

Low-Complexity Robust MISO Downlink Precoder Design With Per-Antenna Power Constraints

Mostafa Medra Timothy N. Davidson

Abstract—This paper considers the design of the beamformers for a multiple-input single-output (MISO) downlink system that seeks to mitigate the impact of the imperfections in the channel state information (CSI) that is available at the base station (BS). The goal of the design is to minimize the outage probability of specified signal-to-interference-and-noise ratio (SINR) targets, while satisfying per-antenna power constraints (PAPCs), and to do so at a low computational cost. Based on insights from the offset maximization technique for robust beamforming, and observations regarding the structure of the optimality conditions, low-complexity iterative algorithms that involve the evaluation of closed-form expressions are developed. To further reduce the computational cost, algorithms are developed for per-antenna power-constrained variants of the zero-forcing (ZF) and maximum ratio transmission (MRT) beamforming directions. In the MRT case, our low-complexity version for systems with a large number of antennas may be of independent interest. The proposed algorithms are extended to systems with both PAPCs and a total power constraint. Simulation results show that the proposed robust designs can provide substantial gains in the outage probability while satisfying the PAPCs.

Index Terms—Broadcast channel, downlink beamforming, robust precoding, outage, per-antenna power constraints, massive MIMO, zero-forcing, maximum ratio transmission.

I. INTRODUCTION

The spatial multiplexing capabilities of base stations (BSs) with multiple antennas offer the potential for substantial gains in the quality of service (QoS) that can be offered to users in a downlink system; e.g., [1]. In particular, linear beamforming schemes have been developed to simultaneously serve multiple users at their requested signal-to-interference-and-noise ratio (SINR) targets [2]–[6]. However, the performance of those beamforming schemes can be quite sensitive to the accuracy of the channel state information (CSI) that is available at the BS. Since that information is typically obtained through estimation on the uplink (in time division duplexing, TDD, systems) or through estimation on the downlink and quantized feedback (in frequency division duplexing, FDD, systems), the CSI at the BS is inherently uncertain. That observation has spawned the development of a variety of design strategies that incorporate different models for the uncertainty into the design. One strategy is to require the requested SINR to be met for all channels that are within a specified distance of

the BS's model for the channel [7]–[12]. However, in many scenarios that is a rather conservative approximation of the outage that occurs in practice. Furthermore, although this strategy, or a mild approximation thereof, often results in a convex optimization problem for finding the beamformers, the computational cost of solving those problems can be quite significant. Fortunately, different approaches to approximating the outage probability can yield alternative design strategies that provide excellent performance in practice, even when the uncertainties in the CSI are quite substantial, and do so in a computationally inexpensive way. One such strategy is the offset maximization algorithm [13], in which the beamformers are designed to maximize a carefully structured offset on the performance specification (see Section II-B).

The above-mentioned design strategies seek to jointly design the beamforming directions and the power allocated to each direction. However, significant reductions in the computational cost can be obtained by computing the beamforming directions using a (computationally cheap) conventional technique and then developing a robust power loading algorithm. The beamforming directions in this approach are typically chosen to be either the maximum ratio transmission (MRT) [14] or zero-forcing (ZF) directions [15]. For the case of additive Gaussian uncertainties in the BS's CSI, single-integral expressions for the outage probability can be obtained [16] and an effective algorithm for finding the power loading that minimizes the power required to meet the specified outage constraint has been developed [17]. However, that algorithm is rather computationally expensive. In [18], insights from bounds on the cumulative distributive function were used to develop a new robust power loading technique that provides performance close to that of the optimal algorithm in [17], but has significantly lower computational cost.

The existing literature on robust downlink beamforming has tended to focus on designs that impose a constraint on the total power transmitted by the BS. In practice, each antenna will typically be driven by its own power amplifier, and hence the design ought to include constraints on the power transmitted from each antenna, as well as the total power. In the case of perfect CSI, a number of downlink beamforming algorithms that incorporate per-antenna power constraints (PAPCs) have been developed [19]–[23]. For robust beamforming designs that can be formulated as convex problems (e.g., [8], [9], [12]) and are solved using generic solvers, incorporating these additional constraints is quite straightforward. However, doing so increases the computational cost of what are, in comparison to the perfect CSI case, already quite expensive algorithms. The goal of this paper is to develop robust beamforming designs

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant RGPIN-2015-06631.

A preliminary version of a portion of this work appears in *Proc. 2015 IEEE Int. Wkshp Signal Process. Adv. Wireless Commun.*

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada. (E-mail: medramm@mcmaster.ca; davidson@mcmaster.ca).

that incorporate PAPCs and have reasonable computational costs. Our technique is based on insights developed from the offset maximization approach to robust beamforming [13], a closely related power loading technique [18], and observations regarding the structure of the optimality conditions for the design problem. These observations enable us to develop a low-complexity dual update optimization strategy related to that in [23] that involves the evaluation of a sequence of closed-form expressions. After extending that algorithm to systems that have both PAPCs and a total power constraint, we make the observation that a large fraction of the computational cost arises from the design of the beamforming directions. To reduce that cost, we develop PAPCed variants of the ZF and MRT directions, and show how these can be incorporated into our design approach. Furthermore, we develop a low-complexity version of our PAPCed MRT beamforming algorithm for “massive MIMO” systems with a large number of antennas. As scaling techniques for large MRT beamformers have been recently proposed [24], that algorithm may be of independent interest.

II. SYSTEM MODEL AND DESIGN APPROACH

We consider a narrowband multiple-input single-output (MISO) downlink in which an N_t -antenna BS sends independent messages to K single-antenna users. The transmitted signal at a given signalling instant is constructed using linear beamforming as $\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k$, where s_k is the power-normalized data symbol for user k , and \mathbf{w}_k is the associated beamformer. In some settings we will refer to $\mathbf{u}_k = \mathbf{w}_k / \|\mathbf{w}_k\|$ as the direction of the beamformer, and $\beta_k = \mathbf{w}_k^H \mathbf{w}_k$ as the power allocated to that direction. That enables us to write

$$\mathbf{w}_k = \sqrt{\beta_k} \mathbf{u}_k.$$

The received signal at user k can be written as

$$y_k = \mathbf{h}_k^H \mathbf{w}_k s_k + \sum_{j \neq k} \mathbf{h}_k^H \mathbf{w}_j s_j + n_k, \quad (1)$$

where \mathbf{h}_k^H denotes the channel between the BS and receiver k , and n_k represents the additive zero-mean circular complex Gaussian noise at that user.

In the problems that we will consider, each user specifies the SINR that it will require in order to support the service that it desires. This constraint takes the form

$$\text{SINR}_k = \frac{\mathbf{h}_k^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_k}{\mathbf{h}_k^H (\sum_{j \neq k} \mathbf{w}_j \mathbf{w}_j^H) \mathbf{h}_k + \sigma_k^2} \geq \gamma_k, \quad (2)$$

where σ_k^2 is the noise variance at receiver k , and γ_k is the required SINR. We will find it convenient to rewrite that constraint as

$$\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2 \geq 0,$$

where

$$\mathbf{Q}_k = \mathbf{w}_k \mathbf{w}_k^H / \gamma_k - \sum_{j \neq k} \mathbf{w}_j \mathbf{w}_j^H. \quad (3)$$

If we denote the signal transmitted from antenna i by x_i , then the power constraint on the BS as a whole can be written as $\sum_{i=1}^{N_t} E\{|x_i|^2\} = \sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \leq P_t$, where we have used

the assumptions that the messages are independent and that the symbols s_k are normalized. If we let p_i denote the maximum power that can be transmitted from antenna i , the PAPC can be written as $E\{|x_i|^2\} = [\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H]_{i,i} \leq p_i$, where $[\cdot]_{i,i}$ denotes the (i,i) th entry of the given matrix.

In order for a BS to be able to evaluate whether a candidate set of beamformers $\{\mathbf{w}_k\}_{k=1}^K$ satisfies the K SINR constraints in (2), the BS must know each channel vector \mathbf{h}_k ; e.g., [2]. However, typically the BS will only have an estimate of each channel, denoted \mathbf{h}_{e_k} . To incorporate the uncertainty in that channel estimate into the design, we will postulate a conditional distribution, $p(\mathbf{h}_k | \mathbf{h}_{e_k})$, and convert the deterministic QoS constraint $\text{SINR}_k \geq \gamma_k$ into the chance constraint $\text{Prob}(\text{SINR}_k \geq \gamma_k) \geq 1 - \delta_k$, where δ_k is the required outage probability. In this paper, we will model the uncertainty additively; i.e.,

$$\mathbf{h}_k = \mathbf{h}_{e_k} + \mathbf{e}_k, \quad (4)$$

with \mathbf{e}_k having zero-mean and being independent of the channel and data. Our results will focus on the case where \mathbf{e}_k is a zero-mean circular Gaussian random variable of covariance $\sigma_{e_k}^2 \mathbf{I}$. Among a number of scenarios, that model is appropriate in certain TDD systems in which channels are estimated during the uplink training phase.

