = 4$. The power constraint for the nth antenna in Fig. 1 is $P_n = P/N$, i.e., equal power constraints, and in Fig. 2 is $P_n = \frac{P}{\sum_{k=1}^{N} k} n$, i.e the power constraint at the nth antenna is proportional to n (unequal power constraints).

For equal power constraints shown in Fig. 1, we can see that the optimal beamformer design with PAPCs yields almost the same sum rate as that with the sum power constraint, especially as P increases since the equal power allocation is proved to be optimal in the high SNR regime. Moreover, Fig. 1 indicates that suboptimal design I is slightly better than suboptimal design II, and and both of them achieve a significant fraction of sum rate of the optimal design. This is partly due to the fact that the PAPCs are likely to be active at the optimum in both suboptimal designs (recall that in this case the matrix $\boldsymbol{A}$ in (24) is invertible). The results with unequal power constraints in Fig. 2 reveal a remarkable difference between the sum rate with the SPC and that with PAPCs. When the power at each antenna is constrained differently from each other, the suboptimal design I becomes inferior to other designs. However, this suboptimal design requires significantly lower complexity than other design methods.

VI. Conclusions

We have addressed the beamformer design for ZF-DPC under a sum power constraint and per-antenna power constraints. For the SPC, we prove that the QRD-based design, introduced in [3], is optimal. For PAPCs, the QRD-based design is no longer optimal and we propose an optimal beamformer design based on a convex-optimization framework, which involves a relaxation technique. The relaxed problem is shown to be equivalent to the original problem. Numerical results indicate that the sum rates of ZF-DPC under a SPC and PAPCs are almost the same with equal power constraints at each antenna, and remarkably different with unequal power constraints. In addition, we present two suboptimal beamformer designs for ZF-DPC, one based on the QRD and the other based on maximizing the product of the effective channel gains. The future research items include the decomposition of the computation to enable distributed implementation in the CoMP context with limited antenna cooperation.

Appendix A
Proof of Lemma 1

First, from (9) and the fact that $\boldsymbol{G}$ is invertible, the ZF constraints can be rewritten as

$$\begin{aligned} \boldsymbol{QW} &= \boldsymbol{G}^{-1}\mathsf{L}(\sqrt{\boldsymbol{q}}) \\ &= \boldsymbol{G}^{-1}(\mathrm{diag}(\sqrt{\boldsymbol{q}}) + \boldsymbol{L}) \\ &= \boldsymbol{G}^{-1}\,\mathrm{diag}(\sqrt{\boldsymbol{q}}) + \tilde{\boldsymbol{L}}, \end{aligned} \quad (32)$$

where $\boldsymbol{L}$ and $\tilde{\boldsymbol{L}} = \boldsymbol{G}^{-1}\boldsymbol{L}$ are strictly lower triangular matrices. Since $\boldsymbol{G}$ is a lower triangular matrix, $\boldsymbol{G}^{-1}$ is also lower triangular and

$$\boldsymbol{G}^{-1} = \mathrm{diag}(1/\boldsymbol{g}) + \boldsymbol{G}_L, \quad (33)$$

where $\boldsymbol{G}_L$ is a strictly lower triangular matrix. Plugging (33) into (32) yields

$$\boldsymbol{QW} = \mathrm{diag}(\sqrt{\boldsymbol{q}}/\boldsymbol{g}) + \boldsymbol{G}_L\,\mathrm{diag}(\sqrt{\boldsymbol{q}}) + \tilde{\boldsymbol{L}}. \quad (34)$$

The general form of a solution for (34) is given by

$$\boldsymbol{W} = \boldsymbol{Q}^H(\mathrm{diag}(\sqrt{\boldsymbol{q}}/\boldsymbol{g}) + \boldsymbol{G}_L\,\mathrm{diag}(\sqrt{\boldsymbol{q}}) + \tilde{\boldsymbol{L}}) + \boldsymbol{U}, \quad (35)$$

where $\boldsymbol{U} \in \mathbb{C}^{N \times K}$ is an arbitrary matrix that lies in $\mathcal{N}(\boldsymbol{Q})$, i.e., $\boldsymbol{QU} = \boldsymbol{0}$.

APPENDIX B
PROOF OF LEMMA 2

In this appendix, we prove that the rank of optimal solutions to (22) is less than or equal to 1. The proof follows similar arguments to those in [14]. We begin by reformulating (22) as

$$\begin{array}{ll} \underset{\tilde{\boldsymbol{S}}_k}{\text{maximize}} & \sum_{k=1}^{K} \log(1 + \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{S}}_k \tilde{\boldsymbol{h}}_k) \\ \text{subject to} & \sum_{k=1}^{K} \mathrm{tr}(\tilde{\boldsymbol{S}}_k \boldsymbol{A}_k^{(n)}) \le P_n, \ \forall n \\ & \tilde{\boldsymbol{S}}_k \succeq \boldsymbol{0}, \ \forall k, \end{array} \quad (36)$$

where $\tilde{\boldsymbol{h}}_k = \boldsymbol{V}_k^H \boldsymbol{h}_k \in \mathbb{C}^{(N-k+1)\times 1}$, and $\boldsymbol{A}_k^{(n)}$ is defined as

$$\boldsymbol{A}_k^{(n)} \triangleq \boldsymbol{V}_k^H \,\mathrm{diag}(\underbrace{0, \ldots, 0}_{n-1}, 1, \underbrace{0, \ldots, 0}_{N-n}) \boldsymbol{V}_k.$$

