Robust Downlink Beamforming With Partial Channel State Information for Conventional and Cognitive Radio Networks

Imran Wajid, Marius Pesavento, *Member, IEEE*, Yonina C. Eldar, *Fellow, IEEE*, and Dana Ciochina, *Student Member, IEEE*

***Abstract*—We address the problem of robust multiuser downlink beamforming under the assumption that the transmitter has partial covariance-based channel state information (CSI). In our approach, the uncertainty on the channel covariance matrices is assumed to be confined in an ellipsoid of given size and shape, where prior knowledge about the statistical distribution of the CSI mismatch is taken into account. The goal is to minimize the transmitted power under the worst-case quality-of-service (QoS) constraints. We extend the developed robust problem to downlink beamforming in cognitive radio (CR) networks where QoS constraints apply to the users of the secondary network (SN) and interference leaked to the primary users (PUs) is required to be below a given interference threshold. We avoid the coarse approximations used by previous solutions and obtain exact reformulations for both worst-case problems based on Lagrange duality. The resulting problems can then be approximated using semidefinite relaxation (SDR). Further, we consider a popular alternative robust approach that is based on probabilistic QoS and interference constraints and show that both approaches are generally equivalent. Computer simulations show that the proposed techniques provide substantial performance improvements over earlier robust downlink beamforming techniques for both the conventional and the CR scenarios.**

***Index Terms*—Downlink beamforming, cognitive radio, convex optimization, user quality-of-service.**

I. Introduction

THE increasing demand for wireless services has urged researchers to seek an efficient way of utilizing the available radio spectrum. The reliability and throughput of wireless communication systems can be improved by employing multiple antennas at the transmitter and/or receivers which helps in efficiently controlling multiuser interference by exploiting the spatial domain. With the introduction of advanced multiantenna standards, like WiMAX, LTE and LTE-A and efficient multiantenna base stations, multiuser downlink beamforming has become an active area of research in recent years (see [1] and the references therein). The performance improvements brought by the use of downlink beamforming can be realized if accurate channel state information (CSI) is available at the transmitter. In such a scenario, using precoding at the multiantenna transmitters, the beam-pattern can be adjusted to mitigate multiuser interference and improve user quality-of-service (QoS) [2].

Several beamforming techniques have been developed assuming perfect CSI. In the scheme of [2], the availability of instantaneous CSI at the transmitter is considered while the techniques in [3] and [4] rely on perfect covariance-based CSI at the transmitter. In both cases the assumption of exact CSI is, however, not practical. In real cellular systems, the CSI is either estimated at the receiver using training sequences and then signaled back to the transmitter via feedback channels or, alternatively, the CSI may be determined in the uplink, i.e., in time division duplexing (TDD) systems, assuming channel reciprocity. Estimation errors are inevitable in these cases [5]. Due to the limited capacity of the feedback channels [6], CSI is also prone to quantization errors. Furthermore, due to latencies in channel state feedback and short channel coherence time, CSI at the transmitter may be imprecise. Since the methods developed assuming perfect CSI are quite sensitive to such channel uncertainties, several recent works have focused on the design of downlink beamforming techniques that are robust to CSI errors; see [7]–[9] for instantaneous CSI based approaches and [10]–[12] for methods utilizing covariance-based CSI. Since the second order statistics of the channel change slowly compared to the channel itself, the feedback requirements for covariance-based CSI are significantly lower compared to instantaneous channel feedback. Therefore, the use of covariance based CSI is generally more practical, especially in fast fading channels.

The aforementioned robust approaches are based on worst-case error considerations. These methods assume that the CSI errors are within an ellipsoid of a given size without making any assumption on the distribution of the error. The goal then is to design beamformers that satisfy the constraints for all possible errors within this ellipsoid. Another approach for developing robust techniques, used in both transmit and receive beamforming, is based on the design of the outage probabilities [13]–[18]. In this case, the distribution of the CSI errors is assumed to be known and the worst-case constraints are replaced by probabilistic constraints. The beamformers are designed such that the outage probability of the constraints is below a predefined value.

