Robust MISO Downlink Precoder Design With Per-Antenna Power Constraints

Mostafa Medra Timothy N. Davidson
Department of Electrical and Computer Engineering
McMaster University, Hamilton, Ontario, Canada

Abstract—We consider the design of a linear precoder 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 signal-to-interference-and-noise ratio (SINR) targets specified by the receivers, while satisfying constraints on the power transmitted by each antenna and a total transmitted power constraint. The proposed algorithm exploits a recently developed broadly applicable approximation of the outage probability. That approximation results in a problem formulation that can be efficiently solved by adapting existing low-computational-cost algorithms for scenarios with perfect CSI and per-antenna power constraints. Simulation results show that robust designs with a uniform and near-uniform power distributions across the transmitting antennas can be obtained without a substantial degradation in the outage probability.

I. INTRODUCTION

The spatial multiplexing capabilities of base stations (BSs) with multiple antennas offer the potential for substantial gains in the quality of service that can be afforded to users in a downlink system; e.g., [1]. In particular, for users with inelastic traffic, linear beamforming schemes have been developed to simultaneously serve multiple users at their requested rates [2]–[4]. 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; e.g., [5]–[12].

The approaches in [5]–[10] involve approximations that result in designs that are formulated as convex optimization problems, but the computational cost of solving those problems is quite significant. Furthermore, the approximations of the outage constraint can be quite conservative, especially as the size of the uncertainty increases, and this can result in the approximated problem being infeasible even when the underlying problem has a solution. To address those issues, a different approach to approximating the outage constraint was developed in [13]. Although that approximation does not guarantee that the outage constraint holds, it performs well in the presence of significant uncertainties, and result in a computationally inexpensive algorithm that is almost always feasible.

The existing literature on robust downlink beamforming has focused 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 have been developed [14]–[18]. For robust beamforming designs that are formulated as second-order cone or semidefinite programming problems (e.g., [6], [7], [10]) and 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 a robust beamforming design that incorporates per-antenna power constraints and has a reasonable computational cost. We do so by showing how the iterative algorithms in [18] for the perfect CSI case with per-antenna power constraints can be adapted to solve the problem that is obtained when the per-antenna power constraints are incorporated into the robust downlink beamforming formulation that was developed in [13].

II. SYSTEM MODEL AND DESIGN APPROACH

We consider a narrowband multiple-input single-output (MISO) downlink with an $N_t$ antenna BS and $K$ single antenna users. The transmitted signal $\mathbf{x}$ is constructed using linear beamforming as $\mathbf{x} = \sum_{i=1}^{K} \mathbf{w}_i s_i$, where $s_i$ is the normalized data symbol for user $i$, and $\mathbf{w}_i$ is the associated beamformer. The received signal at user $i$ is

$$y_i = \mathbf{h}_i^H \mathbf{w}_i s_i + \sum_{j \neq i} \mathbf{h}_i^H \mathbf{w}_j s_j + n_i,$$

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

In the problems that we will consider, each user specifies the signal-to-interference-and-noise ratio (SINR) that it will

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.
require in order to support the service that it desires. This constraint takes the form
\[
\text{SINR}_i = \frac{h_i^H w_i w_i^H h_i}{h_i^H (\sum_{j \neq i} w_j w_j^H ) h_i + \sigma_i^2} \geq \gamma_i, \tag{2}
\]
or equivalently
\[
h_i^H Q_i h_i - \sigma_i^2 \geq 0, \tag{3}
\]
where \(\sigma_i^2\) is the noise variance at receiver \(i\). 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 \(\sum_{i=1}^{N_t} E|x_i|^2 \leq P_t\). In case of zero mean independent data symbols of normalized power, this constraint becomes \(\sum_{i=1}^{K} w_i^H w_i \leq P_t\). If we let \(P_i\) denote the maximum power that could be transmitted using antenna \(i\), the per-antenna power constraint can be written as \(E|x_i|^2 \leq P_i\). Using the previous assumptions for the data symbols, this per-antenna power constraint can be written as \(\sum_{i=1}^{K} w_i^H w_i \leq P_i\).

