Offset-Based Beamforming: A New Approach to Robust Downlink Transmission

Mostafa Medra, Yongwei Huang, *Senior Member, IEEE*, and Timothy N. Davidson, *Senior Member, IEEE*

***Abstract*—The design of beamformers for the multi-user multiple-input single-output downlink that provide prespecified levels of quality-of-service can be quite challenging when the channel state information is not perfectly known at the base station. The constraint of having the signal-to-interference-and-noise ratio (SINR) satisfy a given threshold with high probability is intractable in general, which results in problems that are fundamentally hard to solve. We develop a high-quality approximation of the SINR outage constraint that, along with a semidefinite relaxation, enables us to formulate the robust beamformer design problem for Gaussian channel estimation errors as a convex optimization problem that can be efficiently solved. For systems with small uncertainties, further approximations yield design algorithms based on iterative evaluations of closed-form expressions that have substantially lower computational cost. Since finding the beamforming directions incurs most of that cost, analogous power loading algorithms for predefined beamforming directions are developed and their performance is shown to be close to optimal. When the system contains a large number of antennas, the proposed power loading can be obtained at a computational cost that grows only linearly in the number of antennas. The proposed power loading algorithm provides an explicit relationship between the outage probability required and the power consumed, which allows us to precisely control the power consumption, and automatically identifies users who are consuming most of the power resources. The flexibility of the proposed approach is illustrated by developing a power loading technique that minimizes an average notion of outage.**

***Index Terms*—Broadcast channel, downlink beamforming, robust precoding, outage, low-complexity, channel uncertainty.**

Manuscript received October 25, 2017; revised May 12, 2018 and August 31, 2018; accepted October 14, 2018. Date of publication November 2, 2018; date of current version November 20, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Tsung-Hui Chang. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN-2015-06631, and in part by the National Natural Science Foundation of China under Grant 11871168. This paper was presented in part at the 49th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, November 2015, in part at the 9th IEEE Sensor Array and Multichannel Signal Processing Workshop, Rio de Janeiro, Brazil, July 2016, and in part at the 50th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, November 2016.
(Corresponding author: Mostafa Medra.)

M. Medra was with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada. He is now with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: mmedra@ece.utoronto.ca).

Y. Huang is with the School of Information Engineering, Guangdong University of Technology, University Town, Guangzhou 510006, China (e-mail: ywhuang@gdut.edu.cn).

T. N. Davidson is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada (e-mail: davidson@mcmaster.ca).

Digital Object Identifier 10.1109/TSP.2018.2879049

I. Introduction

The directional signalling capabilities of base stations (BSs) that have multiple transmit antennas enable a variety of techniques [1] for simultaneously transmitting independent messages to multiple single-antenna receivers, including dirty paper coding [2], vector perturbation precoding [3], lattice reduction precoding [4], Tomlinson-Harashima precoding [5], rate splitting [6], per-symbol beamforming [7], and conventional linear beamforming [8]. Of these signalling techniques, conventional linear beamforming has the simplest implementation and will be the focus of this paper. In particular, we will consider scenarios in which the users that have been scheduled for transmission specify the quality-of-service (QoS) that they expect to receive. In that setting, the BS designs the set of beamformers to ensure that the signal-to-interference-and-noise ratio (SINR) at each receiver meets the target level that is implicitly specified by that user's QoS requirements. When the BS has perfect knowledge of the channel to each user, the beamformers that minimize the total transmitted power required to achieve the SINR targets can be efficiently found [9]–[12]. However, in practice these channels are estimated and possibly predicted. In time division duplexing (TDD) systems the estimation is typically performed during the training phase on the uplink, whereas in frequency division duplexing (FDD) systems, each receiver estimates its channel and feeds back a quantized version of that estimate to the BS. Since the BS has only estimates of the users' channels, it can only estimate the receivers' SINRs. Those estimates are, quite naturally, uncertain and hence there is a possibility that a design performed using the estimated channels will fail to meet the SINR targets when the beamformers are implemented.

A prominent approach to designing a precoder that can control the consequent outage is to postulate a model for the uncertainty in the channel estimates and to seek designs that control the outage probability under that uncertainty model. In some cases the approach involves jointly designing the beamforming directions and the power allocated to these directions (e.g., [13]–[16]), while in other cases the beamforming directions are designed based on the channel estimates only, and the uncertainty model is incorporated into the design of the power loading; e.g., [17]–[19]. Unfortunately, in most settings the outage constraint has proven to be intractable (an exception is the case in [17]), and hence the goal has been to develop computationally efficient algorithms that can effectively manage the outage probability. One possible strategy for doing so is to seek "safe" approximations of the robust optimization problem [20]. When such approximations result in a feasible design problem, the solution is guaranteed to satisfy the constraints of the original problem, but these approximations can be quite conservative, and the corresponding design problems often require significant computational effort; e.g., [14], [15]. The alternative strategy that we will take in this paper is to develop approximations of the outage constraints that typically provide good performance and are amenable to design algorithms of much lower complexity (at least in the case of Gaussian channel estimation errors), but might not necessarily guarantee that their solution is feasible for the original problem.

The development of the proposed offset-based approach begins with the rewriting of the SINR constraint as the non-negativity of a random variable. That random variable is a non-convex quadratic function of the uncertainties, in which the quadratic kernel is a quartic function of the beamformers. The key step in the proposed approach is to approximate the non-negativity constraint on the random variable by the constraint that its mean is larger than a given multiple of its standard deviation. For the case of Gaussian channel uncertainties, the mean and standard deviation are quadratic and quartic function of the beamformers, respectively. That fact enables the application of semidefinite relaxation techniques to obtain a convex formulation of (a relaxed version of) the approximated problem. While that design technique is quite effective, the computational cost of solving the convex conic program with semidefinite constraints is significant. By making a further approximation that is suitable for systems with reasonably small uncertainties, we obtain a design formulation for which the KKT optimality conditions have a simpler structure. That simpler structure facilitates the development of an approximate solution method that only requires the iterative evaluation of closed-form expressions. An additional approximation reveals an intriguing connection to a robust downlink beamforming technique that was obtained in a fundamentally different way [16]. The "constant offset" approach in [16] was derived from an approximation of the probabilistic outage constraint by a deterministic constraint that guarantees that the SINR constraints are satisfied for all channels in a certain neighbourhood of the channel estimate. The connection to the constant offset approach shows that the approach proposed herein offers a more sophisticated approximation of the outage constraint and consequently better performance, but that it does incur a modestly larger computational cost.

An analysis of the computational cost of these precoder design techniques shows that it is the calculation of the beamforming directions that consumes most of the required computational resources, and that when these directions are defined in advance, the computational load can be significantly reduced. Accordingly, we develop variants of our precoder design algorithms that perform power loading on a set of fixed beamforming directions. These algorithms have low computational costs, and provide performance that is close to that of the optimal power loading algorithm [17]. Furthermore, for systems with a large number of antennas (i.e., "massive MIMO") in which the channel hardens, we develop a variant of our power loading algorithm that has a computational cost that grows only linearly with the number of antennas.

In practice, the BS has limited power available for transmission, and it is possible that the power required to serve the scheduled users with the required outage probabilities may exceed that limit. In some scenarios, some users suffer from a weak channel, or from having their channels closely aligned with those of other users. When that happens, such users consume most of the power transmitted by the BS. This suggests opportunities to reschedule users. On the other hand, some users might be close to the BS and experiencing a relatively strong channel; a case that suggests opportunities for doing some sort of power saving. The proposed power loading algorithm provides an explicit relationship between the required outage probabilities and the consumed power, which allows us to address these issues. Using this explicit power-outage relationship we can reduce the required power when the resulting increases in the outage probabilities are tolerable, and we can identify users that consume excessive amounts of power.

The above-mentioned designs are "fair" in the sense that they seek to provide each user with their specified outage probability. However, the proposed design techniques are quite flexible, and can accommodate other objectives, such as the sum of the outage probabilities. As we will demonstrate, such designs can improve the average performance of the users.

II. System Model

We consider a scenario in which a BS that has N_t antennas communicates with K single-antenna users over a narrow-band channel. In the linear beamforming transmission case, the transmitted signal can be written as $\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k$, where s_k is the normalized data symbol intended for user k, and $\mathbf{w}_k$ is the associated beamformer vector. For later reference we let $\mathbf{u}_k = \mathbf{w}_k / \|\mathbf{w}_k\|$ denote the beamforming direction for user k, and let $\beta_k = \|\mathbf{w}_k\|^2$ denote the power allocated to that direction. Hence, $\mathbf{w}_k = \sqrt{\beta_k}\mathbf{u}_k$. The received signal at user k is modelled as

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

where $\mathbf{h}_k^H$ is the vector of complex channel gains between the antennas at the BS and user k, and n_k is the additive zero-mean circular complex Gaussian noise at that user. Under this model, if we let σ_k^2 denote the noise variance, then the SINR at user k is

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j \neq k} |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma_k^2}. \tag{2}$$

The design of a set of beamformers $\{\mathbf{w}_k\}_{k=1}^K$ so that the SINRs satisfy specified target values (i.e., $\text{SINR}_k \geq \gamma_k$) requires the knowledge of the channel vectors $\{\mathbf{h}_k\}_{k=1}^K$. However, the BS has only estimates of $\{\mathbf{h}_k\}_{k=1}^K$, and hence its estimates of the SINRs at the receivers are uncertain. Accordingly, we will incorporate the channel uncertainty model into the design process. In particular, we will consider systems in which the uncertainty can be modelled using the simple additive model,

$$\mathbf{h}_k = \mathbf{h}_{e_k} + \mathbf{e}_k, \tag{3}$$

where $\mathbf{h}_{e_k}$ is the BS's estimate of the channel to user k, and the uncertainty in that estimate is characterized by the distribution of the elements of $\mathbf{e}_k$. In this paper, we will focus on scenarios in which $\mathbf{e}_k$ can be modelled as a circular complex Gaussian random variable with mean $\mathbf{m}_k$ and covariance $\mathbf{C}_k$; i.e., $\mathbf{e}_k \backsim \mathcal{CN}(\mathbf{m}_k, \mathbf{C}_k)$. One scenario in which that model is applicable is that of a TDD scheme operating in a slow fading environment, in which the BS estimates the channel on the uplink using a linear estimator and exploits channel reciprocity. When the channel gains are uncorrelated and the BS employs the best linear unbiased estimator (BLUE), $\mathbf{e}_k \backsim \mathcal{CN}(0, \sigma_{e_k}^2 \mathbf{I})$, and we will pay particular attention to that case. (Robust beamforming schemes for uncertainty models tailored to the FDD case were developed in [16].)