A. Design approach

With the uncertainty modeled as described above, one approach to the design of the downlink beamformers \mathbf{w}_k is to seek to minimize the probability of outage of the SINR targets, subject to a total power constraint and PAPCs; i.e.,

$$\min_{\mathbf{w}_k, \delta_k} \max_k \delta_k \quad (5a)$$

$$\text{s.t.} \quad \sum_k \mathbf{w}_k^H \mathbf{w}_k \leq P_t, \quad (5b)$$

$$\left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} \leq p_i, \quad \forall i, \quad (5c)$$

$$\text{Prob}(\text{SINR}_k \geq \gamma_k) \geq 1 - \delta_k, \quad \forall k. \quad (5d)$$

This problem is hard to solve even without the PAPCs. However, in the case that the PAPCs are omitted, the offset maximization algorithm [13] is a low-complexity algorithm that has been shown to provide good performance. The goal of this paper is to use insights from the development of the offset maximization approach to develop an effective low-complexity algorithm for the PAPCed case. One observation that we will use is that the performance of the offset maximization approach can be improved by applying the robust power loading algorithm in [18] to the beamforming directions generated by the offset maximization. Doing so reveals that robust beamformers can be obtained with a computational cost that is similar to that of beamformer design in the perfect CSI case. (Many existing approaches to robust beamforming are much more expensive than the perfect CSI case; e.g., [11], [12].) However, like the perfect CSI case, it is the computation of the directions that dominate the computational cost. Therefore, we also propose to apply the principles that underlie the power loading in [18] to beamforming directions that can be computed more efficiently, such as PAPCed variants,

derived herein, of the classical ZF and MRT directions; see Sections IV, and V. In the latter case, a further approximation that is suitable for scenarios with a large number of antennas at the BS substantially reduces the computational cost, and has almost the same outage performance.

To lay the groundwork for the development of the proposed beamforming schemes, in the following subsections we briefly review the offset maximization approach to beamformer design under a total power constraint [13], and the low-complexity robust power loading technique for systems with a total power constraint that was developed in [18].

B. Offset maximization beamforming directions

The offset maximization beamformers [13] can be found by solving the following problem:

$$r_t^* = \max_{\mathbf{w}_k, r} r \quad (6a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \leq P_t, \quad (6b)$$

$$\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 - r \geq 0, \quad \forall k. \quad (6c)$$

It is implicit in (6c) that this algorithm tries to find the largest noise-plus-interference power each user can endure, under the total power constraint. In [13] an efficient method to solve (6) was developed by considering the following problem, in which, for now, it is assumed that the optimal value for (6), r_t^* , is known:

$$P^* = \min_{\mathbf{w}_k} \sum_k \mathbf{w}_k^H \mathbf{w}_k \quad (7a)$$

$$\text{s.t.} \quad \mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 - r_t^* \geq 0, \quad \forall k. \quad (7b)$$

It can be shown [13] that the optimal value of the problem in (7) is P_t , and that any set of beamformers that optimize (7) are also optimal for (6). Also, at optimality, all the constraints are satisfied with equality.

The advantage of the connection between problems (6) and (7) is that a highly efficient algorithm for the problem in (7) with $r = 0$ (i.e., the perfect CSI case) was developed in [3]; see also [6]. That algorithm can be extended to jointly find the optimal beamformers and the optimal offset, r_t^* , for the problem in (6). In particular, if we let ν_k denote the Lagrange multiplier for the SINR constraint in (7b), then from the KKT conditions of (7) we can find the offset maximization directions by solving the eigen problem

$$\mathbf{u}_k = \left(\frac{\nu_k}{\gamma_k} \mathbf{h}_{e_k} \mathbf{h}_{e_k}^H - \sum_{j \neq k} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H \right) \mathbf{u}_k, \quad (8)$$

where the Lagrange multipliers must satisfy the fixed-point relation

$$\nu_k^{-1} = \mathbf{h}_{e_k}^H \left(\mathbf{I}_{N_t} + \sum_j \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H \right)^{-1} \mathbf{h}_{e_k} \left(1 + \frac{1}{\gamma_k} \right). \quad (9)$$

Since (8) can be solved using a power method, the complexity of finding the directions is dominated by the matrix inversion in (9), which requires $\mathcal{O}(N_t^3)$ operations. Having found those directions, the offset maximization power loading and the optimal offset can be found by solving the $K + 1$ linear equations that arise when the constraints in (6b) and (6c) hold with equality.

C. Robust power loading

The offset maximization algorithm described above uses the same offset r to increase the robustness of each user to channel uncertainty. The goal of robust power loading approach in [18] is to provide a computationally-efficient way to adapt the offset to the characteristics of each user's channel. For an arbitrary set of beamforming directions $\{\mathbf{u}_k\}$, the generic power loading problem can be stated as

$$\min_{\beta_k, \delta_k} \max_k \delta_k \quad (10a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \beta_k \leq P_t, \quad (10b)$$

$$\text{Prob}(\text{SINR}_k \geq \gamma_k) \geq 1 - \delta_k, \quad \forall k. \quad (10c)$$

The derivation of the algorithm developed in [18] for producing good solutions to (10) begins by observing the under the additive uncertainty model in (4), the probability that $\text{SINR}_k \geq \gamma_k$ is equal to the probability that

$$f_k(\mathbf{e}_k) = \mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} + 2\text{Re}(\mathbf{e}_k^H \mathbf{Q}_k \mathbf{h}_{e_k}) + \mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k - \sigma_k^2 \geq 0. \quad (11)$$

If we assume that the norms of the errors \mathbf{e}_k are small, as they will need to be for reliable operation [25], then we can approximate the quadratic term $\mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k$ by a Gaussian random variable of the same mean and variance. In that case, the distribution of $f_k(\mathbf{e}_k)$ becomes Gaussian. (Recall that we are focusing on the case where \mathbf{e}_k is Gaussian, with zero mean and of covariance $\sigma_{e_k}^2 \mathbf{I}$; cf. (4).) Under that approximation, if we design the power loading so that the mean, μ_{f_k} , of $f_k(\mathbf{e}_k)$ is a significant multiple of its standard deviation, σ_{f_k} , then that user will achieve a low outage probability. Indeed, we can choose a value for that multiple so that the target outage probability is guaranteed to be satisfied; see, e.g., [7]. We also note that the optimal solution of (10) has equal values for δ_k . If that were not the case, the user(s) with higher outage probability could be allocated more power and the other user(s) less, which would reduce the objective value, and thus contradict the assumed optimality. Therefore, it is natural to choose the same multiple, r , for each user in the approximation of the outage constraint in (10c). The resulting approximation of the problem in (10) can be written as [18]

$$\max_{\beta_k, r} r \quad (12a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \beta_k \leq P_t, \quad (12b)$$

$$\mu_{f_k} \geq r \sigma_{f_k}, \quad \forall k. \quad (12c)$$

From the definition of $f_k(\mathbf{e}_k)$ in (11) and the channel uncertainty model in (4), it can be shown that

$$\mu_{f_k} = \mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 + \sigma_e^2 \beta_k (1/\gamma_k + 1) - \sigma_e^2 P_t, \quad (13)$$

which is linear in the design variables $\{\beta_k\}_{k=1}^K$. (Recall from (3) that $\mathbf{Q}_k = \beta_k \mathbf{u}_k \mathbf{u}_k^H / \gamma_k - \sum_{j \neq k} \beta_j \mathbf{u}_j \mathbf{u}_j^H$.) Similarly, we have that

$$\sigma_{f_k}^2 = \text{var}\{f_k(\mathbf{e}_k)\} = 2\sigma_e^2 \mathbf{h}_{e_k}^H \mathbf{Q}_k^2 \mathbf{h}_{e_k} + \sigma_e^4 \text{tr}(\mathbf{Q}_k^2). \quad (14)$$

The structure of the problem in (12) is such that the constraints hold with equality at optimality [18]. Since σ_{f_k} is not a linear function in β , that results in a set of non-linear equations for the power loading. The following iterative linearization technique has been shown in [18] to be an effective way to obtain good solutions to (12):

- 1) Initialize each $\sigma_{f_k} = 1$.
- 2) Find $\{\beta_k\}$ and r by solving the set of linear equations that arise from equality in (12b) and (12c) for the current values of σ_{f_k} , where μ_{f_k} is defined in (13).
- 3) Update each σ_{f_k} using (14).
- 4) Return to (2) until a convergence criterion is satisfied.

We note that the matrix that relates $\{\beta_k\}$ to σ_{f_k} and r in step 2 is constant, and, accordingly, we need only invert this matrix once [18]. In practice, this algorithm converges quickly with a high probability [18]. In [26], the performance of this algorithm was shown to provide very similar performance to the optimal power loading in [17], and at a cost that is dominated by the $\mathcal{O}(K^3)$ operations that result from the initial matrix inversion.

III. OFFSET MAXIMIZATION DESIGNS WITH PAPCS

To simplify the development of the proposed robust beamforming technique, we will first consider the addition of PAPCs to the offset maximization problem in (6). We will then modify the resulting algorithm using insights from the above robust power loading algorithm.