Manuscript received October 26, 2012; revised March 21, 2013; accepted May 11, 2013. Date of publication May 30, 2013; date of current version June 25, 2013. The work of I. Wajid was supported by the Higher Education Commission of Pakistan. The work of I. Wajid, M. Pesavento and Dana Ciochina has been supported in parts by the European Commission under FP7 Project CROWN, and the European Research Council (ERC) Advanced Investigator Grants Program under Grant 227477-ROSE. The work of Y. C Eldar was supported by the European Commission FP7 Network of Excellence in Wireless COMmunications NEWCOM++.

I. Wajid, M. Pesavento, and D. Ciochina are with the Technische Universität Darmstadt, D-64283 Darmstadt, Germany (e-mail: pesavento@nt.tu-darmstadt.de).

Y. C. Eldar is with the Technion—Israel Institute of Technology, 32000 Haifa, Israel.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2265682

Providing reliable indoor service coverage is a major challenge in current cellular systems. In modern state of the art wireless standards like UMTS, LTE and LTE-A, femtocells are deployed in homes, offices and buildings to improve the indoor service coverage; see [19], [20] and references therein. A femtocell base station may use the same frequency band as the surrounding macro network. One major problem with the deployment of femtocells is that, unlike macrocells, careful cell planning cannot be carried out. Therefore, there exists a strong demand to control interference caused by femtocells so that macrocell users are not adversely affected. In this setup the macrocell is treated as being privileged over the femtocell, so that an underlay cognitive radio (CR) framework can be adopted where the femtocell acts as the secondary network (SN) in which the secondary users (SUs) are served [21]. The macrocell network refers to the primary network (PN) and its users are the primary users (PUs) that according to the underlay paradigm are assumed to be non-cooperative. The use of multiple antennas at the transmitter of the SN, enables spatial processing to enhance QoS at the SN users and to control interference at the PN users. From a mathematical point of view, the CR beamforming problem can be viewed as a conventional beamforming problem with additional constraints on the interference to the PUs [21]–[23]. Several CR beamforming techniques have been developed which assume the availability of perfect CSI, see [23]–[26]. Methods considering erroneous CSI are considered in [27]–[32]. The authors of [27] extended the technique of [10] to the CR beamforming problem. However, there are several approximations used in both [10] and [27]. In particular, conservative modifications of the QoS and PU interference constraints are applied that use some coarse approximations. Moreover, these approaches ignore the positive semidefinitness constraints on the mismatched downlink channel covariance matrices. In [12], an improved conventional downlink beamforming approach is presented that avoids the conservative modifications used in [10]. The techniques of [28] and [31] extend the idea of [12] for the CR beamforming case. The method of [28], however, uses a conservative design that does not consider the positive semidefinitness constraints on the mismatched covariance matrices.

In this paper, we propose new worst-case optimization-based robust techniques for both the conventional and CR beamforming problems. In our approach, we use the weighted Frobenius norm to bound the channel uncertainties instead of the conventional Frobenius norm as in practical systems the covariance matrix errors follow specific distributions that can, e.g., be estimated at the transmitter or the receivers. This can lead to significant improvements in practical implementations. We also avoid the aforementioned approximations used in [10], [27] and [28]. In particular, our formulations explicitly take into account the positive-semidefinitness property of the downlink covariance matrices. We use Lagrange duality to reformulate the QoS and PU interference constraints exactly. We then apply semidefinite relaxation (SDR) to convert the resulting problems into convex semidefinite programming (SDP) problems. We further consider an alternative robust beamforming approach using the popular probabilistic model. Interestingly, after some manipulations and simplifications, we show that the final problem formulations for both the worst-case and probabilistic approaches are mathematically equivalent. A similar relationship has been derived for receive beamforming problems in [18].