As shown in [2], for the BS to be able to evaluate whether a candidate set of beamformers \(\{w_i\}_{i=1}^{K}\) satisfies the K-SINR constraints, the BS must know each channel vector \(h_i\). However, typically the BS will only have an estimate of that channel, \(h_i^\hat{}\). One way in which the BS can incorporate the resulting channel uncertainty into the design is to postulate a conditional distribution, \(p(h_i|h_i^\hat{})\), and seek to design the beamformers \(w_i\) so as to minimize the probability of outage of the SINR targets, subject to a total power constraint; i.e.,
\[
\min_{w_i, \delta_i} \max_{\epsilon_i} \delta_i \tag{4a}
\]
s.t. \(\sum_i ||w_i||^2 \leq P_t\), \(\text{Prob(SINR}_i \geq \gamma_i) \geq 1 - \delta_i\). \tag{4b}

where we will leave it implicit that (4c) must hold for all \(i \in \{1, 2, ..., K\}\). Since the probability constraint in (4c) is intricate and non-convex, many approaches to robust design replace this constraint by another constraint that imposes zero outage for a set of channel vectors having a probability of no less than \(1 - \delta_i\); e.g., [9], [10]. This assumption is conservative by nature, and does not consider the performance outside that "zero-outage region". Furthermore, the characterization of the zero-outage region often involves linear matrix inequalities. Although such constraints are convex, the computational cost of implementing them can be quite considerable.

These weaknesses of the zero-outage region approach were part of the motivation for the development of the computationally-efficient "offset maximization" approach that was developed in [13]. This approach maintains good performance within the zero-outage region and seeks to provide better performance outside that region. Given the set of channel estimates at the BS \(\{h_i\}\), the offset maximization approach seeks solutions to the following problem:
\[
r_i^* = \max_{w_i,r} \quad r \tag{5a}
\]
s.t. \(\sum_{i=1}^{K} w_i^H w_i \leq P_t\), \(h_i^H Q_i h_i - \sigma_i^2 - r \geq 0\). \tag{5c}

It is implicit in (5c) that this algorithm tries to find the largest noise plus interference power each user can endure, under a total power constraint. When \(r_i^* > 0\) the solution to (5) provides robustness against channel estimation errors. If \(r_i^* < 0\), then even in the case in which the set of channel estimates \(\{h_i\}\) is perfect, there is no encoder that can achieve the specified SINR targets with total power at most \(P_t\). In that case, \(r_i^*\) implicitly specifies the degradation in the SINRs that is required to reduce the transmitted power to \(P_t\).

In [13] an efficient method to solve (5) was developed by considering the following problem, in which, for now, it is assumed that the optimal value for (5), \(r_i^*\), is known:
\[
P^* = \min_{w_i} \quad \sum_i ||w_i||^2 \tag{6a}
\]
h_i^H Q_i h_i - \sigma_i^2 - r \geq 0. \tag{6b}

The optimal value of the problem in (6), \(P^*\), cannot be greater than \(P_t\) because the solution of (5) is feasible for (6) and has an objective value of \(P_t\). Furthermore, \(P^* \neq P_t\), because otherwise the solution of (5) would have an optimal value that is larger than \(r_i^*\). Therefore, \(P^* = P_t\) and any set of beamformers that optimize (6) is also optimal for (5).

The advantage of the connection between problems (5) and (6) is that a highly efficient iterative algorithm for the problem in (6) with \(r_i^* = 0\) (i.e., the perfect CSI case) was developed in [2]. That algorithm exploits the structure of the KKT conditions. The directions of the beamformers are obtained from an eigen equation, the matrix of which is updated via a fixed point equation. The \(K\) unknown amplitude squares \(\beta_i = ||w_i||^2\) are calculated from the \(K\) linear equations that are derived from the fact that at optimality all the constraints in (6b) hold with equality. To adapt that algorithm to solve the problem in (5) we note that we have one extra variable, \(r\), and we have also one extra constraint \(\sum_{i=1}^{K} \beta_i = P_t\). Since the variable \(r\) enters the SINR constraints linearly, and since the additional power constraint is linear in \(\beta_i\), the technique in [2] for the problem in (6) can be extended to solve the problem in (5) simply by replacing the \(K\) linear equations for finding the optimal values of \(\{\beta_i\}_{i=1}^{K}\) by the \(K + 1\) linear equations for finding the optimal values for \(\{\beta_i\}_{i=1}^{K}\) and \(r\).