When we add the PAPCs to the offset maximization problem in (6), the design problem becomes

$$r_{tpa}^* = \max_{\mathbf{w}_k, r} r \quad (15a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \leq P_t, \quad (15b)$$

$$\left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} \leq p_i, \quad \forall i, \quad (15c)$$

$$\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 - r \geq 0, \quad \forall k. \quad (15d)$$

Although the formulation in (15) is not convex, it can be transformed in a straightforward way into a second order cone program, using the technique that was used for the case of perfect CSI; cf. [4], [23]. While that formulation can be solved using a generic interior point method (e.g., [27]), such generic methods do not exploit the structure of the problem, and the development of tailored algorithms that do exploit the structure offers the potential for improved computational efficiency.

In the following subsections, we will first develop a low-complexity algorithm for the case where we have PAPCs only, with no total power constraint. Then we will tackle the general problem with both types of power constraints. The development will use insights from algorithms developed for the perfect CSI case [23] and insights from the robust power loading algorithm described in Section II-C.

A. Dominant PAPCs

If $P_t > \sum_{i=1}^{N_t} p_i$, the total power constraint can never be active and the problem in (15) can be rewritten as

$$r_{pa}^* = \max_{\mathbf{w}_k, r} r \quad (16a)$$

$$\text{s.t.} \quad \left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} \leq p_i, \quad \forall i, \quad (16b)$$

$$\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 - r \geq 0, \quad \forall k. \quad (16c)$$

Motivated by the way that a customized algorithm for (7) was adapted [13] to solve the problem in (6), we consider the following problem in which, for now, r_{pa}^* is presumed to be known,

$$\min_{\mathbf{w}_k, \alpha} \alpha \sum_{i=1}^{N_t} p_i \quad (17a)$$

$$\text{s.t.} \quad \left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} \leq \alpha p_i, \quad \forall i, \quad (17b)$$

$$\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 - r_{pa}^* \geq 0, \quad \forall k. \quad (17c)$$

In the context of (17), the constant term $\sum_{i=1}^{N_t} p_i$ in the objective is superfluous, but it will simplify the interpretation of the Lagrangian. Using arguments analogous to those in [13], [23], it can be shown that any set of beamformers that is optimal for (17) is also optimal for (16), and the optimal value of α in (17) is one.

Now, let q_i denote the dual variable of the i th condition in (17b) and ν_k denote the dual variable of the k th condition in (17c). Let us also define the diagonal matrix $\hat{\mathbf{Q}}$, such that $[\hat{\mathbf{Q}}]_{i,i} = q_i$. These definitions enable us to write the Lagrangian of the problem in (17) as

$$\begin{aligned} \mathcal{L}(\mathbf{w}_k, \alpha, \nu_k, q_i) = & \sum_{k=1}^K \nu_k (\sigma_k^2 + r_{pa}^*) + \alpha \left(\sum_{i=1}^{N_t} p_i - \sum_{i=1}^{N_t} q_i p_i \right) \\ & + \sum_{k=1}^K \mathbf{w}_k^H \left(\hat{\mathbf{Q}} + \sum_{j \neq k} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H - \nu_k / \gamma_k \mathbf{h}_{e_k} \mathbf{h}_{e_k}^H \right) \mathbf{w}_k. \end{aligned} \quad (18)$$

Using the notion of complementary slackness, since the optimal value of α is one, at optimality we have that $\sum_{i=1}^{N_t} p_i - \sum_{i=1}^{N_t} q_i p_i = 0$. Also, at optimality we have $\hat{\mathbf{Q}} + \sum_{j \neq k} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H - \nu_k / \gamma_k \mathbf{h}_{e_k} \mathbf{h}_{e_k}^H \succeq \mathbf{0}$, with \mathbf{w}_k lying in the null space of this matrix. This can be simplified to show that \mathbf{w}_k and

$$\hat{\mathbf{w}}_k = \left(\hat{\mathbf{Q}} + \sum_{k} \nu_k \mathbf{h}_{e_k} \mathbf{h}_{e_k}^H \right)^\dagger \mathbf{h}_{e_k}, \quad (19)$$

where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse, should be in the same direction. Further simplifications show that the dual variable ν_k in (19) should satisfy the fixed point equation

$$\nu_k^{-1} = \mathbf{h}_{e_k}^H \left(\hat{\mathbf{Q}} + \sum_j \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H \right)^\dagger \mathbf{h}_{e_k} \left(1 + \frac{1}{\gamma_k} \right). \quad (20)$$

From (20) we observe that if we were given the optimal $\hat{\mathbf{Q}}$, we could find the optimal values for $\{\nu_k\}$ using (20) and then the optimal directions $\{\mathbf{u}_k\}$ by normalizing the $\{\hat{\mathbf{w}}_k\}$ obtained using (19). After doing so, we could complete the solution of (17) by finding the optimal values for $\beta_k = \|\mathbf{w}_k\|^2$. That can be done by solving the set of K linear equations that

arise from the fact that at optimality (17c) holds with equality. (If this were not the case for condition k in (17c), then the amplitude of \mathbf{w}_k could be decreased which would allow a smaller value of α while satisfy all the other constraints.) To adapt that approach to solve (16), in the final step we must simultaneously solve for $\{\beta_k\}$ and r_{pa}^* . To do so we observe that r_{pa}^* enters linearly into (17c), and hence all we need is one more linearly independent equation. To obtain that equation we observe that if $q_i > 0$, then the i th component of (17b) holds with equality. By summing over all the active constraints in (17b) we obtain the following equation

$$\sum_{i, \forall q_i \neq 0} \left[\sum_{k=1}^K \beta_k \mathbf{u}_k \mathbf{u}_k^H \right]_{i,i} = \sum_{i, \forall q_i \neq 0} p_i. \quad (21)$$

In the case that all the q_i are positive — a case that happens quite often — the equation in (21) simplifies to $\sum_{k=1}^K \beta_k = \sum_{i=1}^{N_t} p_i$.

To complete the algorithm, we need to develop a technique to determine the optimal $\hat{\mathbf{Q}}$. One strategy for doing so is to apply the projected subgradient technique developed in [23]. That involves applying the update equation $\hat{\mathbf{Q}}^{n+1} = \text{proj}(\hat{\mathbf{Q}}^n + t_n \text{diag}(\text{diag}(\sum_i \mathbf{w}_i \mathbf{w}_i^H)))$, where $\text{proj}(\cdot)$ denotes the projection of a matrix on the space of diagonal positive semidefinite matrices that satisfy $\sum_{i=1}^{N_t} q_i p_i = \sum_{i=1}^{N_t} p_i$ and, consistent with the syntax used in MATLAB, when $\text{diag}(\cdot)$ operates on a matrix it produces a vector containing the diagonal elements and when it operates on a vector it produces a diagonal matrix with the elements of the vector on the diagonal. The initialization parameters used in [23] were chosen to be $\hat{\mathbf{Q}}^0 = \mathbf{I}$ and the step size chosen to be $t_n = 1/n$. Although this strategy converges, it can be quite slow [23]. In this paper, we will refine the approach in two ways. First, in Appendix A we develop a computationally cheap quasi-closed-form expression for the projection of $\hat{\mathbf{Q}}^{n+1}$ in a 2-norm sense. Second, based on insights from [28] we will choose a step size of the form $t_n = t_{n-1} - t_{n-1}^2/a$, for some positive scalar a . In addition, in Section III-C we will identify a prediction step that can be used in the first iteration to accelerate the algorithm. One simple termination strategy is to stop the algorithm when $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i < \epsilon_i, \forall i$, where ϵ_i is the maximum allowable violation of the power constraint for the i th antenna. Following the above development, the algorithm can be summarized as shown in Algorithm 1.

Algorithm 1 Offset maximization with PAPCs

- 1: Initialize the diagonal matrix $\hat{\mathbf{Q}}^0$ such that each element is non-negative and $\sum_{i=1}^{N_t} q_i p_i = \sum_{i=1}^{N_t} p_i$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Find $\{\nu_k\}$ using (20).
 - 4: Solve for the directions $\{\mathbf{u}_k\}$ by normalizing the $\{\hat{\mathbf{w}}_k\}$ obtained using (19).
 - 5: Find the power loading $\{\beta_k\}$ and r_{pa}^* by solving the set of linear equations arising from (17c) holding with equality and (21).
 - 6: Update $\hat{\mathbf{Q}}^{n+1}$ using the results in Appendix A.
 - 7: Increment n .
 - 8: **end while**
-

Having developed an efficient algorithm for the offset maximization problem with PAPCs, we now seek to incorporate the principles of the robust power loading discussed in Section II-C. To do so, we note that in the offset maximization design, the directions are independent of the offset term r in (16c); cf. (19) and (20). That suggests that we could simply modify the power loading step. Indeed, once the directions have been obtained in step 4 of Algorithm 1, we can replace the power loading in step 5 by the $\{\beta_k\}$ and r^* that solve (12). Those values can be found using the algorithm in Section II-C; see [18]. Incorporating that robust power loading algorithm into the framework of Algorithm 1 results in Algorithm 2.