Now if we let δ_k denote the maximum tolerable outage probability for user k, the generic QoS joint beamforming and power loading problem can be written as

$$\min_{\mathbf{w}_k} \quad \sum_{k=1}^{K} \mathbf{w}_k^H \mathbf{w}_k \tag{4a}$$

$$\text{subject to} \quad \text{Prob}(\text{SINR}_k \geq \gamma_k) \geq 1 - \delta_k, \quad \forall k. \tag{4b}$$

This problem is hard to solve due to the intractable probabilistic outage constraint in (4b) even when the uncertainty is Gaussian [13]–[15]. In order to resolve that intractability, a variety of approximations of the problem in (4) by problems that are tractable have been proposed [13]–[16]. In many cases, the class of approximations that is considered is restricted to the class of "safe" approximations [20]. Such approximations are structured so that they guarantee that any solution of the approximate problem is feasible for the original problem in (4). However, in the downlink beamforming application, such approximations can be quite conservative, in the sense that the feasible set of the approximate problem is significantly smaller than that of the original problem; cf. (4). That can result in instances of the approximate problem being infeasible when the original problem has a solution, or in beamformer designs that consume significantly more power than necessary. The approximation that we will develop below is not structurally constrained in this way, but it typically performs well in practice. Furthermore, its simple form provides considerable flexibility in its application, and facilitates the development of highly-efficient algorithms.

III. Principles of the Offset-Based Approach

The derivation of the proposed approximation of the outage probability begins by rewriting $\text{SINR}_k \geq \gamma_k$ as $\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2 \geq 0$, where[1]

$$\begin{aligned}\mathbf{Q}_k(\{\mathbf{w}_k\}) &= \mathbf{w}_k \mathbf{w}_k^H/\gamma_k - \sum_{j \neq k} \mathbf{w}_j \mathbf{w}_j^H \\ &= \beta_k \mathbf{u}_k \mathbf{u}_k^H/\gamma_k - \sum_{j \neq k} \beta_j \mathbf{u}_j \mathbf{u}_j^H.\end{aligned} \tag{5}$$

That is, the probability that $\text{SINR}_k \geq \gamma_k$ is the same as the probability that the term $\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2$ is non-negative. Under the additive uncertainty model in (3), we observe that $\mathbf{h}_k^H \mathbf{Q}_k \mathbf{h}_k - \sigma_k^2$ is an indefinite quadratic function of the uncertainty, $\mathbf{e}_k$. In particular, the probability that $\text{SINR}_k \geq \gamma_k$ can be rewritten as the probability that

$$f_k(\mathbf{e}_k) = \mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} + 2\text{Re}(\mathbf{e}_k^H \mathbf{Q}_k \mathbf{h}_{e_k}) + \mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k - \sigma_k^2 \geq 0. \tag{6}$$

The key observation that underlies our offset-based approach to beamformer design is that for uncertainties $\mathbf{e}_k$ that are reasonably concentrated, if we design the beamforming vectors so that the mean value of $f_k(\mathbf{e}_k)$, denoted by μ_{f_k}, is a significant multiple of its standard deviation, denoted by σ_{f_k}, then that user will achieve a low outage probability. If we let r_k denote that multiple for the kth user, then the resulting approximation of the SINR constraint, $\text{Prob}(\text{SINR}_k \geq \gamma_k) \geq 1 - \delta_k$, can be written as

$$\mu_{f_k} \geq r_k \sigma_{f_k}. \tag{7}$$

As we explain in more detail in Appendix A, the insight that underlies this observation is the essence of Cantelli's Inequality, which is sometimes referred to as the "one-sided" Chebyshev Inequality. Indeed, we can use Cantelli's Inequality to show that if we choose $r_k = \sqrt{1/\delta_k - 1}$, then (for any distribution of $\mathbf{e}_k$) the approximation in (7) is "safe", in the sense that if the constraint in (4b) is replaced by (7), as it is in the problem in (10) below, then any solution to the approximated problem is guaranteed to satisfy the original constraint in (7). While that choice of r_k is safe, it is made so that (7) is a safe approximation for any distribution of $f_k(\mathbf{e}_k)$. In our case, $f_k(\mathbf{e}_k)$ is an indefinite quadratic function of Gaussian random variables [21], [22], and for such distributions a direct application of Cantelli's Inequality yields a rather conservative choice for r_k. Indeed, as we outline in Appendix A, in the case of normal downlink operation, in which the norm of the channel uncertainty will typically be small compared to the norm of the channel (cf. [23]), effective choices for r_k (that are not necessarily "safe") can be obtained from a Gaussian approximation of $f_k(\mathbf{e}_k)$.

To be able to use the offset approximation in (7) in a low-complexity design algorithm, we need to obtain expressions for μ_{f_k} and σ_{f_k} in terms of the design variables $\mathbf{w}_k = \sqrt{\beta_k}\mathbf{u}_k$. As shown in Appendix B, when $\mathbf{e}_k \backsim \mathcal{CN}(\mathbf{m}_k, \mathbf{C}_k)$

$$\begin{aligned}\mu_{f_k} &= \mathbb{E}\{f_k(\mathbf{e}_k)\} \\ &= (\mathbf{h}_{e_k} + \mathbf{m}_k)^H \mathbf{Q}_k (\mathbf{h}_{e_k} + \mathbf{m}_k) - \sigma_k^2 + \mathbf{w}_k^H \mathbf{C}_k \mathbf{w}_k/\gamma_k \\ &\quad - \sum_{j \neq k} \mathbf{w}_j^H \mathbf{C}_k \mathbf{w}_j^H,\end{aligned} \tag{8a}$$

$$\begin{aligned}\sigma_{f_k}^2 &= \text{var}\{f_k(\mathbf{e}_k)\} \\ &= 2(\mathbf{h}_{e_k} + \mathbf{m}_k)^H \mathbf{C}_k^{1/2} \mathbf{Q}_k^2 \mathbf{C}_k^{1/2} (\mathbf{h}_{e_k} + \mathbf{m}_k) \\ &\quad + \text{tr}(\mathbf{C}_k^{1/2} \mathbf{Q}_k \mathbf{C}_k^{1/2})^2,\end{aligned} \tag{8b}$$

where tr denotes the trace function.

From the perspective of beamformer design, an important observation is that μ_{f_k} is a non-convex quadratic function of the beamformers $\{\mathbf{w}_k\}_{k=1}^K$, but for fixed beamforming directions $\{\mathbf{u}_k\}_{k=1}^K$ it is a linear function of the power loading $\{\beta_k\}_{k=1}^K$. The variance $\sigma_{f_k}^2$ is a quartic function of the beamformers, and for fixed directions is a non-convex quadratic function of the power loading. In scenarios in which the model $\mathbf{e}_k \backsim \mathcal{CN}(0, \sigma_{e_k}^2 \mathbf{I})$ is appropriate, these expressions simplify to

$$\mu_{f_k} = \mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 + \sigma_{e_k}^2 \left(\beta_k/\gamma_k - \sum_{j \neq k} \beta_j\right), \tag{9a}$$

$$\sigma_{f_k}^2 = 2\sigma_{e_k}^2 \mathbf{h}_{e_k}^H \mathbf{Q}_k^2 \mathbf{h}_{e_k} + \sigma_{e_k}^4 \text{tr}(\mathbf{Q}_k^2). \tag{9b}$$

We will focus on this simplified case in the following sections.

IV. Offset-Based Robust Beamforming

As discussed above, the robust beamforming problem in (4) is fundamentally hard to solve due to the intractability of the probabilistic SINR outage constraint in (4b). If we were to replace

[1] In (5) we have explicitly stated the dependence of $\mathbf{Q}_k$ on the set of beamformers $\{\mathbf{w}_k\}$. For simplicity, in the rest of the paper we will leave that dependence implicit, except in the formulations of the design problems. With a mild abuse of notation, in (11), (13) and (20) we will write $\mathbf{Q}_k(\{\mathbf{W}_k\})$, where $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$.

that constraint with its offset approximation, $\mu_{f_k} \geq r_k \sigma_{f_k}$, then the problem in (4) can be approximated by

$$\min_{\mathbf{w}_k} \quad \sum_{k=1}^{K} \mathbf{w}_k^H \mathbf{w}_k \tag{10a}$$

$$\text{s.t.} \quad \mu_{f_k} \geq r_k \sigma_{f_k}, \quad \forall k. \tag{10b}$$

Some insights into the behaviour of solutions to (10) can be obtained by observing that when the values of r_k are chosen to be the same, the beamforming vectors are designed so that users with a large SINR variance are provided with a larger SINR mean. To do so, those users with a lower SINR variance are not provided with as large mean SINR as they do not need the same protection against the uncertainty.

Although the objective in (10a) is a convex quadratic function of the beamformers $\{\mathbf{w}_k\}$, the mean μ_{f_k} in (10b) is a non-convex quadratic function of $\{\mathbf{w}_k\}$, and the variance $\sigma_{f_k}^2 = 2\sigma_{e_k}^2 \mathbf{h}_{e_k}^H \mathbf{Q}_k^2 \mathbf{h}_{e_k} + \sigma_{e_k}^4 \operatorname{tr}(\mathbf{Q}_k^2)$ is quartic in $\{\mathbf{w}_k\}$. Generically, such problems are difficult to solve. The goal of this section is to develop a low-complexity algorithm for obtaining good solutions to the problem in (10). That algorithm will appear as Algorithm 1 below. As the first step in that development, we now develop a semidefinite relaxation (SDR) of (10). While the SDR of (10) plays a particular role in the development of Algorithm 1, we observe that there are other strategies that could be taken to obtain good solutions to (10), including successive convex approximation; e.g., [24]–[26].

A. A Semidefinite Relaxation of (10)

To develop a semidefinite relaxation of (10), we observe that if we make the substitution $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$, then the functions on either side of the inequality in (10b) become a linear function of $\{\mathbf{W}_k\}$ and the square root of a quadratic function of $\{\mathbf{W}_k\}$, respectively, and the objective becomes linear. As such, the remaining difficulty in the reformulation of the problem is the set of rank-one constraints on $\mathbf{W}_k$. If we relax those constraints we obtain the following semidefinite relaxation of the problem in (10)

$$\min_{\mathbf{W}_k, d_{1k}, d_{2k}} \quad \operatorname{tr}\left(\sum_{k=1}^{K} \mathbf{W}_k\right) \tag{11a}$$

$$\text{s.t.} \quad \mathbf{h}_{e_k}^H \mathbf{Q}_k(\{\mathbf{W}_k\}) \mathbf{h}_{e_k} - \sigma_k^2 + \sigma_{e_k}^2 \operatorname{tr}(\mathbf{W}_k)/\gamma_k - \sigma_{e_k}^2 \operatorname{tr}\left(\sum_{j\neq k} \mathbf{W}_j\right) \geq r_k \|[d_{1k}\ d_{2k}]\|, \tag{11b}$$

$$d_{1k} \geq \sqrt{2}\sigma_{e_k} \|\mathbf{h}_{e_k}^H \mathbf{Q}_k(\{\mathbf{W}_k\})\|, \tag{11c}$$

$$d_{2k} \geq \sigma_{e_k}^2 \|\mathbf{Q}_k(\{\mathbf{W}_k\})\|_F, \tag{11d}$$

$$\mathbf{W}_k \succeq \mathbf{0}, \quad \forall k, \tag{11e}$$

where $\|\cdot\|_F$ represents the Frobenius norm of the matrix. This problem is a convex conic optimization problem and can be efficiently solved using interior-point methods. Two refined implementations of those methods are easily accessible through the MATLAB-based CVX tool [27].[2] In our numerical experience, the rank of the optimal $\mathbf{W}_k$'s in (11) is usually one. When that occurs, the semidefinite relaxation is tight and the optimal beamformer vectors $\mathbf{w}_k$ for the problem in (10) can be directly obtained from the optimal matrices $\mathbf{W}_k$.[3] When that does not occur, good solutions to (10) can typically be obtained by using the Gaussian randomization technique; cf. [30].