III. PER-ANTENNA POWER CONSTRAINED DESIGNS

If we wish to obtain the robustness provided by the offset maximization algorithm, while satisfying constraints on the power transmitted by each antenna, as well as the total transmitted power, the design problem becomes
\[
r_{ipa}^* = \max_{w_i,r} \quad r \tag{7}
\]
s.t. \(\sum_{i=1}^{K} w_i^H w_i \leq P_t\), \(\sum_{i=1}^{K} w_i^H w_i \leq P_t, \tag{7a}
\]
h_i^H Q_i h_i - \sigma_i^2 - r \geq 0. \tag{7b}

Although the formulation in (7) is not convex, it can be transformed in a straightforward way into a second order cone program, as shown for the case of perfect CSI in [18]. That
form is convex and can be solved using the convenient “cvx” interface [19] to implementations of generic interior point methods. However, these generic algorithms 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 sections, we will show that in the case in which either the per-antenna power constraints dominate or at most one per-antenna constraint can be active, insights from algorithms developed for the perfect CSI case [18] can be adapted to develop highly efficient algorithms that exploit the underlying structure of the problem.

A. Dominant per-antenna power constraints

If \( P_c \geq \sum_{i=1}^{N_t} P_i \), the total power constraint can never be active and the problem in (7) can be rewritten as

\[
\begin{align*}
\max_{\mathbf{w}_i, r} & \quad r \\
\text{s.t.} & \quad \left[ \sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H \right]_{i,i} \leq P_i, \\
& \quad \mathbf{h}_i^H \mathbf{Q} \mathbf{h}_{e_i} - \sigma_i^2 - r \geq 0.
\end{align*}
\]

Using the notion of complementary slackness, since \( \alpha \) must be a positive number, at optimality we have that \( \sum_{i=1}^{N_t} P_i - \sum_{i=1}^{N_t} q_i P_i = 0 \). In this paper we will focus on the case of equal \( P_i \), and in that case this condition is equivalent to \( \sum_{i=1}^{N_t} q_i = N_t \). Also, at optimality we have \( \mathbf{Q} + \sum_{j \neq i} \lambda_j \mathbf{h}_i \mathbf{h}_j^H = \lambda_i / \gamma_i \mathbf{h}_i \mathbf{h}_i^H \geq 0 \) and \( \mathbf{w}_i \) lying in the null space of this matrix. This can be simplified to show that \( \mathbf{w}_i \)

\[
\mathbf{w}_i = \left( \mathbf{Q} + \sum_{j} \lambda_j \mathbf{h}_j \mathbf{h}_j^H \right)^{-1} \mathbf{h}_{e_i}
\]

should be in the same direction. Further simplifications show that the dual variable \( \lambda_i \) should satisfy the fixed point equation

\[
\lambda_i^{-1} = \mathbf{h}_i^H \left( \mathbf{Q} + \sum_{j} \lambda_j \mathbf{h}_j \mathbf{h}_j^H \right)^{-1} \mathbf{h}_{e_i} \left( 1 + \frac{1}{\gamma_i} \right).
\]