Algorithm 2 PAPCed offset maximization with robust power loading

- 1: Initialize the diagonal matrix $\hat{\mathbf{Q}}^0$ such that each element is non-negative and $\sum_{i=1}^{N_t} q_i p_i = \sum_{i=1}^{N_t} p_i$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Find $\{\nu_k\}$ using (20).
 - 4: Solve for the directions $\{\mathbf{u}_k\}$ by normalizing $\{\hat{\mathbf{w}}_k\}$ obtained using (19).
 - 5: Find $\{\beta_k\}$ and r^* by solving $\mathbb{E}(\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2) = \sigma_{s_k} r^*$ and (21) using the method provided in Section II-C.
 - 6: Update $\hat{\mathbf{Q}}^{n+1}$ using the results in Appendix A.
 - 7: Increment n .
 - 8: **end while**
-

B. Total and PAPCed algorithm

Using the principles outlined in Section II-B and the previous section, we can develop an algorithm for solving the general problem in (15), which has PAPCs and a total power constraint. In this section, we will focus on the case when P_t is sufficiently smaller than $\sum_i p_i$ to ensure that the total power constraint is active. (Otherwise, the problem can be solved by the techniques in the previous section.) Similar to the previous section, we will obtain the beamforming directions by normalizing the beamformers resulting from the following problem

$$\min_{\mathbf{w}_k} \sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \quad (22a)$$

$$\text{s.t.} \quad \left[\sum_{j=1}^K \mathbf{w}_j \mathbf{w}_j^H \right]_{i,i} \leq p_i, \quad \forall i \quad (22b)$$

$$\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 - r_{tpa}^* \geq 0, \quad \forall k, \quad (22c)$$

and then we will refine the power loading using the method described in Section II-C. As in the previous development, the Lagrangian of (22) plays a key role. It can be written as

$$\mathcal{L}(\mathbf{w}_k, \nu_k, q_i) = \sum_{k=1}^K \nu_k (\sigma_k^2 + r_{tpa}^*) - \sum_{i=1}^{N_t} q_i p_i + \sum_{k=1}^K \mathbf{w}_k^H \left(\mathbf{I} + \hat{\mathbf{Q}} + \sum_{j \neq k} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H - \nu_k / \gamma_k \mathbf{h}_{e_k} \mathbf{h}_{e_k}^H \right) \mathbf{w}_k. \quad (23)$$

Using the KKT conditions, for a given value for $\hat{\mathbf{Q}}$ we can compute the corresponding directions and then the robust power loading in Section II-C. Furthermore, the subgradient

used in the previous section remains a subgradient in this case. However, the structure of the KKT conditions is simpler in this case, which results in a more straightforward projection for the $\hat{\mathbf{Q}}$ matrix. Indeed, since the only constraint on q_i in this case is that it is non-negative, the update equation for $\hat{\mathbf{Q}}$ can be written as

$$\hat{\mathbf{Q}}^{n+1} = \max\left(\hat{\mathbf{Q}}^n + t_n \text{diag}\left(\text{diag}\left(\sum_k \mathbf{w}_k \mathbf{w}_k^H\right) - \mathbf{p}\right), \mathbf{0}\right), \quad (24)$$

where the maximum operator is defined element-wise, and \mathbf{p} is the vector whose i th element is p_i . Therefore, we can construct an algorithm that has a similar structure to that in Algorithm 2. Having said that, in the case of PAPCs only there is a strong likelihood that the PAPCs will be active at optimality, and hence it makes sense to initialize the algorithm with a positive definite matrix $\hat{\mathbf{Q}}^0$. In the general case, the PAPCs are less likely to be active at optimality, and hence we will initialize the algorithm with $\hat{\mathbf{Q}}^0 = \mathbf{0}$. The resulting algorithm is provided in Algorithm 3.

Algorithm 3 Generalized offset maximization

- 1: Initialize $\hat{\mathbf{Q}}^0 = \mathbf{0}$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Find ν_k using the fixed point equations

$$\nu_k^{-1} = \mathbf{h}_{e_k}^H \left(\mathbf{I} + \hat{\mathbf{Q}}^n + \sum_j \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H \right)^{-1} \mathbf{h}_{e_k} \left(1 + 1/\gamma_k \right).$$
 - 4: Solve for the directions $\mathbf{u}_k = \hat{\mathbf{w}}_k / \|\hat{\mathbf{w}}_k\|$, where

$$\hat{\mathbf{w}}_k = \left(\mathbf{I} + \hat{\mathbf{Q}}^n + \sum_j \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H \right)^{-1} \mathbf{h}_{e_k}.$$
 - 5: Find $\{\beta_k\}$ and r^* by solving $\mathbb{E}(\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2) = \sigma_{s_k} r^*$ and $\sum_k \beta_k = P_t$ using the method provided in Section II-C.
 - 6: Update $\hat{\mathbf{Q}}^{n+1}$ using (24).
 - 7: Increment n .
 - 8: **end while**
-

C. Algorithm acceleration

As will be apparent in the simulations in Section VI, the modified update in Appendix A and the improved step size selection result in a substantial reduction of the number of iterations required over the number required using the choices made in [23]. Furthermore, we have observed that $\hat{\mathbf{Q}}^1$ and the corresponding matrix $\hat{\mathbf{Q}}^n$ at the termination of the algorithm are typically closely related. If that relationship can be determined with reasonable accuracy, this observation suggests that a predictive step could be used to further reduce the number of iterations. As an example of what can be done, in Section VI we illustrate how replacing $\hat{\mathbf{Q}}^1$ with a simple affine prediction, $\hat{\mathbf{Q}}_p^1$, of the terminating matrix $\hat{\mathbf{Q}}^n$ results in substantial reduction in the number of iterations.

IV. CONVENTIONAL ZF BEAMFORMING WITH PER-ANTENNA POWER CONSTRAINTS

Even though the computational cost of each iteration of the PAPCed offset maximization beamforming algorithms in the previous section is dominated by terms that are only $\mathcal{O}(N_t^3)$, when the BS has a large number of antennas the

resulting computational load can still be substantial. The dominating components arise from determining the beamforming directions, and the fact that these directions are updated at each iteration. That suggests that we may be able to develop lower cost algorithms for systems with a large number of antennas if we could find a way to simplify the computation of the beamforming directions. In this section we will do that by developing variants of the nominal ZF directions, and we will integrate them with the robust power loading technique while ensuring that the required PAPCs are satisfied. In the following section we will develop analogous techniques based on variants of the MRT directions. For the ZF case, the beamforming directions are obtained using techniques developed in [29], but in the MRT case, the design of the beamforming directions appears to be new.

To develop PAPCed variants of the conventional ZF and MRT beamformers, we observe that in contrast to QoS-based designs, in which the SINR is controlled directly (e.g., (15d)), the conventional ZF and MRT designs focus on the desired signal power and interference components of the SINR separately. In particular, given that the SINR for user k is $\text{SINR}_k = \frac{\mathbf{h}_k^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_k}{\mathbf{h}_k^H (\sum_{j \neq k} \mathbf{w}_j \mathbf{w}_j^H) \mathbf{h}_k + \sigma_k^2}$, if we were to maximize the minimum nominal received signal power subject to a total power constraint (i.e., $\max_{\{\mathbf{w}_k\}} \min_k \mathbf{h}_{e_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{e_k}$ subject to $\sum_k \mathbf{w}_k^H \mathbf{w}_k \leq P_t$) we would obtain beamformers that are a particular power loading of the nominal MRT directions. If we were to add the nominal ZF constraints on the interference into that problem (i.e., $\mathbf{h}_{e_j}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{e_j} = 0, \forall k \neq j$), then we would obtain beamformers that are a particular power loading of the ZF directions [29]. Due to the structure of the total power constraint, in many simple beamforming problems the optimization of the beamforming directions decouples from the power loading. That is indeed the case for our formulation for MRT and ZF beamforming directions. As an example, if we were to maximize the minimum value of $\mathbf{h}_{e_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{e_k} / \|\mathbf{h}_{e_k}\|^2$, which is the power of the signal transmitted in the direction of user k , rather than the power received by that user, we would obtain a set of beamformers in the MRT or ZF directions, but with a different power loading.

When the total power constraint is replaced by PAPCs, the optimization of the beamforming directions becomes coupled with the power loading and hence the choice of the metric to optimize changes both the power loading and the directions. While our approach will work for either metric, and indeed for several others, we will focus on the second metric $\mathbf{h}_{e_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{e_k} / \|\mathbf{h}_{e_k}\|^2$. The rationale for this choice is that while the received signal power is suitable for the ZF problem in the perfect CSI case, where the ZF constraints will eliminate the interference [29], it can be quite sensitive to the interference incurred due to channel estimation errors. (This is illustrated in our simulation results in Section VI.) Accordingly, we define the normalized channel directions $\mathbf{h}_{n_k} = \mathbf{h}_{e_k} / \|\mathbf{h}_{e_k}\|$ and we formulate the following generic problem to obtain PAPCed versions of the conventional beam-

formers

$$\max_{\mathbf{w}_k, t} \quad t \quad (25a)$$

$$\text{s.t.} \quad \left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} \leq p_i, \quad \forall i \quad (25b)$$

$$\mathbf{h}_{n_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{n_k} \geq t \quad \forall k, \quad (25c)$$

$$\mathbf{h}_{n_j}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{n_j} \leq \varepsilon \quad \forall k \neq j. \quad (25d)$$

The value of ε determines whether the problem is of the ZF type, the MRT type, or a variant thereof. When ε is negligible compared to the noise power, the formulation describes a ZF-based approach, and when ε is of the order of the noise power this represents a regularized ZF-based approach; cf. [15]. When ε is sufficiently large, the constraints in (25d) become inactive, and accordingly the formulation describes an MRT-based approach.