B. Low-Complexity Precoding Algorithm

Although the problem in (11) is convex, its solution still imposes a significant computational load, even for a moderate number of antennas. In this section, we will show how insights into the KKT conditions of a mild approximation of a problem obtained from the problem in (11) enable the development of a low-complexity design algorithm consisting of the iterative evaluation of a sequence of closed-form expressions.

Our approximation of (11) is based on the observation, made above, that in practical downlink systems the uncertainty in the channel estimates must be small in order for the system to support reasonable rates [23]. In such scenarios, the term in (10b) containing $\sigma_{e_k}^4$ will typically be significantly smaller than the other term. Accordingly, $\sigma_{f_k}^2 \approx 2\sigma_{e_k}^2 \mathbf{h}_{e_k}^H \mathbf{Q}_k^2 \mathbf{h}_{e_k}$ is a reasonable approximation. Applying this approximation in the context of the problem in (10) we obtain the following approximation of (10b)

$$\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 + \sigma_{e_k}^2 \mathbf{w}_k^H \mathbf{w}_k/\gamma_k - \sigma_{e_k}^2 \sum_{j\neq k} \mathbf{w}_j^H \mathbf{w}_j \geq r_k \sqrt{2}\sigma_{e_k} \|\mathbf{h}_{e_k}^H \mathbf{Q}_k\|. \tag{12}$$

The semidefinite relaxation of the resulting approximation of the problem in (10) can be written as

$$\min_{\mathbf{W}_k, \mathbf{d}_k} \quad \operatorname{tr}\left(\sum_{k=1}^{K} \mathbf{W}_k\right) \tag{13a}$$

$$\text{s.t.} \quad \mathbf{h}_{e_k}^H \mathbf{Q}_k(\{\mathbf{W}_k\}) \mathbf{h}_{e_k} - \sigma_k^2 + \sigma_{e_k}^2 \operatorname{tr}(\mathbf{W}_k)/\gamma_k - \sigma_{e_k}^2 \operatorname{tr}\left(\sum_{j\neq k} \mathbf{W}_k\right) \geq \|\mathbf{d}_k\|, \tag{13b}$$

$$\mathbf{d}_k = r_k \sqrt{2}\sigma_{e_k} \mathbf{Q}_k(\{\mathbf{W}_k\}) \mathbf{h}_{e_k}, \tag{13c}$$

$$\mathbf{W}_k \succeq \mathbf{0}, \quad \forall k. \tag{13d}$$

As was the case in (11), we note that the problem in (13) is over parameterized; the vectors $\mathbf{d}_k$ are not needed. However, this over parameterization will simplify the following analysis. The problem in (13) is another convex conic program, and when the approximation that led to (12) is accurate, good solutions to (10) can typically be obtained using the techniques described after (11). However, the presence of the semidefinite constraints means that considerable computational effort is still required.

To develop a more efficient algorithm for generating good solutions to (10) based on the approximation in (12), we will

[2] In the particular formulation in (11), each SINR constraint in (10b) is replaced by three second order cone (SOC) constraints that involve two additional variables, d_{1k} and d_{2k}. While those additional variables will play a role in the development of Algorithm 1, there are equivalent formulations that are more compact.

[3] For some related, but structurally different, robust beamforming problems with bounded uncertainties, conditions under which the semidefinite relaxation is tight have been established [28], [29].

consider the "beamforming variant" of (13). That is, we will substitute $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$ into (13), remove the constraint in (13d), and consider the problem as being an optimization problem over $\{\mathbf{w}_k\}$ and $\{\mathbf{d}_k\}$. Then we will seek beamformers $\mathbf{w}_k$, vectors $\mathbf{d}_k$, and the associated Lagrange multipliers that approximate a stationary solution to the beamforming variant of (13), in the sense that they approximately satisfy the KKT conditions for that problem.

To formalize that approach, we let ν_k denote the dual variable for the constraint in (13b), and let $\boldsymbol{\psi}_{f_k}$ denote the vector of dual variables for the equality constraint in (13c). The Lagrangian of the beamforming variant of (13) can then be written as

$$\mathcal{L}(\mathbf{w}_k, \mathbf{d}_k, \nu_k, \boldsymbol{\psi}_{f_k}) = \sum_{k=1}^{K} \mathbf{w}_k^H \mathbf{w}_k - \sum_{k=1}^{K} \nu_k \Bigg(\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k} - \sigma_k^2 + \sigma_{e_k}^2 \mathbf{w}_k^H \mathbf{w}_k / \gamma_k - \sigma_{e_k}^2 \sum_{j \neq k} \mathbf{w}_j^H \mathbf{w}_j - \|\mathbf{d}_k\| \Bigg) - \sum_{k=1}^{K} \boldsymbol{\psi}_{f_k}^H (\mathbf{d}_k - r_k \sqrt{2} \sigma_{e_k} \mathbf{Q}_k \mathbf{h}_{e_k}). \tag{14}$$

From the KKT conditions of the beamforming variant of (13), we can deduce that

$$\mathbf{w}_k = \Bigg(\frac{\nu_k}{\gamma_k} \mathbf{h}_{e_k} \mathbf{h}_{e_k}^H - \sum_{j \neq k} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H + \frac{\nu_k \sigma_{e_k}^2}{\gamma_k} \mathbf{I} - \sum_{j \neq k} \nu_j \sigma_{e_j}^2 \mathbf{I} - \frac{r_k \sigma_{e_k}}{\sqrt{2}\gamma_k} (\boldsymbol{\psi}_{f_k} \mathbf{h}_{e_k}^H + \mathbf{h}_{e_k} \boldsymbol{\psi}_{f_k}^H) + \sum_{j \neq k} \frac{r_j \sigma_{e_j}}{\sqrt{2}} (\boldsymbol{\psi}_{f_j} \mathbf{h}_{e_j}^H + \mathbf{h}_{e_j} \boldsymbol{\psi}_{f_j}^H) \Bigg) \mathbf{w}_k, \tag{15}$$

which is an eigen equation for the direction $\mathbf{u}_k$. Using a similar approach to the perfect CSI case [12], we can rearrange this equation to obtain the following expression for ν_k,

$$\nu_k^{-1} = \mathbf{h}_{e_k}^H \Bigg(\mathbf{I} + \sum_{j} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H - \frac{\nu_k \sigma_{e_k}^2}{\gamma_k} \mathbf{I} + \sum_{j \neq k} \nu_j \sigma_{e_j}^2 \mathbf{I} + \frac{r_k \sigma_{e_k}}{\sqrt{2}\gamma_k} (\boldsymbol{\psi}_{f_k} \mathbf{h}_{e_k}^H + \mathbf{h}_{e_k} \boldsymbol{\psi}_{f_k}^H) - \sum_{j \neq k} \frac{r_j \sigma_{e_j}}{\sqrt{2}} (\boldsymbol{\psi}_{f_j} \mathbf{h}_{e_j}^H + \mathbf{h}_{e_j} \boldsymbol{\psi}_{f_j}^H) \Bigg)^{-1} \mathbf{h}_{e_k} \left(1 + \frac{1}{\gamma_k} \right). \tag{16}$$

Therefore, if we can determine $\{\nu_k\}$ and $\{\boldsymbol{\psi}_{f_k}\}$ that satisfy the set of K equations in (16), then we can determine the directions $\{\mathbf{u}_k\}$ using (15). With those directions determined, finding a "KKT point" of the beamforming variant of (13) reduces to a power loading problem for the powers $\{\beta_k\}$ to be allocated to each direction. While those powers can be efficiently obtained by using the fact that the constraints in (13b) holding with equality at optimality, we will show in Section V that for given beamforming directions the optimal power loading for the original offset problem in (10) can also be efficiently obtained. Since those powers are computed without the approximation that led to (13), we will use the power loading algorithm in Section V in our approach. Therefore, the problem that remains is to efficiently determine the $\{\nu_k\}$ and $\{\boldsymbol{\psi}_{f_k}\}$ that satisfy the K equations in (16).

The expressions in (16) share a similar structure to those obtained for the corresponding QoS problem in the case of perfect CSI at the BS [12], but the matrix components of each equation contain four additional terms that are dependent on the variance of the channel estimation error. The structure of those terms suggests that if we could find a reasonable approximation for the vectors $\boldsymbol{\psi}_{f_k}$, we would obtain an iterative closed-form solution. To do so, we observe that the variable $\mathbf{d}_k$ in (13c) appears in the Lagrangian in the term $\nu_k \|\mathbf{d}_k\| - \boldsymbol{\psi}_{f_k}^H \mathbf{d}_k$. Accordingly, from the stationarity component of the KKT conditions we have that $\|\boldsymbol{\psi}_{f_k}\| = \nu_k$ and that $\mathbf{d}_k$ and $\boldsymbol{\psi}_{f_k}$ are in the same direction; i.e., $\mathbf{d}_k / \|\mathbf{d}_k\| = \boldsymbol{\psi}_{f_k} / \|\boldsymbol{\psi}_{f_k}\|$. Accordingly, we can write

$$\boldsymbol{\psi}_{f_k} = \nu_k \mathbf{d}_k / \|\mathbf{d}_k\|. \tag{17}$$

Since $\mathbf{d}_k = r_k \sqrt{2} \sigma_{e_k} \mathbf{Q}_k \mathbf{h}_{e_k}$, $\boldsymbol{\psi}_{f_k}$ explicitly depends on the beamforming directions, which have not yet been determined. However, we observe that if we substitute (17) into (16), the terms involving $\mathbf{d}_k$ are multiplied by the standard deviation of the error, σ_{e_k}. As we have already argued in the derivation of the approximations that lead to (11), σ_{e_k} will be small in effective downlink beamforming schemes, and this suggests that reasonable initial approximations of the directions should yield a good approximation of $\{\nu_k\}$, and hence a good set of beamforming directions. We suggest the use of the zero-forcing (ZF) directions [31] for the estimated channels, which we will denote by $\mathbf{u}_{z_k}$. When we use that initialization, the initial direction of $\mathbf{d}_k$ will be the same as $\mathbf{u}_{z_k}$, which allows us to simplify (16) to

$$\nu_k^{-1} = \mathbf{h}_{e_k}^H \Bigg(\mathbf{I} + \sum_{j} \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H - \frac{\nu_k \sigma_{e_k}^2}{\gamma_k} \mathbf{I} + \sum_{j \neq k} \nu_j \sigma_{e_j}^2 \mathbf{I} + \frac{r_k \sigma_{e_k} \nu_k}{\sqrt{2}\gamma_k} (\mathbf{u}_{z_k} \mathbf{h}_{e_k}^H + \mathbf{h}_{e_k} \mathbf{u}_{z_k}^H) - \sum_{j \neq k} \frac{r_j \sigma_{e_j} \nu_j}{\sqrt{2}} (\mathbf{u}_{z_j} \mathbf{h}_{e_j}^H + \mathbf{h}_{e_j} \mathbf{u}_{z_j}^H) \Bigg)^{-1} \mathbf{h}_{e_k} \left(1 + \frac{1}{\gamma_k} \right). \tag{18}$$

Since the terms other than $\{\nu_k\}$ in (18) are constant, that set of K equations forms a set of fixed-point equations for $\{\nu_k\}$. In the limit of small estimation errors, the conventional fixed-point approach to solving (18) is known to converge to the optimal values for the dual variables [32]; cf. (22) below. While a formal proof of convergence in the presence of estimation errors remains a work in progress, in our simulations that approach was applied to (18) and was observed to converge to the optimal dual variables. (When the optimal solution to (13) has rank-one matrices, the optimal dual variables $\{\nu_k\}$ are implicitly generated by the application to (13) of primal-dual interior-point methods, such as those that can be accessed through CVX [27].)