From (13) we observe that if we were given \( \mathbf{Q} \), we could find the optimal \( \{ \lambda_i \} \) using (13) and then the optimal directions for \( \{ \mathbf{w}_i \} \) using (12). After doing so, we could complete the solution of (9) by solving a set of linear equations for the optimal values of \( \beta_i = \| \mathbf{w}_i \|_2^2 \) and \( r_{pa} \). In order to adapt that approach to solve (8) we observe that \( r_{pa} \) enters linearly into (9c) and that at optimality conditions in (9c) must be satisfied with equality; if this is not the case for condition \( i \), then the amplitude of \( \mathbf{w}_i \) can be decreased which will allow a smaller value for \( \alpha \) and yet satisfy all the conditions. We also observe that if \( \mathbf{Q} = 0 \), then the constraints in (9b) hold with equality. By summing the constraints over the antennas we obtain the additional linear equation \( \sum_{i=1}^{K} \beta_i = \sum_{i=1}^{N_t} P_i \), and we can obtain \( \{ \beta_i \}_{i=1}^{K} \) and \( r_{pa} \) by solving these \( K + 1 \) linear equations. If \( \mathbf{Q} \) contains a zero diagonal element, then the corresponding constraint in (9b) might not hold with equality at optimality. As in [18], if that occurs the above approach will yield a feasible solution for (8), but that solution may be suboptimal. Fortunately, under typical operating conditions this is a rare event; cf. [18].

To complete the development of the algorithm, we need to develop a technique to determine \( \mathbf{Q} \). Using the subgradient projection method, it is shown in [18] that we can use the update equation \( \mathbf{Q}^{n+1} = \text{proj}(\mathbf{Q}^n + t_n \text{diag}(\sum_{i} \mathbf{w}_i \mathbf{w}_i^H)) \), where \( t_n \) is the step size and \( \text{proj}() \) denotes the projection of a matrix on the space of diagonal positive semidefinite matrices having a trace equal to \( N_t \). (Recall that we have focused on the case of equal \( P_i \).) If we denote the projected diagonal values of \( \mathbf{Q} \) by \( \gamma \) and the current values by \( q_i \), that projection can be computed by solving \( \min_{\mathbf{y}} \| \mathbf{y} - \mathbf{q} \|_2^2 \), s.t. \( \mathbf{y} \succeq 0 \) and \( \sum_i \mathbf{y}_i = N_t \). Using the dual of that problem and the KKT conditions, it can be shown that \( y_i = q_i - t_n P_i / N_t \) is optimal when it is non-negative. Accordingly, we can employ the following update equation instead of the original one

\[
\mathbf{Q}^{n+1} = \mathbf{Q}^n + t_n \text{diag}(\sum_{i} \mathbf{w}_i \mathbf{w}_i^H - P_i / N_t). \tag{14}
\]

Under this update, if \( \text{tr}(\mathbf{Q}^n) = N_t \) then \( \text{tr}(\mathbf{Q}^{n+1}) = N_t \), too. If an element of \( \mathbf{Q}^{n+1} \) is negative it is projected back to zero, and \( \mathbf{Q}^{n+1} \) is rescaled so that its trace equals \( N_t \).
Following [18], we will initialize $\hat{Q}$ as an identity matrix. The numerical results in [18] show a very slow convergence rate using $t_n = 1/n$, so we will use the step size update equation $t_n = t_{n-1} - r_{n-1}/1000$ [20], where $t_0 = 1/P$. The algorithm can be summarized as:

1) Initialize $\hat{Q}$ such that $\text{tr}(\hat{Q}^2) = N_t$. Set $n = 0$.
2) Find $\lambda_i$ using (13).
3) Solve for the directions $\hat{w}_i$ using (12).
4) Find the beamformers magnitudes and $r^*_{tpa}$ by solving the corresponding set of linear equations.
5) Update $\hat{Q}^{n+1}$ using (14).
6) Increment $n$, test for convergence. If the test fails, return to 2.

One simple termination strategy is to stop the algorithm when $[\sum_i w_i w_i^H]_{i,i} - P_t < \epsilon_i$, where $\epsilon_i$ is the maximum allowable violation of the power constraint for the $i$th antenna. The modified update in (14) and the improved step size selection result in a substantial reduction of the number of iterations required over the number required using the choices in [18]. However, it is possible to use predictive steps to further reduce the number of iterations required. As one can be active; i.e., $\hat{Q}$ has at most one non-zero entry and we have at most one non-zero dual variable $q_i$ to find. Following previous approaches, to develop a customized algorithm, we consider the following problem

\[
\begin{align*}
\min_{w_i} & \quad \sum_{i=1}^{K} w_i^H w_i, \\
\text{s.t.} & \quad \sum_{i=1}^{K} w_i w_i^H \leq P_t, \\
& \quad h_i^H Q_i h_i - \sigma^2_i - r^*_{tpa} \geq 0.
\end{align*}
\]