One strategy for solving (25) is to employ a semidefinite relaxation [30]. As in related beamforming methods based on semidefinite relaxation (e.g., [2]), that approach involves the solution of a convex optimization problem for a set of matrices and a post-processing step that extracts good beamformers from these matrices. However, the computational cost of solving the convex optimization problem is even higher than that of the offset maximization algorithm, and that is only the cost of determining the beamforming directions. Accordingly, in the following sections we will present low-cost algorithms for robust beamforming with PAPCed variants of the ZF and MRT beamforming directions.

A. ZF beamforming with PAPCs only

When $\varepsilon = 0$, the problem in (25) involves finding the beamforming vectors that remove the interference at the receivers under the nominal channel conditions and satisfy the PAPCs. The essence of this problem was addressed in [29] using a re-parametrization technique. In particular, let us define the matrix \mathbf{H} as the matrix whose k th column is \mathbf{h}_{n_k} and the matrix $\tilde{\mathbf{U}}_{\text{ZF}} = \mathbf{H}(\mathbf{H}^H \mathbf{H})^{-1}$. The k th column of $\tilde{\mathbf{U}}_{\text{ZF}}$, denoted $\tilde{\mathbf{u}}_{\text{ZF}_k}$ is a zero-forcing direction for the k th user with a unit signal gain; i.e., $\tilde{\mathbf{u}}_{\text{ZF}_k}^H \mathbf{h}_{n_k} = 1$. If we let \mathbf{H}_\perp denote a matrix whose columns form a basis for the null space of \mathbf{H} , then the set of all ZF directions for the k th user is given by the k th column of $\tilde{\mathbf{U}}_{\text{ZF}} + \mathbf{H}_\perp \mathbf{M}$, for an arbitrary scaling matrix \mathbf{M} . Accordingly, the solution to the problem in (25) takes the form

$$\mathbf{w}_k = \sqrt{t}(\tilde{\mathbf{u}}_{\text{ZF}_k} + \mathbf{H}_\perp \mathbf{m}_k), \quad (26)$$

where \mathbf{m}_k is the k th column of matrix \mathbf{M} [29]. Note that the constraints in (25c) and (25d) (with $\varepsilon = 0$) are automatically satisfied by designing the precoding vectors \mathbf{w}_k in the form in (26). The conditions that remain to be met are the PAPCs, and that can be done by adjusting the scaling matrix \mathbf{M} . In [29], this problem was formulated as a convex quadratically-constrained program that can be efficiently solved

$$\min_{\mathbf{M}, \hat{p}} \quad \hat{p} \quad (27a)$$

$$\text{s.t.} \quad \|(\tilde{\mathbf{U}}_{\text{ZF}} + \mathbf{H}_\perp \mathbf{M})^H \tilde{\mathbf{e}}_i\|^2 \leq \hat{p}, \quad \forall i, \quad (27b)$$

where $\tilde{\mathbf{e}}_i$ is the i th column of the identity matrix. To complete the design, we choose the largest value for t such that $t\hat{p} \leq p_i, \forall i$, which means that the beamformers of the form in (26) satisfy the remaining constraints; i.e., those in (25b).

If we let the (k, k) th entry of the diagonal matrix $\hat{\mathbf{Q}}$ denote the dual variable of the k th PAPC in (27b), then the KKT conditions of the dual problem of (27) show that the scaling matrix should satisfy $\mathbf{M} = -(\mathbf{H}_\perp^H \hat{\mathbf{Q}} \mathbf{H}_\perp)^\dagger (\mathbf{H}_\perp^H \hat{\mathbf{Q}} \tilde{\mathbf{U}}_{\text{ZF}})$. Although such a relation does not allow for a closed-form solution, as we do not know $\hat{\mathbf{Q}}$, it does allow for the integration of the robust power loading method in [18], as an alternative to giving all the users the same nominal signal strength t . Furthermore, the explicit relation between $\hat{\mathbf{Q}}$ and \mathbf{M} allows us to use the sub-gradient algorithm for $\hat{\mathbf{Q}}$ and to calculate \mathbf{M} accordingly. The proposed algorithm is summarized as Algorithm 4.

Algorithm 4 ZF with PAPCs and robust power loading

- 1: Find \mathbf{H}_\perp and $\tilde{\mathbf{U}}_{\text{ZF}}$. Initialize $\hat{\mathbf{Q}}^0 = \mathbf{I}$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Compute $\mathbf{M} = -(\mathbf{H}_\perp^H \hat{\mathbf{Q}}^n \mathbf{H}_\perp)^\dagger (\mathbf{H}_\perp^H \hat{\mathbf{Q}}^n \tilde{\mathbf{U}}_{\text{ZF}})$.
 - 4: Find the beamformers directions $\{\mathbf{u}_k\}$ by normalizing $\tilde{\mathbf{u}}_{\text{ZF}_k} + \mathbf{H}_\perp \mathbf{m}_k$.
 - 5: Find $\{\beta_k\}$ and r^* by solving $\mathbb{E}(\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2) = \sigma_{s_k} r^*$ and (21) using the method provided in Section II-C.
 - 6: Update $\hat{\mathbf{Q}}^{n+1}$ using the results in Appendix A.
 - 7: Increment n .
 - 8: **end while**
-

From a computational perspective, the key steps in the initialization of this algorithm are the finding of the ZF directions and the null space of \mathbf{H} , which requires $\mathcal{O}(N_t^2 K)$ operations. Each iteration of the algorithm involves the iterative solution of the $K + 1$ linear equations in step 5, which, as explained in Section II-C, requires $\mathcal{O}(K^3)$ operations, and the matrix operations required to update \mathbf{M} in step 3, which require $\mathcal{O}((N_t - K)^3)$ operations. When the number of antennas N_t is close to the number of users K , the dimensions of the matrix $\mathbf{H}_\perp^H \hat{\mathbf{Q}} \mathbf{H}_\perp$ are small, which means that in that case the computational cost of this algorithm is dominated by finding the ZF directions and the null space in the initialization step.

B. Generalized ZF beamforming

The extension of the ZF design with PAPCs to accommodate a total power constraint is straightforward, and follows the same steps that were used in the generalized offset maximization problem; see Section III-B. The generalized ZF problem can be formulated by adding the total power constraint $\sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \leq P_t$ to the constraints in (25). Then we consider the equivalent power minimization problem, assuming, for now, that the optimal t is known

$$\min_{\mathbf{M}} \quad t \sum_i \|(\tilde{\mathbf{U}}_{\text{ZF}} + \mathbf{H}_\perp \mathbf{M}) \tilde{\mathbf{e}}_i\|^2 \quad (28a)$$

$$\text{s.t.} \quad t \|(\tilde{\mathbf{U}}_{\text{ZF}} + \mathbf{H}_\perp \mathbf{M})^H \tilde{\mathbf{e}}_i\|^2 \leq p_i, \quad \forall i. \quad (28b)$$

Consistent with our previous analysis, we will let $\hat{\mathbf{Q}}$ denote the diagonal matrix with the dual variables of the PAPCs on

its diagonal. From the KKT conditions we can then show that $\mathbf{M} = -(\mathbf{H}_\perp^H(\hat{\mathbf{Q}} + \mathbf{I}_{N_t})\mathbf{H}_\perp)^{-1}(\mathbf{H}_\perp^H(\hat{\mathbf{Q}} + \mathbf{I}_{N_t})\tilde{\mathbf{U}}_{\text{ZF}})$. Furthermore, as in the previous algorithm we replace the uniform power loading, t , with the robust power loading from [18]. The resulting modified version of Algorithm 4 is stated in Algorithm 5. As is apparent from Algorithm 5, the order of its computational cost is the same as that of Algorithm 4.

Algorithm 5 Generalized ZF

- 1: Find \mathbf{H}_\perp and $\tilde{\mathbf{U}}_{\text{ZF}}$. Initialize $\hat{\mathbf{Q}}^0 = \mathbf{0}$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Compute $\mathbf{M} = -(\mathbf{H}_\perp^H(\hat{\mathbf{Q}}^n + \mathbf{I}_{N_t})\mathbf{H}_\perp)^{-1}(\mathbf{H}_\perp^H(\hat{\mathbf{Q}}^n + \mathbf{I}_{N_t})\tilde{\mathbf{U}}_{\text{ZF}})$.
 - 4: Find the beamformers directions $\{\mathbf{u}_k\}$ by normalizing $\tilde{\mathbf{U}}_{\text{ZF}} + \mathbf{H}_\perp \mathbf{m}_k$.
 - 5: Find $\{\beta_k\}$ and r^* by solving $\mathbb{E}(\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2) = \sigma_{s_k} r^*$ and $\sum_k \beta_k = P_t$ using the method provided in Section II-C.
 - 6: Update $\hat{\mathbf{Q}}^{n+1}$ using (24).
 - 7: Increment n .
 - 8: **end while**
-

V. CONVENTIONAL MRT WITH PER-ANTENNA POWER CONSTRAINTS

As we have seen in the previous section, our approach to imposing PAPCs on the class of ZF beamformers can result in an algorithm of lower computational cost than that of offset maximization with PAPCs. However, any advantage is dependent on the size of the null space of the channel matrix. In settings with a large number of antennas and a small number of users, such as those arise in massive MIMO, the size of the null space can be quite large. In this section, we will show how the complexity can be further reduced by using an MRT-based approach rather than the ZF-based approach.