Algorithm 1: Single-Cycle Closed-Form Beamformer Design.

1: Find the ZF directions $\{\mathbf{u}_{z_k}\}$.
2: Find each ν_k using (18).
3: Find each $\mathbf{u}_k$ using the variant of (15) in which $\boldsymbol{\psi}_{f_k} = \nu_k \mathbf{u}_{z_k}$.
4: Apply the power loading developed in Section V.

The first cycle of the algorithm derived above is summarized in the sequence of closed-form operations in Algorithm 1. While the design obtained from that first cycle can be improved by using the beamformers obtained in steps 3 and 4 to obtain a refined estimate of the direction of $\mathbf{d}_k$ and returning to step 2 of the algorithm, the simulation results in Section VI suggest that the single-cycle approach taken in Algorithm 1 produces a solution whose performance is quite close to that of the original offset-based design formulation in (11). That suggests that in the scenarios that we have considered, the underlying approximations are working quite well.

C. Connection to the Constant-Offset Algorithm in [16]

As is apparent from the derivation in the previous section, one of the challenges that complicates the design procedure, and in particular the closed-form calculations, is the quartic dependence of the variances $\sigma^2_{f_k}$ on the beamforming vectors $\mathbf{w}_k$ on the right hand side of the inequality $\mu_{f_k} \geq r_k \sigma_{f_k}$ in (7). One ad-hoc way in which these complications can be reduced is to modify the offset approximation in (7) so that the mean is constrained to be greater than a constant rather than a multiple of its standard deviation; i.e., the SINR outage constraint in (4b) is replaced by

$$\mu_{f_k} \geq c_k. \tag{19}$$

Replacing $r_k \sigma_{f_k}$, which depends on the beamformers, by the constant c_k results in an approximation that is, structurally, a more coarse approximation than the one we are using in this paper; cf. (7). However, it does simplify the design problem. In particular, if we make the approximation that the channel estimation errors are small enough that the third term on the right hand side of (9a) can be neglected, the variant of the problem in (10) that uses the approximation in (19) can be written as

$$\min_{\mathbf{W}_k} \quad \operatorname{tr}\left(\sum_{k=1}^{K} \mathbf{W}_k\right) \tag{20a}$$

$$\text{s.t.} \quad \mathbf{h}_{e_k}^H \mathbf{Q}_k(\{\mathbf{W}_k\})\mathbf{h}_{e_k} - \sigma_k^2 \geq c_k, \tag{20b}$$

$$\mathbf{W}_k \succeq \mathbf{0}, \quad \forall k, \tag{20c}$$

Interestingly, the problem in (20) also arose in a fundamentally different approach to robust downlink beamforming design that was originally developed for FDD systems [16]. In that work, the outage probability was approximated using a "zero-outage region" approach, in which the SINR constraints are required to hold for all the channels in a neighbourhood of the estimated channel. In doing so, the chance-constrained robust design problem in (4) is approximated by a worst-case robust design problem. An approximation of a linear matrix inequality in a reformulation of the worst-case robust design problem led to the design problem in (20).

The structure of the problem in (20) is the same as that of the semidefinite relaxation of the nominal SINR constrained beamformer design problem in which the channel estimates are deemed to be precise. The difference is that the term that plays the role of the noise variance at user k is modified to $\sigma_k^2 + c_k$; [16]. Therefore, the semidefinite relaxation is tight [10] and a provably-convergent iterative closed-form solution is available [12], [30]. However, guidance on appropriate choices for the constant offset c_k in (19) is difficult to obtain. In contrast, as outlined in Appendix A, Cantelli's Inequality and the Gaussian approximation of $f_k(\mathbf{e}_k)$ provide ample guidance on the choice of the multiplier r_k in the offset approximation in (7).

We will use the optimal beamforming directions of the constant offset approach (which are the same as the optimal directions for the nominal problem [16]) as a benchmark in our simulations, and hence we briefly review its iterative closed-form algorithm. We also develop a reduced-complexity variant of that algorithm that is suitable for "massive MIMO" applications.

The iterative closed-form solution to (20) has a similar structure to that in Algorithm 1, but given the simpler structure of the problem, the Lagrange multipliers $\boldsymbol{\psi}_{f_k}$ disappear, and the expressions in (15) and (16) simplify to

$$\mathbf{w}_k = \left(\frac{\nu_k}{\gamma_k}\mathbf{h}_{e_k}\mathbf{h}_{e_k}^H - \sum_{j\neq k}\nu_j \mathbf{h}_{e_j}\mathbf{h}_{e_j}^H\right)\mathbf{w}_k, \tag{21}$$

$$\nu_k^{-1} = \mathbf{h}_{e_k}^H\left(\mathbf{I} + \sum_j \nu_j \mathbf{h}_{e_j}\mathbf{h}_{e_j}^H\right)^{-1}\mathbf{h}_{e_k}\left(1 + \frac{1}{\gamma_k}\right). \tag{22}$$

Conventional fixed-point algorithms to solve (22) and convergence analysis were proposed in [32], and were shown therein that the convergence is guaranteed. In Appendix C, we provide a more efficient algorithm to obtain $\{\nu_k\}$ than that proposed in [32].

After obtaining the beamforming directions from (22) and (21), the power loading in [16] is performed based on the fact that the constraints in (20b) are satisfied with equality at optimality. (If this were not the case for constraint k, then the power allocated to $\mathbf{w}_k$ could be reduced in a way that will still satisfy all the constraints and provide a lower objective value, contradicting the presumed optimality.) While doing so generates a solution to (20), significant performance gains can be obtained when the beamforming directions obtained from (21) are combined with the power loading algorithm presented in Section V.

D. Complexity Analysis and Further Approximations

The problems in (11) and (13) are convex optimization problems with SOC and semidefinite constraints. General purpose interior-point methods for such problems require $\mathcal{O}(N_t^6)$ per iteration, which represents a significant computational load. In contrast, the key computational steps in the iterative closed-form approximation, Algorithm 1, are those in (15), (18) and the calculation of the ZF directions that are used in the initialization. The ZF directions can be obtained in $\mathcal{O}(N_t K^2)$ operations. The computational cost of solving (18) is dominated by the matrix inversion required for each user and hence it grows as $\mathcal{O}(N_t^3 K)$. We can exploit the factorized matrix structure in (15) which allows for an efficient use of the power iteration method. Therefore, the cost of step 3 grows as $\mathcal{O}(N_t K^2)$. We can see that it is the computation of the Lagrange multipliers (18) that requires most of the resources to compute the beamforming directions.

The constant-offset algorithm [16] that was reviewed in Section IV-C does not require an initial set of directions and the expression for ν_k is significantly simpler. In particular, the matrix to be inverted is the same for each user, which reduces the number of computations required to $\mathcal{O}(N_t^3)$. Furthermore, additional approximations can be applied to avoid the matrix inversion all together. When the channels are nearly orthogonal, as they tend to be in massive MISO channels that "harden" as the number of antennas increases [33], then if we let $\alpha_k = \|\mathbf{h}_{e_k}\|^2$, we can write $\sum_j \nu_j \mathbf{h}_{e_j} \mathbf{h}_{e_j}^H$ in the form of an eigen decomposition $\sum_j \nu_j \alpha_j \frac{\mathbf{h}_{e_j}}{\sqrt{\alpha_j}} \frac{\mathbf{h}_{e_j}^H}{\sqrt{\alpha_j}}$, and hence,

$$\mathbf{h}_{e_k}^H \left(\mathbf{I} + \sum_j \nu_j \alpha_j \frac{\mathbf{h}_{e_j}}{\sqrt{\alpha_j}} \frac{\mathbf{h}_{e_j}^H}{\sqrt{\alpha_j}} \right)^{-1} \mathbf{h}_{e_k} \approx \frac{\alpha_k}{1 + \nu_k \alpha_k}.$$

Accordingly, we can approximate (22) by

$$\nu_k \approx \gamma_k / \alpha_k.$$

To find the channel norms $\alpha_k = \|\mathbf{h}_{e_k}\|^2$ we need only $\mathcal{O}(N_t)$ operations. Hence, that approximation enables us to compute all ν_ks in only $\mathcal{O}(N_t K)$ operations.

V. Offset-Based Robust Power Loading

In the previous section, we considered the problem of directly designing the beamformers $\mathbf{w}_k = \sqrt{\beta_k}\mathbf{u}_k$. In the case of the approximated problem in (13), we developed a low-complexity approximate algorithm in which the design of the directions $\{\mathbf{u}_k\}$ and the power allocation $\{\beta_k\}$ naturally decouples. In this section, we will show that when the beamforming directions are chosen separately, we can obtain a low-complexity algorithm for the power loading variant of the original offset-based design problem in (10), without having to make the approximation in (12). Examples of choices for those directions include the maximum ratio transmission (MRT), ZF, or regularized ZF (RZF) directions, which are calculated from the estimated channels, or any of the directions generated by the previously described algorithms. The algorithm developed in this section may be of particular interest in massive MIMO settings, where the use of simple beamforming directions is often preferred in order to reduce the computational cost.