This problem can be solved using the following simple steps. First we obtain the closed form solution of (5) using (6) as explained in Section II. Then we check the power transmitted by each antenna. If there are no violations of the constraints, the problem is solved. If there is a violation, there can be at most one, and hence there is only one Lagrange multiplier to find. That value can be found using a simple bisection search in which the decisions to remove part of the search region is based on the satisfaction of the per-antenna power constraint. Once that value has been found, we have $Q_i$ and hence we can find the optimal $w_i$ and the optimal $r^*_{tpa}$ using the techniques in Section III.A.

IV. SIMULATION RESULTS

In this section, we will show how the application of per-antenna power constraints can substantially reduce the dynamic range of the power transmitted from each antenna, without significantly degrading the outage probability of the system. We consider a system with a BS with 4 antennas that serves 3 single-antenna users. The noise power variance is set $\sigma^2 = 0.01$, the users’ SINR targets are set to 3dB, and the channels are modelled using an i.i.d. Rayleigh fading model with equal path loss. The path loss is normalized to one. The channel uncertainty in the BS’s estimates of the channel is modelled as if the system was a TDD system with quasi-static channels and channel estimation being performed by the BS on the uplink. Under our model, independent training yields zero mean uncorrelated Gaussian estimation errors. We will set the variance of these errors to $\sigma^2_i = 0.04$.

In Fig. 1 we plot the output power per antenna for five different algorithms. The first three are the offset maximization algorithm with BS total power constraint $P_t = 1$ and no per-antenna power constraint in (5), the general algorithm having both $P_t = 1$ and per-antenna power constraints with $P_i = 0.4, \forall i$, in (7) (In all scenarios we considered, the solution to (7) with $P_t = 0.4$ has at most one active constraint and hence the same solution could have been obtained by the method in Section III.B), and the per-antenna power constrained problem in (8) with $P_i = P_t/N_t, \forall i$. Those benchmark results were obtained using generic convex optimization algorithms accessed through the cvx interface [19]. Fig. 1 also contains the results for the formulation in (8) that were obtained using both customized iterative algorithms developed in Section III.A (the original, “Orig.”, and accelerated, “Acc.”, algorithms), for termination conditions corresponding to at most 10% and 5% errors; i.e., $\epsilon_i = P_t/10$ and $\epsilon_i = P_t/20$, respectively. Fig. 1 shows each algorithm’s power distribution over the 4 BS antennas. We observe that, as expected, the modified iterative algorithms for (8) offers an output power distribution that is similar to the original one and that these solutions become closer as $\epsilon_i$ decreases. However, from Table 1 we observe that the average number of iterations is much lower for the accelerated algorithm. To assess the impact of the per-antenna power constraints on the outage probability, in Fig. 2 we have plotted the outage probability against the total transmitted power, $P_t$, for the formulation without per-antenna power constraints, a formulation with $P_i = 0.4P_t$, and formulation with $P_i = P_t/N_t, \forall i$. We also include results for the iterative solutions with $P_i = P_t/N_t$. As expected, imposing per-antenna power constraints degrades the outage performance, but in light of the significant reduction of the dynamic range requirements in the per-antenna constrained case (see Fig. 1) this degradation can be viewed as being rather modest.

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.

583
20,000 channel realizations. The SINR targets are $\gamma = 3$ dB, the noise variance is $\sigma^2 = 0.01$ and the channel estimation error variance is $\sigma_v^2 = 0.04$.

Table 1: Average number of iterations of original and accelerated algorithms, for different $P_1$ and different termination conditions.

| $P_1$  | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.35 | 0.4 | 0.45 | 0.5 |
|--------|-----|------|-----|------|-----|------|-----|------|-----|
| Orig. 10% | 44  | 38  | 30  | 26  | 21  | 18.6 | 16  | 14.5 | 13.1 |
| Acc. 10% | 3.9 | 3.7 | 3.5 | 3.7 | 3.6 | 3.7 | 3.6 | 3.5 | 3.7 |