A. MRT with PAPCs

In the MRT case, the interference conditions $\mathbf{h}_{n_j}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{n_j} \leq \epsilon$ are omitted from the problem in (25), and the problem of finding nominal MRT-based beamformers that satisfy PAPCs can be written as

$$\max_{\mathbf{w}_k, t} \quad t \quad (29a)$$

$$\text{s.t.} \quad \left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} \leq p_i, \quad \forall i \quad (29b)$$

$$\mathbf{h}_{n_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{n_k} \geq t, \quad \forall k. \quad (29c)$$

Following a similar analysis to those performed earlier, if we let q_i denote the dual variable for the i th PAPC, define the diagonal matrix $\hat{\mathbf{Q}}$ such that $[\hat{\mathbf{Q}}]_{i,i} = q_i$, and define ν_k to be the dual variable for the k th condition in (29c), then the Lagrangian of the problem in (29) can be written as:

$$\begin{aligned} \mathcal{L}(t, \mathbf{w}_k, \nu_k, q_i) = & -t + \sum_{i=1}^{N_t} q_i \left(\left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} - p_i \right) \\ & - \sum_{k=1}^K \nu_k (\mathbf{h}_{n_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{n_k} - t). \end{aligned} \quad (30)$$

Accordingly, we can state the KKT conditions in a simplified form as $\sum_{k=1}^K \nu_k = 1$, $t = \sum_{i=1}^{N_t} q_i p_i$, and $\hat{\mathbf{Q}} \mathbf{w}_k = \nu_k \mathbf{h}_{n_k} \mathbf{h}_{n_k}^H \mathbf{w}_k$. The last condition can be re-written as $\mathbf{w}_k = \nu_k \mathbf{h}_{n_k}^H \mathbf{w}_k \hat{\mathbf{Q}}^{-1} \mathbf{h}_{n_k}$, which means that \mathbf{w}_k and $\hat{\mathbf{Q}}^{-1} \mathbf{h}_{n_k}$ have the same direction. We note that at optimality $[\hat{\mathbf{Q}}]_{i,i}$ is equal to the i th element of \mathbf{h}_{n_k} , scaled by $\nu_k \mathbf{h}_{n_k}^H \mathbf{w}_k$, then divided by the i th element of \mathbf{w}_k . This equation does not allow any optimal $[\hat{\mathbf{Q}}]_{i,i}$ to be zero except if the channel vector \mathbf{h}_{n_k} contains a zero, which, under most reasonable channel models, is a “zero-probability” event. Since each q_i is positive, the constraints in (29b) are all active, and accordingly $\sum_k \beta_k = \sum_i p_i$.

Similar to the analysis of the previous problems, if we know $\hat{\mathbf{Q}}$, then we can find the beamforming directions using $\hat{\mathbf{Q}}^{-1} \mathbf{h}_{n_k}$ and subsequently solve for the power loading using the linear equations that arise when (29c) holds with equality. The value of t in (29c) can be calculated using the KKT equation $t = \sum_{i=1}^{N_t} q_i p_i$. If the equations in (29c) were not satisfied with equality at optimality, we could rescale the beamforming vectors to get a larger value of t , which would contradict the assumed optimality. This observation is similar to the observation in the offset maximization section that enabled the use of the subgradient algorithm to find $\hat{\mathbf{Q}}$. Accordingly, we can suggest the iterative algorithm in Algorithm 6.

Algorithm 6 Nominal MRT with PAPCs

- 1: Initialize $\hat{\mathbf{Q}}^0 = \mathbf{I}$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Solve for the directions using $(\hat{\mathbf{Q}}^n)^{-1} \mathbf{h}_{n_k}$.
 - 4: Find the beamformer magnitudes $\{\beta_k\}$ and t using the linear equations that arise when the constraints in (29c) are satisfied with equality and $\sum_k \beta_k = \sum_i p_i$.
 - 5: Update $\hat{\mathbf{Q}}^{n+1}$ using Appendix A.
 - 6: Increment n .
 - 7: **end while**
-

Algorithm 6 provides an iterative way to find the values of $\hat{\mathbf{Q}}$, and, accordingly, the optimal precoding vectors. Its complexity per iteration is no more than linear in N_t . However, we will now develop a closed-form expression that approximates the optimal solution of Algorithm 6 when the PAPCs are the same; i.e., $p_i = p, \forall i$. This closed-form removes the need for any iterations, which allows for an algorithm that is suitable for massive MIMO settings. To develop the approximation, we first note that the PAPCs and the MRT constraints hold with equality at optimality. That means that at optimality

$$\begin{aligned} \left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{i,i} &= \sum_{k=1}^K \nu_k^2 |\mathbf{h}_{n_k}^H \mathbf{w}_k|^2 \left| [\hat{\mathbf{Q}}^{-1} \mathbf{h}_{n_k}]_i \right|^2 \\ &= \sum_{k=1}^K \nu_k^2 t \left| [\hat{\mathbf{Q}}^{-1} \mathbf{h}_{n_k}]_i \right|^2. \end{aligned} \quad (31)$$

Now let us define

$$g_i = \sum_{k=1}^K \nu_k^2 \left| [\mathbf{h}_{n_k}]_i \right|^2. \quad (32)$$

Using (31), we can also write $g_i = pq_i^2/t = pq_i^2/(\sum_{i=1}^{N_t} q_i p)$. Accordingly, we can calculate q_i from $\{g_i\}$ as $q_i = (\sum_j \sqrt{g_j})\sqrt{g_i}$. The objective of maximizing $t = \sum_{i=1}^{N_t} q_i p$ is, therefore, equivalent to maximizing $\sum_j \sqrt{g_j}$. Since the dual variables ν_k^2 enter (32) as weighting variables for the power gains of the components of \mathbf{h}_{n_k} , the optimal values of ν_k are influenced by the relative values of the elements of each set $\{[\mathbf{h}_{n_k}]_i\}_{i=1}^K$. When these elements have the same distribution, the optimal values of ν_k tend to get closer as the number of antennas grows. Since $\sum_k \nu_k = 1$, that suggests the approximation $\nu_k \approx 1/K$. Since the approximation only holds in the limit, there will be discrepancy between the actual power on the antennas and p , but as the number of antennas grows, that difference decreases. For a finite number of antennas, we may rescale the result so that the PAPCs are satisfied. That is done in steps 6 and 7 in Algorithm 7.

Algorithm 7 One-shot approximate nominal MRT with PAPCs

- 1: Approximate $\nu_k \approx 1/K$.
 - 2: Calculate $g_i = \sum_{k=1}^K \nu_k^2 |[\mathbf{h}_{n_k}]_i|^2$ and $t = \sum_j \sqrt{g_j}$.
 - 3: Calculate $\hat{\mathbf{Q}}$ using $q_i = (\sum_j \sqrt{g_j})\sqrt{g_i}$.
 - 4: Calculate the beamformer directions $\{\mathbf{u}_k\}$ by normalizing $\hat{\mathbf{Q}}^{-1}\mathbf{h}_{n_k}$.
 - 5: Find $\beta_k = t/\mathbf{h}_{n_k}^H \mathbf{u}_k \mathbf{u}_k^H \mathbf{h}_{n_k}$.
 - 6: Form the vector \mathbf{y} , such that $y_i = [\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H]_{i,i}$.
 - 7: Form the correction vector \mathbf{z} such that $z_i = \sqrt{p_i/y_i}$.
 - 8: Correct each beamformer vector by element-wise multiplying each \mathbf{w}_k by the correction vector \mathbf{z} .
-

Both of the algorithms for the nominal MRT-based approach (Algos 6 and 7) result in beamformers that satisfy the PAPCs. As we will see in the simulation section, the resulting beamformers provide similar outage performance even for relatively small number of antennas. However, both of the algorithms are based on nominal performance criteria and any robustness that is obtained arises only implicitly. To address that point, we observe that Algorithm 6 updates $\hat{\mathbf{Q}}$ iteratively using the sub-gradient algorithm, which allows for the incorporation of the robust power loading described in Section II-C. The ability to incorporate that power loading can significantly reduce the outage probability by allocating each user an appropriate amount of power rather than forcing the nominal signal power of different users to be the same value t . The resulting algorithm is stated in Algorithm 8. In scenarios in which it is reasonable to use the same value of t for all users, or when we can pre-define different weights for the value of t , Algorithm 7 can provide a closed-form solution that is close to the optimal one, without the need for any iterations.

The complexity of Algorithm 8 is dominated by operations that are linear in the number of antennas for each user. This means that the complexity per iteration is of the order of $\mathcal{O}(N_t K)$ operations. The robust power loading can be effectively approximated in the massive MIMO settings so that it requires only $\mathcal{O}(N_t K)$ operations, beside the $\mathcal{O}(K^3)$ operations for the initial matrix inversion [18].