For a given set of beamforming directions $\{\mathbf{u}_k\}$, the original offset-based design problem in (10) reduces to

$$\min_{\beta_k} \quad \sum_{k=1}^{K} \beta_k \tag{23a}$$

$$\text{s.t.} \quad \mu_{f_k} \geq r_k \sigma_{f_k}, \quad \forall k, \tag{23b}$$

where the expressions for μ_{f_k} and σ_{f_k} in (9a) and (9b) simplify to

$$\mu_{f_k} = |\mathbf{h}_{e_k}^H \mathbf{u}_k|^2 \beta_k/\gamma_k - \sum_{j \neq k} |\mathbf{h}_{e_k}^H \mathbf{u}_j|^2 \beta_j - \sigma_k^2 + \sigma_{e_k}^2 \left(\beta_k/\gamma_k - \sum_{j \neq k} \beta_j \right). \tag{24a}$$

$$\sigma_{f_k}^2 = 2\sigma_{e_k}^2 \mathbf{h}_{e_k}^H \left(\beta_k \mathbf{u}_k \mathbf{u}_k^H/\gamma_k - \sum_{j \neq k} \beta_j \mathbf{u}_j \mathbf{u}_j^H \right)^2 \mathbf{h}_{e_k} + \sigma_{e_k}^4 \operatorname{tr} \left(\beta_k \mathbf{u}_k \mathbf{u}_k^H/\gamma_k - \sum_{j \neq k} \beta_j \mathbf{u}_j \mathbf{u}_j^H \right)^2. \tag{24b}$$

Since μ_{f_k} is linear in $\{\beta_k\}$ and σ_{f_k} is a convex quadratic function of $\{\beta_k\}$, the problem in (23) can be rewritten as an SOC programming problem, and an optimal solution can be efficiently obtained using generic interior-point methods. However, the structure of the constraints suggests that an iterative linearization approach would enable the development of a more efficient algorithm. In particular, if we were to calculate $\{\sigma_{f_k}\}$ using the current values for $\{\beta_k\}$ and then treat those values for $\{\sigma_{f_k}\}$ as constants, the problem in (23) would become a linear program. Furthermore, the constraints in (23b) would hold with equality at optimality. If that were not the case for constraint k, then β_k could be reduced in a way that still satisfies the constraints and yet provides a lower objective value, which would contradict the presumed optimality. The fact that the linearized version of (23b) would hold with equality at optimality would provide a set of K linear equations for the K updated values for $\{\beta_k\}$. Those updated values could then be used to update the values for $\{\sigma_{f_k}\}$ and the next iteration would begin.

In order to state that algorithm more precisely, let us define $\boldsymbol{\beta} = [\beta_1, \beta_2, ..., \beta_K]^T$, $\boldsymbol{\sigma}_f = [\sigma_{f_1}, \sigma_{f_2}, ..., \sigma_{f_K}]^T$, $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, ..., \sigma_K]^T$, $\mathbf{r} = [r_1, r_2, ..., r_k]^T$, and let us also define the matrix $\mathbf{A}$ such that $[\mathbf{A}]_{ii} = |\mathbf{h}_{e_i}^H \mathbf{u}_i|^2/\gamma_i + \sigma_{e_i}^2/\gamma_i$, and $[\mathbf{A}]_{ij} = -|\mathbf{h}_{e_i}^H \mathbf{u}_j|^2 - \sigma_{e_i}^2, \forall i \neq j$. Given the current values for $\{\sigma_{f_k}\}$, which are computed from the current values for $\{\beta_k\}$ and are to be treated as constants, the set of linear equations for updating the values for $\{\beta_k\}$ can be written as

$$\mathbf{A}\boldsymbol{\beta} = \boldsymbol{\sigma}^2 + \boldsymbol{\sigma}_f \odot \mathbf{r}, \tag{25}$$

in which $\odot$ represents element-by-element multiplication. Once the values of $\{\beta_k\}$ have been updated, we can update the value of $\boldsymbol{\sigma}_f$ using (24b). The resulting iterative linearization algorithm for solving (23) is summarized in Algorithm 2.

The computational effort required to initialize Algorithm 2 consists of $\mathcal{O}(K^2 N_t)$ operations to compute $\mathbf{A}$ and $\mathcal{O}(K^3)$ operations to compute $\mathbf{A}^{-1}$. In each iteration, the computational cost for step 2 is $\mathcal{O}(K^2)$ operations, and the cost of step 3 is $\mathcal{O}(K^2 N_t)$ operations. Our numerical experience, some of which is reported in Section VI, suggests that Algorithm 2 converges, and that the number of iterations needed for near-optimal performance is very small. While a formal convergence proof remains a topic for future work, some insight can be obtained by observing the dependence of $\boldsymbol{\sigma}_f$ on $\boldsymbol{\beta}$ in (24b), and viewing Algorithm 2 as a fixed-point technique for solving $\boldsymbol{\beta} = \mathbf{A}^{-1}\boldsymbol{\sigma}^2 + \mathbf{A}^{-1}(\boldsymbol{\sigma}_f \odot \mathbf{r})$. The eigenvalues of $\mathbf{A}^{-1}$ influence the convergence properties for these fixed-point equations.

Algorithm 2: The Power Loading Algorithm.

1: Initialize $\sigma_{f_k} = 1$. Compute $\mathbf{A}$ and $\mathbf{A}^{-1}$.
2: Find $\boldsymbol{\beta}$ by solving the set of linear equations in (25).
3: Update each σ_{f_k} using (24b).
4: Return to 2 until a termination criterion is satisfied.

Since the matrix $\mathbf{A}$ typically has large diagonal values representing the signal powers, and lower values on the off-diagonal elements representing the interference powers, the eigen values of $\mathbf{A}^{-1}$ will typically be less than one.

A. Simplifying the SINR Variance Calculation

The above analysis shows that the step in Algorithm 2 that is computationally intensive is the computation of σ_{f_k}. In massive MISO systems, the resulting computational load can be significant. To reduce the required computations, we observe that when the number of antennas is large and the channels are uncorrelated, the inner product between different channels will typically be relatively small. Since the beamforming directions will typically be closely aligned with the channel vectors, the inner product between different beamforming vectors will likely be small as well. This observation suggests removing the cross terms $\mathbf{u}_j^H \mathbf{u}_k, \forall j \neq k$ in (24b). That would yield the following approximations

$$\begin{aligned}\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{Q}_k \mathbf{h}_{e_k} &= \mathbf{h}_{e_k}^H \left(\beta_k \mathbf{u}_k \mathbf{u}_k^H/\gamma_k - \sum_{j\neq k} \beta_j \mathbf{u}_j \mathbf{u}_j^H\right)^2 \mathbf{h}_{e_k} \\ &\approx |\mathbf{h}_{e_k}^H \mathbf{u}_k|^2 \beta_k^2/\gamma_k^2 + \sum_{j\neq k} |\mathbf{h}_{e_k}^H \mathbf{u}_j|^2 \beta_j^2,\end{aligned} \tag{26}$$

and

$$\begin{aligned}\mathrm{tr}(\mathbf{Q}_k^2) &= \mathrm{tr}\left(\beta_k \mathbf{u}_k \mathbf{u}_k^H/\gamma_k - \sum_{j\neq k} \beta_j \mathbf{u}_j \mathbf{u}_j^H\right)^2 \\ &\approx \mathrm{tr}\left(\beta_k^2 \mathbf{u}_k \mathbf{u}_k^H \mathbf{u}_k \mathbf{u}_k^H/\gamma_k^2 + \sum_{j\neq k} \beta_j^2 \mathbf{u}_j \mathbf{u}_j^H \mathbf{u}_j \mathbf{u}_j^H\right) \\ &= \beta_k^2/\gamma_k^2 + \sum_{j\neq k} \beta_j^2.\end{aligned} \tag{27}$$

The numerical results presented in Section VI indicate that these approximations result in designs that are very close in performance to those obtained from the original formulations, even when the number of antennas is quite small. Furthermore, since the terms $|\mathbf{h}_{e_k}^H \mathbf{u}_j|^2$ are already computed in the initialization step that constructs the matrix $\mathbf{A}$, these approximations reduce the computational cost of updating $\boldsymbol{\sigma}_f$ in step 3 of Algorithm 2 from $\mathcal{O}(N_t K^2)$ to $\mathcal{O}(K^2)$.

B. User Rescheduling

One of the fundamental characteristics of the original outage constrained beamformer design problem in (4) is that for a certain set of channel estimates the problem may be infeasible. That is, there may be no set of beamformers that can satisfy the outage constraints. Furthermore, even when the problem is feasible, the solution may be impractical in the sense that the minimum transmission power required to satisfy the outage constraints may exceed the capability of the BS. The approximations of the original formulation in (11) and (13) retain these characteristics, and the power loading problem in (23) retains them, too. Fortunately, as we now explain, for systems in which each user specifies the same value for r, the structure of a closely-related power loading problem provides insights into which users should be rescheduled in order for the problem in (23) to be feasible, and for the solution of the problem to be within the capabilities of the BS. The auxiliary power loading problem that we will consider is that of maximizing a common offset coefficient subject to an explicit power constraint, namely

$$\max_{\beta_k, r} \quad r \tag{28a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \beta_k \leq P_t, \tag{28b}$$

$$\mu_k \geq r\sigma_{f_k}, \quad \forall k, \tag{28c}$$

where P_t denotes the maximum transmission power of the BS. The problem in (28) complements that in (23), in the sense that it provide the "greatest" robustness (in the offset sense) for a given power constraint, whereas the problem in (23) seeks to minimize the power required to meet each user's offset-based outage target. This complementary relationship is analogous to the relationship between "max-min" fairness and QoS problems in the perfect CSI case; e.g., [12].

From the perspective of efficient user rescheduling, an important feature of the problem in (28) is that it is always feasible whenever all the estimated channels are different. (The value of r can be decreased until all components of (28c) can be satisfied using a power loading that satisfies (28b).) However, negative values and small positive values of r correspond to cases with high probability of outage. It is that feature that we will exploit for user rescheduling.

The problem in (28) can be solved using an algorithm similar to that in Algorithm 2. However, at the step analogous to step 2 of Algorithm 2, we need an additional equation to determine the value for r. That equation arises from observing that the power constraint in (28b) holds with equality at optimality, and hence, from (25) and (28b) we have that

$$r = \frac{P_t - \mathbf{1}^T \mathbf{A}^{-1} \boldsymbol{\sigma}^2}{\mathbf{1}^T \mathbf{A}^{-1} \boldsymbol{\sigma}_f},$$

where $\mathbf{1}$ is the vector with all elements equal to one. This equation clearly demonstrates the relationship between the power budget and the robustness. More importantly, it indicates that the users that correspond to the largest elements of $\mathbf{A}^{-1}\boldsymbol{\sigma}^2$ play a large role in constraining the extent of robustness that can be obtained. That suggests that if the optimal value of r in (28) is not large enough to provide the desired robustness level, one or more of those users corresponding to large values of $\mathbf{A}^{-1}\boldsymbol{\sigma}^2$ should be rescheduled.[4]

A simple way to do that would be to perform a "greedy" search in which we incrementally reschedule active users corresponding to the largest value of $\mathbf{A}^{-1}\boldsymbol{\sigma}^2$, recalculate the reduced

[4]We note that the use of good user selection algorithms, e.g., [34], prior to the design of the beamforming directions will reduce the need to reschedule users, but the inherent capability of the proposed power loading algorithms to perform rescheduling provides significant performance gains when the initial user selection is imperfect.

dimension matrices $\mathbf{A}$ and $\mathbf{A}^{-1}$, and solve the reduced dimension version of (28). Once the optimal value of the reduced dimension version of (28) exceeds the desired value for r, the power minimization problem in (23) can be solved. In order to implement this approach we need to select a desired value for r. As outlined in Appendix A, we could take the "safe" approximation approach suggested by Cantelli's Inequality, but that choice could be quite conservative, resulting in many users being rescheduled. The discussion in Appendix A suggests that in the context of robust downlink beamforming, the guidance on the choice of r provided by the Gaussian approximation of $f_k(\mathbf{e}_k)$ is more suitable. For an outage probability target δ_k that is between 10% and 0.1%, the desired value of r_k would usually be chosen to be in the range of 1.3 to 3.1; cf. Appendix A.