The Lagrangian function of (36) is given by

$$\mathcal{L}(\tilde{\boldsymbol{S}}_k, \boldsymbol{\lambda}, \boldsymbol{\Phi}_k) = \sum_{k=1}^{K} \log(1 + \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{S}}_k \tilde{\boldsymbol{h}}_k) - \sum_{n=1}^{N} \lambda_n \left(\sum_{k=1}^{K} \mathrm{tr}(\tilde{\boldsymbol{S}}_k \boldsymbol{A}_k^{(n)}) - P_n \right) + \mathrm{tr}(\boldsymbol{\Phi}_k \tilde{\boldsymbol{S}}_k), \quad (37)$$

where $\boldsymbol{\lambda}$ are dual variables associated with the PAPCs, and $\boldsymbol{\Phi}_k \succeq \boldsymbol{0}$ is the dual variable for the positive semidefinite constraint. Denote $\boldsymbol{P} = \mathrm{diag}(P_1, P_2, \ldots, P_N)$, $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$, and $\boldsymbol{\Lambda}_k = \boldsymbol{V}_k^H \boldsymbol{\Lambda} \boldsymbol{V}_k$. We can then rewrite (37) as

$$\mathcal{L}(\tilde{\boldsymbol{S}}_k, \boldsymbol{\lambda}, \boldsymbol{\Phi}_k) = \sum_{k=1}^{K} \log(1 + \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{S}}_k \tilde{\boldsymbol{h}}_k) - \mathrm{tr}(\boldsymbol{\Lambda}_k \tilde{\boldsymbol{S}}_k - \boldsymbol{\Lambda P}) + \mathrm{tr}(\boldsymbol{\Phi}_k \tilde{\boldsymbol{S}}_k). \quad (38)$$

We now show that the dual optimal variables of (22) are strictly positive, $\lambda_n > 0$ for all $1 \le n \le N$. As proof, consider the dual objective of (22), which can be expressed as,

$$g(\boldsymbol{\lambda}, \boldsymbol{\Phi}_k) = \max \ \mathcal{L}(\tilde{\boldsymbol{S}}_k, \boldsymbol{\lambda}, \boldsymbol{\Phi}_k). \quad (39)$$

By contradiction, suppose $\lambda_i = 0$ for some $1 \le i \le N$. We construct a set of $\tilde{\boldsymbol{S}}_k$ such that $\tilde{\boldsymbol{S}}_1 = \mathrm{diag}(\underbrace{0, \ldots, 0}_{i-1}, \alpha, \underbrace{0, \ldots, 0}_{N-i})$, and $\tilde{\boldsymbol{S}}_k = \boldsymbol{0}$ for $2 \le k \le K$. Then, the objective function in (39) becomes

$$\mathcal{L}(\tilde{\boldsymbol{S}}_k, \boldsymbol{\lambda}, \boldsymbol{\Phi}_k) = \log(1 + \alpha |h_{1,i}|^2) + \mathrm{tr}(\boldsymbol{\Phi}_1 \tilde{\boldsymbol{S}}_1). \quad (40)$$

We can see that the objective function in (40) is unbounded above if $\alpha \to \infty$.

Since we are only interested in the case where $g(\boldsymbol{\lambda}, \boldsymbol{\Phi}_k)$ is finite, we conclude that $\lambda_i > 0$ for all $1 \le i \le N$. We continue with the proof of Lemma 2. At the optimum, we have

$$\frac{1}{1 + \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{S}}_k \tilde{\boldsymbol{h}}_k} \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^H - \boldsymbol{\Lambda}_k + \boldsymbol{\Phi}_k = \boldsymbol{0}. \quad (41)$$

Using the complementary slackness property $\boldsymbol{\Phi}_k \tilde{\boldsymbol{S}}_k = \boldsymbol{0}$, we obtain

$$\frac{1}{1 + \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{S}}_k \tilde{\boldsymbol{h}}_k} \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^H \tilde{\boldsymbol{S}}_k = \boldsymbol{\Lambda}_k \tilde{\boldsymbol{S}}_k. \quad (42)$$

Since $\boldsymbol{\Lambda}$ must be positive definite, $\boldsymbol{\Lambda}_k$ is invertible. It follows from (42) that $\mathrm{rank}(\tilde{\boldsymbol{S}}_k^\star) \le \mathrm{rank}(\tilde{\boldsymbol{h}}_k) = 1$, which completes the proof.

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun*, vol. 10, pp. 585–598, Nov. 1999.

[2] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

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

[4] J. Jiang, R. Buehrer, and W. Tranter, "Greedy scheduling performance for a zero-forcing dirty-paper coded system," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 789–793, May 2006.

[5] Z. Tu and R. Blum, "Multiuser diversity for a dirty paper approach," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 370–372, Aug. 2003.

[6] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409 –4418, 2008.

[7] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx, Apr. 2011.

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

[9] W. Yu, "Sum-capacity computation for the Gaussian vector broadcast channel via dual decomposition," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 754 –759, Feb. 2006.

[10] F. Boccardi and H. Huang, "Zero-forcing precoding for the MIMO broadcast channel under per-antenna power constraints," in *IEEE SPAWC 2006*, 2006, pp. 1–5.

[11] 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.

[12] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan., 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip

[13] S. B. M. Lobo, L. Vandenberghe and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 248, pp. 193–228, Nov. 1998.

[14] M. Vu, "MISO capacity with per-antenna power constraint," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1268–1274, May 2011.