Algorithm 8 Robust MRT with PAPCs

- 1: Initialize $\hat{\mathbf{Q}}^0 = \mathbf{I}$. Set $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Find the beamforming directions $\{\mathbf{u}_k\}$ using $\hat{\mathbf{Q}}^{-1}\mathbf{h}_{n_k}$.
 - 4: Find $\{\beta_k\}$ and r^* by solving $\mathbb{E}(\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2) = \sigma_{s_k} r^*$ and (21) using the method provided in Section II-C.
 - 5: Update $\hat{\mathbf{Q}}^{n+1}$ using Appendix A.
 - 6: Increment n .
 - 7: **end while**
-

B. Generalized MRT

The derivation of the MRT-based algorithm when the total power constraint is added to (29) follows the same steps that were performed in the ZF case and the offset maximization case. The modified algorithm is presented in Algorithm 9.

Algorithm 9 Generalized MRT

- 1: Set $\hat{\mathbf{Q}}^0 = \mathbf{0}$, and $n = 0$.
 - 2: **while** $[\sum_k \mathbf{w}_k \mathbf{w}_k^H]_{i,i} - p_i > \epsilon_i$ for any i **do**
 - 3: Find the beamformers directions $\{\mathbf{u}_k\}$ by normalizing $(\mathbf{I}_{N_t} + \hat{\mathbf{Q}})^{-1}\mathbf{h}_{n_k}$.
 - 4: Find $\{\beta_k\}$ and r^* by solving $\mathbb{E}(\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2) = \sigma_{s_k} r^*$ and $\sum_k \beta_k = P_t$ using the method provided in Section II-C.
 - 5: Update $\hat{\mathbf{Q}}^{n+1}$ using (24).
 - 6: Increment n .
 - 7: **end while**
-

VI. SIMULATION RESULTS

In this section, we will show how the application of PAPCs to substantially reduce the dynamic range of the power transmitted from each antenna can be implemented without significantly degrading the outage probability of the system. We consider a system in which a BS with N_t antennas serves K single-antenna users distributed uniformly in a disk of radius 3.2km around the BS. The large scale fading is modelled using a path-loss exponent of 3.52 and log-normal shadow fading with 8dB standard deviation. The small scale fading is modelled using the standard i.i.d. Rayleigh model. We assume an additive channel estimation error of covariance $0.04\mathbf{I}$, and an SINR target of $\gamma = 3\text{dB}$ for all users. For the algorithms with PAPCs only, the PAPC is uniform and is set to $p_i = P_t/N_t$, where P_t is the total power constraint, which is implicit in this case. For the generalized algorithms with both PAPCs and a total power constraint (Algos 3, 5, and 9), the PAPCs are set to be slightly larger, so that the total power constraint is active. For these cases we choose $p_i = 1.2P_t/N_t$. We assume that each user has a signal sensitivity of -90dBm, and we will consider this power as the noise power. The termination parameter for the algorithms is chosen to be $\epsilon_i = 0.1p_i$, and each experiment is repeated on 20,000 channel realizations. A simple channel-strength user selection technique is employed, where users having $\|\mathbf{h}_{e_k}\|^2 P_t / k \sigma_k^2 \geq \gamma_k$ are served.

To demonstrate the application of PAPCs with offset maximization, in Fig. 1 we plot the outage probability versus the total power constraint P_t for six different algorithms in a scenario in which $N_t = 4$ and $K = 3$. The first algorithm is the nominal PAPCed design algorithm presented in [23], with the beamforming vectors scaled so that the total power is equal to P_t . This is equivalent to solving (17) when $r_{pa}^* = 0$, then scaling the resulting beamforming vectors. We compare the performance of [23] to the performance of Algo. 1 with and without the acceleration step, and Algo. 2 with the acceleration step. We note that while the performance of Algo. 1 is close to that of [23], the application of the robust power loading in Algo. 2 provides a significant reduction in the outage probability. To assess the impact of the PAPCs we compare the performance of Algo. 2 to that of the robust offset maximization technique with a total power constraint only [18]. As seen in Fig. 1, Algo. 2 achieves a performance close to that of [18] even though it imposes PAPCs. As expected, the performance of Algo. 3, which imposes a total power constraint and weaker PAPCs, falls in between that of [18] and Algo. 2.

The convergence rate of the subgradient algorithm strongly depends on how the step size is chosen and, hence, this should be tailored to the application. Based on insights from [28] we have chosen a step size that is updated using $t_n = t_{n-1} - t_{n-1}^2/1000$. Our numerical experience has suggested choosing $t_0 = N_t/(P_t K)$. To examine the potential impact of the prediction scheme outlined in Section III-C, we have implemented a linear predictor of the form $\hat{\mathbf{Q}}_p^1 = 2.8 \text{diag}(\mathbf{q}^1) - 1.8\mathbf{I}$. To show the effectiveness of these choices, we plot in Fig. 2 the percentage of violated PAPCs versus the iteration number for the scenario in which $P_t = 40$. We set the violation to one when any antenna is transmitting a power that is more than 10% higher than p_i . (Recall that we set $\epsilon_i = 0.1p_i$.) We observe from Fig. 2 that within the first few iterations, the PAPCs are met in most cases. We also note that the acceleration step can reduce the average number of iterations while providing almost the same outage performance. In order to provide context for these results, we point out that the average number of iterations required by the nominal algorithm in [23] is much higher. Indeed, as shown in [23], it can range from a few tens to hundreds in analogous settings.

In assessing the performance of the ZF-based PAPCed beamforming algorithms, rather than examining the outage performance against the transmission power, we will fix the total power constraint to $P_t = 2$ and examine the performance as the number of antennas, N_t , increases. Other than that, the scenario is the same as the previous one. As performance benchmarks for Algo. 4, we have included the performance of the algorithm in [29] which maximizes the minimum received signal power, $\mathbf{h}_{e_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{e_k}$, and a modified version of the algorithm in [29] that maximizes the minimum value of $\mathbf{h}_{n_k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{n_k}$ instead. We observe that in the case of noisy channel estimates, the normalization step significantly reduces the outage probability. More importantly, the application of the robust power loading in Algo. 4 provides significantly better performance. As a lower bound on the outage achieved

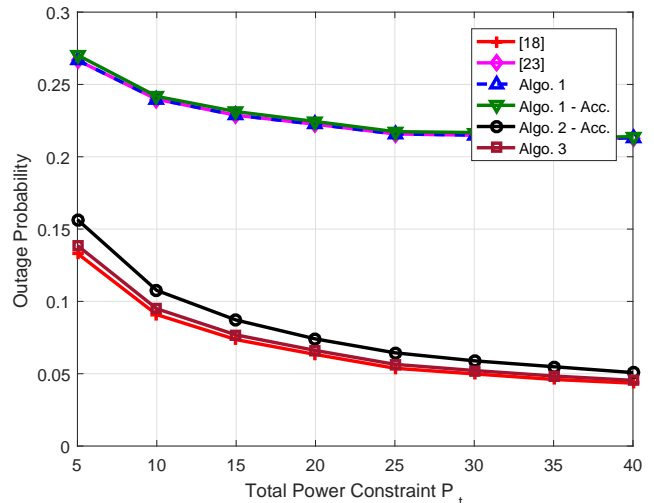


Fig. 1. Outage probability for a 4 antenna BS serving 3 users with a total transmitted power of P_t .

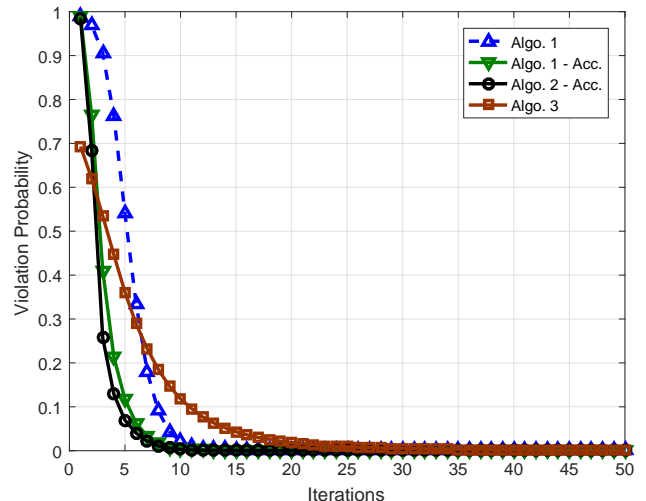


Fig. 2. Convergence behaviour for a 4 antenna BS serving 3 users with a total transmitted power of P_t . The violation probability measures the fraction of the 20,000 realizations for which at least one PAPC was violated by more than 10% at the given iteration of the algorithm.

by Algo. 4 we consider ZF beamforming with the nominal ZF directions and robust power loading with only a total power constraint [18]; i.e., without the PAPCs. The resulting comparison shows that the degradation incurred by imposing the PAPCs is quite small. Finally, as expected, the performance of the generalized algorithm (Algo. 5) lies in between that of Algo. 4 and that of [18].