C. Average Outage

The design formulations that we have considered up until this point have taken the form of minimization of the transmission power subject to (an approximation of) an outage constraint on each user for the current realizations of the channels. However, as we now illustrate, the proposed design approach is quite flexible and can accommodate other notions of outage.

Let us assume that we have the optimal power loading and offset coefficient for the problem in (28), which provide all the users with essentially the same outage probability. We will denote those values by $\{\beta_k^\star\}_{k=1}^K$ and $r^\star$. Given this solution, the goal of this section is to perturb the value of the offset coefficient for each user so as to minimize the average outage probability over the users, and to adjust the power loading accordingly. To do so, we let δ_{r_k} denote the perturbation on the kth user's offset coefficient; i.e., $r_k = r^\star + \delta_{r_k}$. As discussed in Section III, in typical operating scenarios the distribution of $f_k(\mathbf{e}_k)$ can be accurately approximated by a Gaussian distribution. In that case, the outage probability for a given value of the offset coefficient r_k is simply the value of the complementary cumulative distribution function (CCDF) of the standardized normal distribution, $\mathcal{N}(0,1)$, at the value of r_k. If we let $g(\cdot)$ denote the CDF of the standard normal distribution, the the problem of minimizing the outage probability becomes

$$\max_{\beta_k,\delta_{r_k}} \quad \sum_{k=1}^K g(r^\star + \delta_{r_k}) \tag{29a}$$

$$\text{s.t.} \quad \sum_{k=1}^K \beta_k = P_t, \tag{29b}$$

where the condition $\sum_{k=1}^K \beta_k = P_t$ ensures that the power used after perturbation will be the same as that used by the solution to (28). That constraint can be shown to be equivalent to the linear constraint $\mathbf{1}^T\mathbf{A}^{-1}(\boldsymbol{\sigma}_f \odot \boldsymbol{\delta_r}) = 0$, where $\boldsymbol{\delta_r}$ is the vector containing the scalars δ_{r_k}. Furthermore, the CDF $g(\cdot)$ can be well approximated by a quadratic curve; see Fig. 1.

With this approximation in place, the problem can be stated as the following convex problem in $\boldsymbol{\delta}_r$

$$\max_{\boldsymbol{\delta_r}} \quad \sum_{k=1}^K a_0(r^\star + \delta_{r_k})^2 + a_1(r^\star + \delta_{r_k}) + a_2 \tag{30a}$$

$$\text{s.t.} \quad \mathbf{1}^T\mathbf{A}^{-1}(\boldsymbol{\sigma}_f \odot \boldsymbol{\delta_r}) = 0, \tag{30b}$$

where a_0, a_1 and a_2 are the coefficients of the quadratic approximation of $g(r)$. If we let $\mathbf{b} = (\mathbf{1}^T\mathbf{A}^{-1})^T \odot \boldsymbol{\sigma}_f$, then using an analysis of the KKT conditions of (30), we can derive the dual variable ζ of the equality constraint as

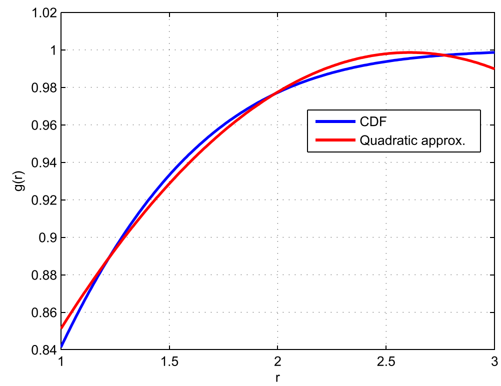

Fig. 1. The CDF of the standardized normal distribution, $\mathcal{N}(0,1)$, denoted $g(r)$, and its least squares quadratic approximation over $r \in [1,3]$.

$$\zeta = \frac{-(2a_0 r^\star + a_1)\mathbf{b}^H\mathbf{1}}{\mathbf{b}^H\mathbf{b}},$$

and the required $\boldsymbol{\delta_r}$ as

$$\boldsymbol{\delta_r} = \frac{-(2a_0 r^\star + a_1)\mathbf{1} - \zeta\mathbf{b}}{2a_0}.$$

Accordingly, whenever we have the optimal solution $\{\beta_k^\star\}_{k=1}^K$ and $r^\star$ of the problem in (28), we can calculate ζ, and the resulting perturbations of the offset coefficient $\boldsymbol{\delta_r}$. The modified offset coefficient vector $\mathbf{r}$ can be updated using $r_k = r^\star + \delta_{r_k}$. The power loading $\{\beta_k\}_{k=1}^K$ is then updated by using the linear equations arising from (28c) holding with equality; i.e., $\boldsymbol{\beta} = \mathbf{A}^{-1}\boldsymbol{\sigma}^2 + \mathbf{A}^{-1}(\boldsymbol{\sigma}_f \odot \mathbf{r})$.

VI. Simulation Results

In this section, we will provide three sets of numerical results. First, we will provide simulation results that show the validity of the offset-based algorithms and compare the performance of the algorithms presented here to that of zero-outage region algorithms that obtain robustness by ensuring that outage does not occur for uncertainties that lie in a given region. Specifically we will compare with the sphere bounding (SB) algorithm presented in [15]. Second, we will provide comparisons between the performance of the offset-based power loading algorithms proposed in Section V, the optimal power loading algorithm in [17], and the perturbation-based power loading algorithm that seeks to minimize the averaged outage, which was presented in Section V-C. In the third set of simulation results, we will demonstrate the performance gains that can be obtained by using the user rescheduling and the power saving described in Section V-B. We will also show the validity of the low-complexity approximations presented in Section IV-D.

For the initial simulation setup, we will we consider a downlink system in which a BS serves three single-antenna users. We will assume that the BS has four antennas, and the three users are randomly distributed within a radius of 1.6 km. The large scale fading is described by a path-loss exponent of 3.52 and log-normal shadow fading with 8 dB standard deviation, and the small scale fading is modelled using the standard i.i.d. Rayleigh model. The channel estimation error is assumed to be zero-mean and Gaussian with covariance $\sigma_{e_k}^2\mathbf{I}$, the power of which is -100 dBm. The receiver noise level is -90 dBm. A

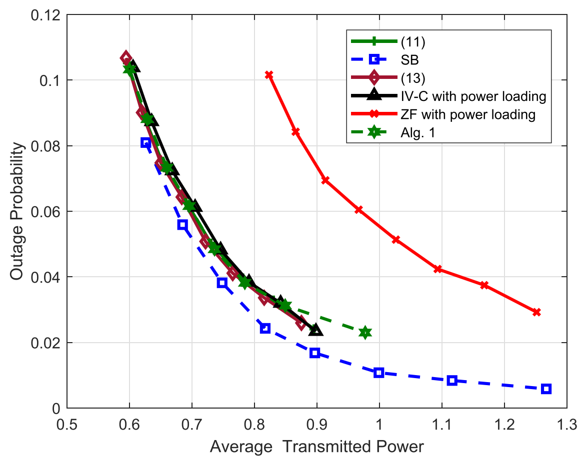

Fig. 2. The outage probability against the average transmitted power for a system with 3 users, 4 BS antennas, and $\gamma = 3$ dB.

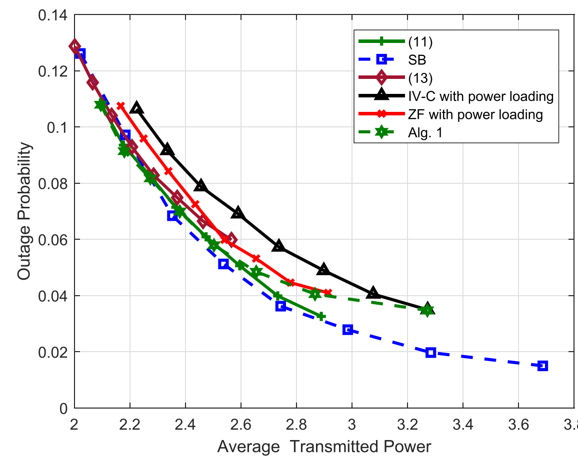

Fig. 3. The outage probability against the average transmitted power for a system with 3 users, 4 BS antennas, and $\gamma = 9$ dB.

simple channel-strength user selection technique is employed, where users are served only if $40\|\mathbf{h}_{e_k}\|^2/\sigma_k^2 \geq \gamma_k$, where we consider 40 here as the implicit total power constraint.

Each of the algorithms that we consider involves a choice of a robustness measure. For the algorithms provided in this paper the robustness measure is the value of the offset coefficient r_k. For the sphere bounding algorithm in [15] it is the size of the zero-outage region, and for the power loading algorithm in [17] it is directly the outage probability. To plot the performance curves, we randomly generate a set of channel realizations and provide the BS with estimates of those channels. Each algorithm is then used to design a set of beamformers that should provide the specified robustness. Using those beamformers we determine whether or not any user in the system with the actual channel realizations is in outage, and we calculate the corresponding transmission power. By repeating this experiment over thousands of channel realizations, we can plot the average outage probability over the users versus the average transmission power for the different algorithms when these algorithms provide a viable solution; by which we mean a solution that satisfies the constraints using a transmitted power that is less than 40. In fairness to all methods, the average is taken over those channel realizations for which all methods produce a viable solution.