In our simulations, the solutions to the resulting SDP problems are generally rank-one which implies that the obtained SDR solutions are optimal. In this paper, we provide a class of specific examples, with large uncertainty thresholds and/or high symmetries among users channel covariance matrices, in which all solutions violate the rank-one constraint. However, by analyzing these examples we conclude that in randomly fading channels and under reasonable assumptions regarding the size of the uncertainty sets, the probability that these cases occur is comparably low. This analysis provides a new insight to the robust beamforming problem, as this result is fundamentally different from the non-robust case where for a small number of interference constraints (≤ 2) rank one solutions can always be obtained [33]. We also demonstrate that, concerning problem feasibility and transmitted power, the proposed beamformers obtain substantial performance improvements in comparison to those of [10] and [27].

The rest of the paper is organized as follows. We present the general system model and problem formulations for the non-robust downlink beamformers in both the conventional and CR scenarios in Section II. In Section III, robust problem formulations for the downlink beamformers are derived and compared to the approaches of [10] and [27]. In Section IV, the proposed robust approaches for the downlink beamforming problems are developed that are free from any coarse approximations. In Section V, we develop a technique based on the outage probability and show its relationship to the proposed worst-case based approach. Then, in Section VI, we discuss a class of special cases in which higher rank solutions are obtained from the SDR technique to gain further insight into the SDR solution. Finally, simulation results are presented in Section VII that show improved performance of our techniques compared to earlier methods.

II. Beamforming for Conventional and CR Scenarios

In this section, we present the non-robust CR downlink beamforming problem that has been addressed in [27]. This problem can be viewed as an extension of conventional downlink beamforming [4] for the operation of a SN base station, where additional PU interference constraints apply.

Consider a CR network where the SN consists of a single base station with N transmit antennas and K single antenna SUs. The corresponding PN is assumed to comprise L single antenna PUs. At time instance t, the SN base station transmits the $N \times 1$ vector

$$\boldsymbol{x}(t) = \sum_{i=1}^{K} s_i(t)\boldsymbol{w}_i \tag{1}$$

where $s_i(t)$ and $\boldsymbol{w}_i$ are the intended signal and the $N \times 1$ complex weight vector for the ith SU, respectively. The signal received by the kth SU is given by

$$y_k(t) = \boldsymbol{h}_k^H \boldsymbol{x}(t) + n_k(t) \tag{2}$$

where $\boldsymbol{h}_k$ is the $N \times 1$ channel vector between the SN base station and the kth SU and $(\cdot)^H$ is the Hermitian transpose. The noise of the kth SU, denoted $n_k(t)$, is assumed to be zero-mean circularly symmetric white Gaussian with variance σ_k^2.

The received SINR of the kth SU can be expressed as [10]

$$\mathrm{SINR}_k = \frac{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k}{\sum_{\substack{i=1\\ i\neq k}}^{K} \boldsymbol{w}_i^H \boldsymbol{R}_k \boldsymbol{w}_i + \sigma_k^2},\ k = 1,\ldots,K \tag{3}$$

where $\boldsymbol{R}_k \triangleq \mathrm{E}\{\boldsymbol{h}_k \boldsymbol{h}_k^H\}$ is the downlink channel covariance matrix for the kth SU. The transmissions from the SN base station cause interference at the PUs. The interference at the lth PU can be written as

$$\mathcal{I}_l = \sum_{i=1}^{K} \boldsymbol{w}_i^H \boldsymbol{R}_{K+l} \boldsymbol{w}_i,\ l = 1,\ldots,L \tag{4}$$

where $\boldsymbol{R}_{K+l} \triangleq \mathrm{E}\{\boldsymbol{h}_{K+l} \boldsymbol{h}_{K+l}^H\}$ is the channel covariance matrix between the SN base station and the lth PU, and $\boldsymbol{h}_{K+l}$ is the corresponding $N \times 1$ channel vector.