To assess the performance of the MRT-based PAPCed algorithms, we will allow for more users, $K = 8$, and set the total power constraint P_t to be 1. As in the ZF case, we examine the outage performance versus the total number of antennas, N_t , but we do so for a larger number of antennas. In Fig. 4, the performance of Algos 6, 7, and 8 is compared to the performance of the algorithm in [24]. We observe that the

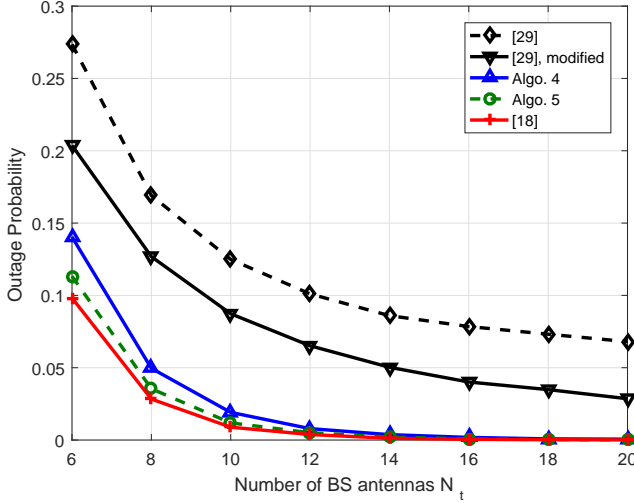


Fig. 3. Outage probability for a BS serving 3 users with a total transmitted power of $P_t = 2$.

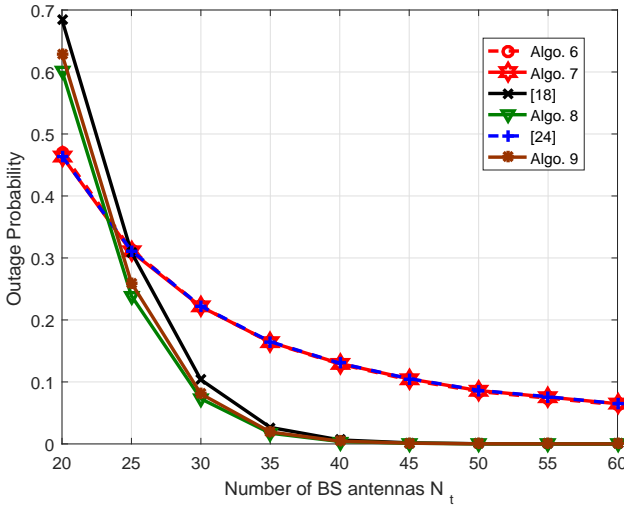


Fig. 4. Outage probability for a BS serving 8 users with a total transmitted power of $P_t = 1$.

performance of Algos 6, and 7 is almost identical to that of the algorithm presented in [24], and that the performance of Algo. 8 is superior. As a benchmark, the performance of the robust MRT beamformer with only a total power constraint (i.e., no PAPCs) [18] is plotted in Fig. 4. The performance of the generalized algorithm (Algo. 9) is also plotted.

VII. CONCLUSION

In this paper, we developed low-complexity algorithms for finding robust beamformers that provide low outage of target SINRs while satisfying specified per-antenna power constraints (PAPCs). Initially, we used insights from the subgradient method for designing PAPCed beamformers in the case of perfect channel state information [23] to obtain PAPCed version of the offset maximization algorithm devel-

oped in [13]. Further reductions in the outage probability were then obtained by incorporating the robust power loading presented in [18] into the design problem. While the proposed algorithms are of low complexity, we identified the evaluation of the beamforming directions as the computational bottleneck. To address that, we developed algorithms that employ PAPCed variants of the conventional zero-forcing (ZF) and maximum ratio transmission (MRT) directions and incorporate the robust power loading. In the process of doing so, we developed a closed-form expression for an MRT-based beamformer that satisfies PAPCs and may be appropriate for massive MIMO systems. Our simulation results revealed that PAPCed beamforming can be achieved without incurring a significant degradation in outage performance.

APPENDIX A Q UPDATE

To determine the updated value for $\hat{\mathbf{Q}}^{n+1}$, we have to determine the projection,

$$\hat{\mathbf{Q}}^{n+1} = \text{proj}(\hat{\mathbf{Q}}^n + t_n \text{diag}(\text{diag}(\sum_i \mathbf{w}_i \mathbf{w}_i^H))).$$

To do so, we let $\mathbf{q} = \text{diag}(\hat{\mathbf{Q}}^{n+1})$, and $\mathbf{q}_o = \text{diag}(\hat{\mathbf{Q}}^n + t_n \text{diag}(\text{diag}(\sum_i \mathbf{w}_i \mathbf{w}_i^H)))$. That enables us to write the projection problem as

$$\min_{\mathbf{q}} \|\mathbf{q} - \mathbf{q}_o\|^2 \quad (33a)$$

$$\text{s.t.} \quad \sum_{i=1}^{N_t} q_i p_i = \sum_{i=1}^{N_t} p_i \quad (33b)$$

$$q_i \geq 0, \quad \forall i. \quad (33c)$$

If we let ζ denote the dual variable of the equality constraint, then from the KKT conditions of (33) we can show that the optimal q_i is

$$q_i = \max(q_{o_i} - p_i \zeta / 2, 0),$$

where $\zeta/2 = (\sum_{i, q_i > 0} p_i q_{o_i} - \sum_i p_i) / \sum_{i, q_i > 0} p_i^2$. Given the nature of dependence of $\{q_i\}$ and ζ on each other, we will solve for their values using a fixed-point approach. First, we initialize $\zeta = 0$, and then we iteratively calculate q_i and ζ from the provided equations until their values stabilize.

In the case of equal p_i (i.e., $p_i = p, \forall i$), and when all the q_i are positive (i.e., all the PAPCs are active), the update equation can be simplified to

$$\hat{\mathbf{Q}}^{n+1} = \hat{\mathbf{Q}}^n + t_n \text{diag}(\text{diag}(\sum_i \mathbf{w}_i \mathbf{w}_i^H - p \mathbf{I})). \quad (34)$$

REFERENCES

- [1] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Proc. Mag.*, vol. 24, no. 5, pp. 36–46, Sept. 2007.
- [2] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. CRC Press, 2001, ch. 18.
- [3] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1313–1324, Oct. 1998.
- [4] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [5] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2324–2334, July 2005.

- [6] E. Bjornson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, July 2014.
- [7] A. Pascual-Iserte, D. Palomar, A. Perez-Neira, and M. Lagunas, "A robust maxmin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 45, pp. 346–360, Jan. 2006.
- [8] M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 714–724, Dec. 2007.
- [9] G. Zheng, K.-K. Wong, and T.-S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 154:1–154:15, Jan. 2008.
- [10] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 714–725, Feb. 2009.
- [11] M. B. Shenouda and T. N. Davidson, "Probabilistically-constrained approaches to the design of the multiple antenna downlink," in *Conf. Rec. 42nd Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Oct. 2008, pp. 1120–1124.
- [12] K.-Y. Wang, A.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
- [13] M. Medra, Y. Huang, W. K. Ma, and T. N. Davidson, "Low-complexity robust MISO downlink precoder design under imperfect CSI," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3237–3249, June 2016.
- [14] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.
- [15] 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.
- [16] T. Y. Al-Naffouri, M. Moinuddin, N. Ajeeb, B. Hassibi, and A. L. Moustakas, "On the distribution of indefinite quadratic forms in Gaussian random variables," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 153–165, Jan. 2016.
- [17] F. Sotiriou and T. N. Davidson, "Coordinate update algorithms for robust power loading for the MU-MISO downlink with outage constraints," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2761–2773, June 2016.
- [18] M. Medra and T. N. Davidson, "Per-user outage-constrained power loading technique for robust MISO downlink," in *Conf. Rec. 49th Asilomar Conf. Signals, Systems, Computers*, Nov. 2015, pp. 1232–1236.
- [19] S. Shen and T.-M. Lok, "Downlink beamforming and power control with per-antenna power constraints," in *Proc. IEEE Wireless Commun. Networking Conf.*, April 2014, pp. 1206–1211.
- [20] G. Dartmann, X. Gong, W. Afzal, and G. Ascheid, "On the duality of the maxmin beamforming problem with per-antenna and per-antenna-array power constraints," *IEEE Trans. Veh. Tech.*, vol. 62, no. 2, pp. 606–619, Feb. 2013.
- [21] T. Liu and J. Chen, "Joint downlink transmit and receive beamforming under per-antenna power constraints," in *Proc. 7th Int. Conf. Inf. Commun. Signal Proc.*, Dec. 2009, pp. 1–5.
- [22] Y. K. Yazarel and D. Aktas, "Downlink beamforming under individual SINR and per antenna power constraints," in *Proc. IEEE Pac. Rim Conf. Commun. Comput. Signal Process.*, Aug. 2007, pp. 422–425.
- [23] 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, June 2007.
- [24] C. Feng and Y. Jing, "Modified MRT and outage probability analysis for massive MIMO downlink under per-antenna power constraint," in *Proc. IEEE Int. Wkshp Signal Process. Adv. Wireless Commun.*, July 2016, pp. 1–6.
- [25] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [26] M. Medra and T. N. Davidson, "Robust MISO downlink: An efficient algorithm for improved beamforming directions," in *Proc. IEEE Signal Process. Wkshp Sensor Array and Multichannel Signal Process.*, July 2016, pp. 1–5.
- [27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [28] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J. S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.
- [29] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409–4418, Sept 2008.
- [30] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.