In Fig. 2, assuming an SINR target of 3 dB, we plot the average outage probability versus the average total transmitted power for the proposed robust beamforming algorithms in (11), (13), Algorithm 1, and that of a system with the constant-offset directions described in Section. IV-C and the suggested power loading in Section V. (For the methods in (11) and (13), the beamformers were extracted from the optimal matrices $\mathbf{W}_k$ by taking the principal component.) As benchmarks, we plot the performance of the SB algorithm [15], and that of a system that employs the ZF directions combined with the power loading in Section V. In Fig. 3, we repeat the experiment for an SINR target of 9 dB. We observe that the performance gap between the proposed algorithms is small, which justifies the validity of our approximations. Indeed, the performance of the low-complexity robust beamforming algorithm in Algorithm 1 is very close to that of the original formulation in (11), and the SB algorithm (both of which incur a significantly larger computational load). The relative performance of the ZF-based algorithm with the proposed power loading algorithm in Section V depends on the SINR target, where comparatively better performance results are obtained when the SINR target is higher. That observation means that while both the offset-based beamforming directions and power loading contribute to the excellent performance for lower SINR targets, as the SINR target increases the role of the offset-based power loading becomes more significant. The performance of the combination of the original constant-offset directions in Section IV-C with the suggested power loading in Section V is not quite as good as that of the other offset-based approaches. However, decoupling the design of the beamforming directions and that of the power loading significantly reduces the computational cost (see Section IV-D), and greatly increases the flexibility of the design, as explained in Section V.

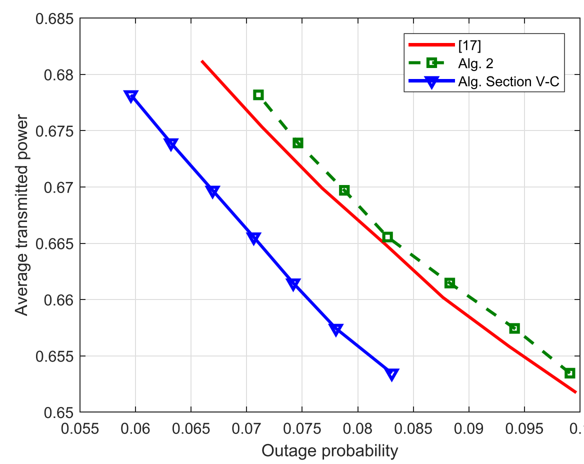

Fig. 4. The average transmitted power against the outage probability for a system with 3 users, 4 BS antennas, and $\gamma = 8$ dB.

The second set of simulation results examines the performance gap between the proposed power loading algorithms in Section V, and the power loading algorithm in [17] when the constant-offset directions are chosen; see Section. IV-C. In Fig. 4, we plot the average outage probability versus the average transmitted power for the power loading algorithm in [17], the power loading in Algorithm 2, and the modified power loading in Section V-C. (For the latter case, the quadratic approximation used in (30) is the least-squares approximation in Fig. 1.) While the algorithm in [17] is optimal in terms of the power required to achieve the specified outage probabilities for each user and

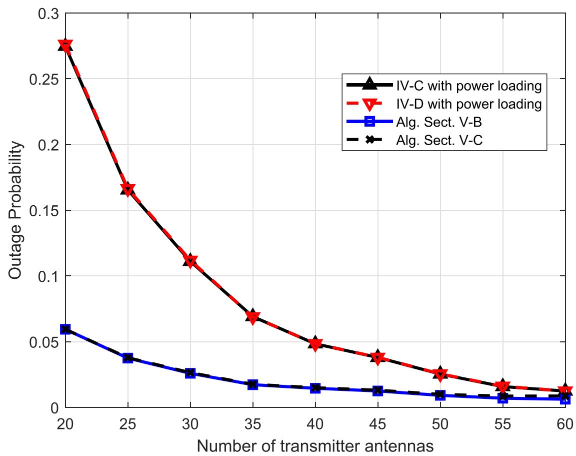

Fig. 5. The outage probability versus the number of antennas for a system with 8 users, and $\gamma = 9$ dB.

for each channel realization, Algorithm 2 provides similar performance at a much lower computational cost, and the modified power loading algorithm in Section V-C provides an even lower average outage probability than that obtained by Algorithm 2 or [17]. This performance is achieved while requiring no more than five iterations in the power loading algorithm in Algorithm 2.

To assess the performance gains that result from the power control capabilities of the proposed power loading algorithms, we plot the outage probability of the problem in (28) with the constant-offset directions in Section IV-C versus the number of antennas. In this case, we set the total power constraint $P_t = 1$, and the number of users to eight. We also plot the corresponding results when the approximations for obtaining the directions in Section IV-D, and those for obtaining the power loading in Section V-A are applied. In addition, we plot the performance of the proposed user rescheduling scheme (Algorithm Section V-B(a)) and the user rescheduling when combined with the power saving (Algorithm Section V-B(b)). We applied user rescheduling whenever the resulting offset r in (28) is smaller than two, and the rescheduled user(s) are considered to be in outage. For the power saving algorithm we upper bound r by 5. We observe from Fig. 5 that the proposed approximations provide almost the same outage performance over the whole range of antenna numbers. We also observe that the user rescheduling technique greatly enhances the outage performance, especially when the number of antennas is relatively low. (When the number of antennas is low, there is a greater probability of the channels not being sufficiently orthogonal.) Fig. 5 shows that the power saving algorithm (Algorithm Section V-B(b)) provides essentially the same performance as the regular algorithms, but significant power can be saved; the average actual transmitted powers used for that algorithm when the number of antennas are $[20, 25, \cdots, 60]$ are $[0.66, 0.58, 0.51, 0.45, 0.40, 0.37, 0.33, 0.30, 0.27]$ all of which are significantly smaller than the total power constraint $P_t = 1$.

VII. Conclusion

In this paper, a new offset-based approach is proposed for robust downlink beamforming. The approach is based on rewriting the SINR outage constraint as a non-negativity constraint on an indefinite quadratic function of the error in the base station's model of the channel. That non-negativity is then approximated by an offset-based constraint in which the mean of the function is required to be larger than a specific multiple of its standard deviation. This approach enabled the formulation of the robust beamforming design problem as a problem that can be transformed into a convex problem through the process of semidefinite relaxation (SDR). The computational complexity of the SDR problem can be further reduced when the uncertainty size is small, allowing for an iterative closed-form solution. When the beamforming directions are defined in advance, the offset-based approach generates a power loading algorithm that provides excellent performance and unique power control capabilities, while incurring only a small computational cost. The demonstrated performance gains, and the significant differences in computational cost exemplify the advantages of using the offset-based approach instead of the sphere bounding approach. Within the suite of algorithms generated by the offset based approach, the separation of the design into the constant-offset directions (Section IV-C) and the proposed power loading (Section V) provides a compelling balance between performance, computational cost and design flexibility.

Appendix A
Choice of r_k

From Cantelli's Inequality, we know that for any random variable $\mathbf{X}$ with mean μ_x and variance σ_x,

$$\text{Prob}(\mathbf{X} - \mu_x \le -r\sigma_x) \le \frac{1}{1+r^2}.$$

Therefore, if we ensure that $\mu_x \ge r\sigma_x$ then $\text{Prob}(\mathbf{X} \le 0) \le \frac{1}{1+r^2}$. Accordingly, if we set $r_k = \sqrt{1/\delta_k - 1}$, then, for any distribution of $\mathbf{e}_k$, the approximation in (7) is "safe", in the sense that any solution to the corresponding problem in (10) is guaranteed to satisfy the original outage constraints in (4).

In the case of Gaussian channel estimation errors, $f_k(\mathbf{e}_k)$ is an indefinite quadratic function of a Gaussian random vector, and considerable insight into its distribution is available; e.g., [21], [22]. For such a distribution, a direct application of Cantelli's Inequality yields a rather conservative choice of r_k. That means that there may be a significant number of scenarios for which the original problem in (4) is feasible, but for which the "safe" approximation in (10) is either infeasible, or requires an excessive transmission power.

In order to gain insight into more effective choices for r_k, we observe that the first and last terms in $f_k(\mathbf{e}_k)$ are constants and that the second term is Gaussian, since we are considering Gaussian estimation errors $\mathbf{e}_k$. Therefore, the term that complicates the calculation of the left tail probability (i.e., $\text{Prob}(f_k(\mathbf{e}_k) < 0)$, which is the outage probability) is the indefinite quadratic term $\mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k$. To have reasonable outage performance, the norm of the channel uncertainty $\mathbf{e}_k$ in (3) should be relatively small compared to the norm of the channel; cf., [23]. In that case, the first and second terms in (6), namely $\mathbf{h}_{e_k}^H \mathbf{Q}_k \mathbf{h}_{e_k}$ and $2\text{Re}(\mathbf{e}_k^H \mathbf{Q}_k \mathbf{h}_{e_k})$, will tend to dominate $\mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k$. Furthermore, since $\mathbf{Q}_k$ generically has one positive eigenvalue and $K-1$ negative eigenvalues, $\mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k$ may take on both positive and negative values. The fact that $\mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k$ will tend to be dominated by the other terms in (6), that it can take on positive and negative values, and that the other random term in (6) is Gaussian suggest that we could gain insight into effective choices for r_k by approximating $\mathbf{e}_k^H \mathbf{Q}_k \mathbf{e}_k$ by a

Gaussian random variable of the same mean and variance. Doing so generates a Gaussian approximation of $f_k(\mathbf{e}_k)$. (The examples in [35, Fig. 1] illustrate the viability of the approximation for the purpose of choosing r_k.) If $f_k(\mathbf{e}_k)$ were in fact Gaussian, then if the beamformers are designed such that $\mu_{f_k} \geq r_k\sigma_{f_k}$ then the outage probability would be $Q(r_k) = \frac{1}{2}\mathrm{erfc}(\frac{r_k}{\sqrt{2}})$, where $\mathrm{erfc}(\cdot)$ is the complementary error function. The analysis above suggests that an effective choice for r_k in our application would be close to $\sqrt{2}\,\mathrm{erfc}^{-1}(2\delta_k)$.