The non-robust CR beamformer design problem in [27] is formulated under the assumption that the channel covariance matrices $\boldsymbol{R}_1,\ldots,\boldsymbol{R}_{K+L}$ are perfectly known. The downlink beamformers are then designed to minimize the total transmitted power at the SN base station, subject to individual QoS constraints for the SUs, while forcing the interference to the PUs to fall below a predefined maximum. The resulting beamforming problem can be written as

$$\begin{aligned} \min_{\{\boldsymbol{w}_i\}} \quad & \sum_{i=1}^{K} \|\boldsymbol{w}_i\|^2 \\ \text{s.t.} \quad & \frac{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k}{\sum_{\substack{i=1\\ i\neq k}}^{K} \boldsymbol{w}_i^H \boldsymbol{R}_k \boldsymbol{w}_i + \sigma_k^2} \geq \gamma_k;\ k = 1,\ldots,K \\ & \sum_{i=1}^{K} \boldsymbol{w}_i^H \boldsymbol{R}_{K+l} \boldsymbol{w}_i \leq \varepsilon_l;\ l = 1,\ldots,L \end{aligned} \tag{5}$$

where $\|\cdot\|$ denotes the Euclidean norm of a vector or the Frobenius norm of a matrix, γ_k is the minimal acceptable SINR of the kth SU, and ε_l is the maximum allowable interference power caused by the SN base station at the lth PU. Note that the beamformers obtained from problem (5) are matched to the second order statistics of the channel vectors and not to their instantaneous values. Additionally, without the PU interference constraints (i.e., for $L = 0$), the CR downlink beamforming problem (5) reduces to the conventional beamforming problem considered in [4].

To solve (5), the authors of [27] extend the SDR approach developed in [4] by introducing the variable transformation

$$\boldsymbol{W}_k \triangleq \boldsymbol{w}_k \boldsymbol{w}_k^H;\ k = 1,\ldots,K \tag{6}$$

and relaxing the non-convex constraints $\mathrm{rank}\{\boldsymbol{W}_k\} = 1$ ($k = 1,\ldots,K$). As a result the original problem (5) is relaxed to the following convex form:

$$\begin{aligned} \min_{\{\boldsymbol{W}_i\}} \quad & \sum_{i=1}^{K} \mathrm{Tr}\{\boldsymbol{W}_i\} \\ \text{s.t.} \quad & \gamma_k \sum_{\substack{i=1\\ i\neq k}}^{K} \mathrm{Tr}\{\boldsymbol{R}_k \boldsymbol{W}_i\} - \mathrm{Tr}\{\boldsymbol{R}_k \boldsymbol{W}_k\} + \gamma_k \sigma_k^2 \leq 0;\ \boldsymbol{W}_k \succeq 0 \\ & \sum_{i=1}^{K} \mathrm{Tr}\{\boldsymbol{R}_{K+l} \boldsymbol{W}_i\} \leq \varepsilon_l;\ k=1,\ldots,K;\ l=1,\ldots,L \end{aligned} \tag{7}$$

where $\mathrm{Tr}\{\cdot\}$ denotes the trace of a matrix. It has been shown in [10], for conventional downlink beamforming ($L = 0$), that the semidefinite program (7) always has at least one rank-one solution for $\boldsymbol{W}_k$. For the case of CR downlink beamforming, the authors of [33] showed that when $L \leq 2$, a rank-one solution of (7) can always be constructed from higher rank solutions. When a rank-one $\boldsymbol{W}_k$ is obtained from problem (7), problems (5) and (7) become equivalent and, therefore, the optimum beamforming vectors $\boldsymbol{w}_k$ can be obtained from the principal eigenvectors of $\boldsymbol{W}_k$.