Appendix B
Mean and Variance Derivations

A Gaussian random variable $\mathbf{e}_k \backsim \mathcal{CN}(\mathbf{m}_k, \mathbf{C}_k)$ can be represented as $\mathbf{e}_k = \mathbf{m}_k + \mathbf{C}_k^{1/2}\hat{\mathbf{e}}_k$, where $\hat{\mathbf{e}}_k \backsim \mathcal{CN}(0, \mathbf{I})$. Using that representation we can write

$$\begin{aligned}
\mu_{f_k} &= \mathbb{E}\{f_k(\mathbf{e}_k)\} \\
&= (\mathbf{h}_{e_k} + \mathbf{m}_k)^H\mathbf{Q}_k(\mathbf{h}_{e_k} + \mathbf{m}_k) - \sigma_k^2 \\
&\quad + \mathbb{E}\{\hat{\mathbf{e}}_k^H\mathbf{C}_k^{1/2}\mathbf{Q}_k\mathbf{C}_k^{1/2}\hat{\mathbf{e}}_k\} \\
&= (\mathbf{h}_{e_k} + \mathbf{m}_k)^H\mathbf{Q}_k(\mathbf{h}_{e_k} + \mathbf{m}_k) - \sigma_k^2 + \mathbf{w}_k^H\mathbf{C}_k\mathbf{w}_k/\gamma_k \\
&\quad - \sum_{j\neq k}\mathbf{w}_j^H\mathbf{C}_k\mathbf{w}_j^H.
\end{aligned} \tag{31}$$

The variance can be expressed as

$$\begin{aligned}
\sigma_{f_k}^2 &= \mathrm{var}\{f_k(\mathbf{e}_k)\} \\
&= \mathrm{var}\{2\mathrm{Re}(\hat{\mathbf{e}}_k^H\mathbf{C}_k^{1/2}\mathbf{Q}_k(\mathbf{h}_{e_k} + \mathbf{m}_k)) \\
&\qquad + \hat{\mathbf{e}}_k^H\mathbf{C}_k^{1/2}\mathbf{Q}_k\mathbf{C}_k^{1/2}\hat{\mathbf{e}}_k\} \\
&= 2(\mathbf{h}_{e_k} + \mathbf{m}_k)^H\mathbf{C}_k^{1/2}\mathbf{Q}_k^2\mathbf{C}_k^{1/2}(\mathbf{h}_{e_k} + \mathbf{m}_k) \\
&\qquad + \mathrm{var}\{\hat{\mathbf{e}}_k^H\mathbf{C}_k^{1/2}\mathbf{Q}_k\mathbf{C}_k^{1/2}\hat{\mathbf{e}}_k\} + 0^* \\
&= 2(\mathbf{h}_{e_k} + \mathbf{m}_k)^H\mathbf{C}_k^{1/2}\mathbf{Q}_k^2\mathbf{C}_k^{1/2}(\mathbf{h}_{e_k} + \mathbf{m}_k) \\
&\qquad + \mathrm{tr}(\mathbf{C}_k^{1/2}\mathbf{Q}_k\mathbf{C}_k^{1/2})^2,
\end{aligned} \tag{32}$$

where tr denotes the trace function. At the point marked with the asterisk we have used the fact that the expectation of the cross terms is equal to zero. This is true because $\mathbb{E}\{2\mathrm{Re}(\mathbf{e}_k^H\mathbf{Q}_k(\mathbf{h}_{e_k} + \mathbf{m}_k))(\hat{\mathbf{e}}_k^H\mathbf{C}_k^{1/2}\mathbf{Q}_k\mathbf{C}_k^{1/2}\hat{\mathbf{e}}_k)\}$ consists of terms containing either similar or different components from the $\hat{\mathbf{e}}_k$ vector. Since $\hat{\mathbf{e}}_k$ has a zero mean, all terms with different indices will have a zero mean, while terms of similar indexes will take the form of a complex Gaussian raised to the power of three, which also has zero mean.

Appendix C
Fixed-Point Algorithm for ν_k

The fixed-point method for solving (22) that was proposed in [32] consists of "Jacobi-style" iterations. At the $(n+1)$th iteration, the values of $\{\nu_k\}$ are updated using

$$1/\nu_k^{n+1} := \mathbf{h}_{e_k}^H\left(\mathbf{I} + \sum_j \nu_j^n\mathbf{h}_{e_j}\mathbf{h}_{e_j}^H\right)^{-1}\mathbf{h}_{e_k}\left(1 + 1/\gamma_k\right).$$

The proposed method is based on an analytic solution to an equation that would rise in "Gauss-Seidel-style" iterations. In particular, we will solve the following equation

$$\begin{aligned}
1/\nu_k^{n+1} = \mathbf{h}_{e_k}^H\Bigg(\mathbf{I} &+ \sum_{j=1}^{k-1}\nu_j^{n+1}\mathbf{h}_{e_j}\mathbf{h}_{e_j}^H + \nu_k^{n+1}\mathbf{h}_{e_k}\mathbf{h}_{e_k}^H \\
&+ \sum_{j=k+1}^{K}\nu_j^n\mathbf{h}_{e_j}\mathbf{h}_{e_j}^H\Bigg)^{-1}\mathbf{h}_{e_k}\left(1 + 1/\gamma_k\right).
\end{aligned} \tag{33}$$

We can solve this equation efficiently for ν_k^{n+1} by using the fact that ν_k^{n+1} is the coefficient of a rank-one term in the matrix inversion. Indeed, by applying that fact and applying Sherman-Morrison-Woodbury formula, we can find ν_k^{n+1} using

$$\nu_k^{n+1} = \gamma_k/\mathbf{h}_{e_k}^H\mathbf{C}_k^n\mathbf{h}_{e_k} - \gamma_k\nu_k^n,$$

where

$$\mathbf{C}_k^n = \left(\mathbf{I} + \sum_{j=1}^{k-1}\nu_j^{n+1}\mathbf{h}_{e_j}\mathbf{h}_{e_j}^H + \sum_{j=k}^{K}\nu_j^n\mathbf{h}_{e_j}\mathbf{h}_{e_j}^H\right)^{-1}.$$

The matrix inversion can be updated recursively without matrix inversion using

$$\mathbf{C}_k^{n+1} = \mathbf{C}_k^n - \frac{(\nu_k^{n+1} - \nu_k^n)\mathbf{C}_k^n\mathbf{h}_{e_k}\mathbf{h}_{e_k}^H\mathbf{C}_k^n}{1 + (\nu_k^{n+1} - \nu_k^n)\mathbf{h}_{e_k}^H\mathbf{C}_k^n\mathbf{h}_{e_k}}.$$

This results in a computational cost per iteration of $\mathcal{O}(N_t^2K)$ operations, which is less than the $\mathcal{O}(N_t^3)$ operation per iteration required by the "Jacobi-style" method in [32].

References

[1] M. Alodeh *et al.*, "Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1733–1757, May 2018.

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

[3] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[4] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-reduction-aided broadcast precoding," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2057–2060, Dec. 2004.

[5] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1305–1316, Jul. 2004.

[6] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser MISO systems: A rate-splitting approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016.

[7] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[8] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.

[9] 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.

[10] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. Boca Raton, FL, USA: CRC Press, 2001, ch. 18.

[11] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[12] 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, Jul. 2014.

[13] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, Jan. 2002.

[14] M. B. Shenouda and T. N. Davidson, "Probabilistically-constrained approaches to the design of the multiple antenna downlink," in *Proc. 42nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2008, pp. 1120–1124.

[15] 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.

[16] 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, Jun. 2016.

[17] F. Sohrabi 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, Jun. 2016.

[18] A. Pascual-Iserte, D. Palomar, A. Perez-Neira, and M. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 346–360, Jan. 2006.

[19] N. Vucic and H. Boche, "A tractable method for chance-constrained power control in downlink multiuser MISO systems with channel uncertainty," *IEEE Signal Process. Lett.*, vol. 16, no. 5, pp. 346–349, May 2009.

[20] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[21] 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.

[22] D. Raphaeli, "Distribution of noncentral indefinite quadratic forms in complex normal variables," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 1002–1007, May 1996.

[23] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[24] G. Scutari, F. Facchinei, and L. Lampariello, "Parallel and distributed methods for constrained nonconvex optimization—Part I: Theory," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 1929–1944, Apr. 2017.

[25] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Kornar, and N. D. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 804–808, Jul. 2015.

[26] L.-N. Tran, M. F. Hanif, and M. Juntti, "A conic quadratic programming approach to physical layer multicasting for large-scale antenna arrays," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 114–117, Jan. 2014.

[27] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," *version 2.1*, Mar. 2014, http://cvxr.com/cvx

[28] E. Song, Q. Shi, M. Sanjabi, R.-Y. Sun, and Z.-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?" *EURASIP J. Wireless Commun. Netw.*, vol. 1, pp. 1–11, 2012.

[29] W. K. Ma, J. Pan, A. M. C. So, and T. H. Chang, "Unraveling the rank-one solution mystery of robust MISO downlink transmit optimization: A verifiable sufficient condition via a new duality result," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1909–1924, Apr. 2017.

[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.

[31] 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.

[32] 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.

[33] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[34] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Topics Signal Process.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[35] M. Medra and T. N. Davidson, "Per-user outage-constrained power loading technique for robust MISO downlink," in *Proc. 49th Asilomar Conf. Signals, Syst., Comput.*, Nov. 2015, pp. 1232–1236.

Mostafa Medra (S'06–M'16) received the B.Sc. and M.Sc. degrees in electrical engineering from Alexandria University, Alexandria, Egypt, in 2009 and 2013, respectively, and the Ph.D. degree from McMaster University, Hamilton, ON, Canada, in 2017. From 2012 to 2013, he held a research position with the Spirtonic research team, working on digital signal processing for non-destructive testing using ultrasonic waves. He is currently a Postdoctoral Researcher with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. His current research interests include MIMO communications, optimization, wireless communications, and signal processing.

Yongwei Huang (M'09–SM'16) received the Bachelor of Science degree in information and computational science, the Master of Science degree in operations research from Chongqing University, Chongqing, China, in 1998 and 2000, respectively, and the Ph.D. degree in systems engineering and engineering management from the Chinese University of Hong Kong, Hong Kong, in 2005.

In 2016, he joined the School of Information Engineering, Guangdong University of Technology, Guangzhou, China, where he is currently a Professor. Prior to the position, he held a faculty position with the Hong Kong Baptist University, and several research appointments, namely, with the Hong Kong University of Science and Technology, the Chinese University of Hong Kong, and the University of Naples "Federico II," Italy. His research interests are related to optimization theory and algorithms, including conic optimization, robust optimization, and combinatorial optimization, and their applications in signal processing for wireless communications and radar.

He was a recipient of the Best Poster Award in the 2007 Workshop on Optimization and Signal Processing. He is an Associate Editor for the IEEE Signal Processing Letters, and a member of Sensor Array and Multichannel Signal Processing Technical Committee of the IEEE Signal Processing Society.

Timothy N. Davidson (M'96–SM'15) received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, Australia, and the D.Phil. degree in engineering science from the University of Oxford, Oxford, U.K., in 1991 and 1995, respectively.

He is a Professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he is currently serving as the Chair of the Department. Previously, he has served as an Acting Director of the School of Computational Engineering and Science for two years, and as an Associate Director for three years. His research interests lie in the general areas of communications, signal processing, and control.

He was the recipient of the 1991 J. A. Wood Memorial Prize from UWA, the 1991 Rhodes Scholarship for Western Australia, and the 2011 Best Paper Award from the IEEE Signal Processing Society. He was an Associate Editor for the IEEE Transactions on Signal Processing, the IEEE Transactions on Wireless Communications, and the IEEE Transactions on Circuits and Systems II. He was also a Guest Co-Editor of issues of the IEEE Journal on Selected Areas in Communications, the IEEE Journal of Selected Topics in Signal Processing, and the *EURASIP Journal on Advances in Signal Processing*. He was a General Co-Chair for the 2014 IEEE International Workshop on Signal Processing Advances in Wireless Communications, a Technical Program Co-Chair for the 2014 IEEE Global Conference on Signal and Information Processing, and the Technical Chair for the 2015 Asilomar Conference on Signals, Systems and Computers. He was also the Chair of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking. He is a Registered Professional Engineer in the Province of Ontario.