III. The Robust Beamforming Problem for Conventional and CR Scenarios

In practical scenarios, the CSI available at the base station is often subject to uncertainty. Therefore, we consider a robust extension of the conventional and CR downlink beamforming problems. As the true covariance matrices $\boldsymbol{R}_k$ and $\boldsymbol{R}_{K+l}$ are unknown, the beamformers generated by solving (5) may not satisfy the constraints in (5) when applied to the true covariance matrices. We model the estimates of the true channel covariance matrices, available at the transmitter, for the kth SU and lth PU as $\hat{\boldsymbol{R}}_k$ ($k = 1,\ldots,K$) and $\hat{\boldsymbol{R}}_{K+l}$ ($l = 1,\ldots,L$), respectively. Let the $N \times N$ Hermitian matrices $\boldsymbol{\Delta}_k$ and $\boldsymbol{\Delta}_{K+l}$ model the corresponding uncertainties in these estimates. Depending on the CSI estimation methods or the feedback quantization scheme, the uncertainty matrices $\boldsymbol{\Delta}_k$ and $\boldsymbol{\Delta}_{K+l}$ follow specific random distributions. We consider the case that the channel covariance matrices are subject to colored noise. Let $\mathrm{vec}(\boldsymbol{X})$ denote the operator that stacks the columns of the matrix $\boldsymbol{X}$ to form a vector. We assume that

$$\boldsymbol{R}_{\boldsymbol{\Delta}_m} \triangleq \mathrm{E}\{\mathrm{vec}(\boldsymbol{\Delta}_m)\mathrm{vec}(\boldsymbol{\Delta}_m)^H\};\ m = 1,\ldots,K+L, \tag{8}$$

denotes the colored covariance matrix of the CSI error vector $\mathrm{vec}(\boldsymbol{\Delta}_m)$. Assume that $\boldsymbol{u}_n(\boldsymbol{R}_{\boldsymbol{\Delta}_m})$ is the eigenvector of $\boldsymbol{R}_{\boldsymbol{\Delta}_m}$ with the corresponding eigenvalue $\lambda_n(\boldsymbol{R}_{\boldsymbol{\Delta}_m})$. Due to the coloring of the CSI errors, the errors along the dominant eigenvectors corresponding to the largest eigenvalues are more prominent than the errors along the minor eigenvectors. To take the effect of the colored CSI errors into account the uncertainty set is appropriately described by an ellipsoid with elliptic radii proportional to the square root of the eigenvalues $\lambda_n(\boldsymbol{R}_{\boldsymbol{\Delta}_m})$, as specified below.

Consider e.g., the weighting matrix defined as

$$\boldsymbol{Q}_m \triangleq \frac{\boldsymbol{R}_{\boldsymbol{\Delta}_m}}{\eta_m};\ m = 1,\ldots,K+L \tag{9}$$

where η_m is a constant such that $\|\boldsymbol{Q}_m\| = 1$. Then we define the $\boldsymbol{P}$-weighted Frobenius norm, for some positive-semidefinite matrix $\boldsymbol{P}$ as

$$\|\boldsymbol{X}\|_{\boldsymbol{P}} \triangleq \|\sqrt{\boldsymbol{P}}\mathrm{vec}(\boldsymbol{X})\| = \sqrt{\mathrm{vec}(\boldsymbol{X})^H \boldsymbol{P} \mathrm{vec}(\boldsymbol{X})}. \tag{10}$$

In our beamforming approach we ensure robustness for all mismatch matrices that are bounded by ellipsoids of known elliptic radii, i.e., $\|\boldsymbol{\Delta}_k\|_{\boldsymbol{Q}_k^{-1}} \leq \alpha_k$ and $\|\boldsymbol{\Delta}_{K+l}\|_{\boldsymbol{Q}_{K+l}^{-1}} \leq \alpha_{K+l}$. We remark that if $\lambda_n(\boldsymbol{Q}_k)$ denotes the nth eigenvalue of the weighting matrix $\boldsymbol{Q}_k$ and $\|\boldsymbol{\Delta}_k\|_{\boldsymbol{Q}_k^{-1}} \leq \alpha_k$, then the nth elliptic radius is given by $\sqrt{\lambda_n(\boldsymbol{Q}_k)}\alpha_k$. In this sense, we consider the cases in which the true channel covariance matrices lie in the set of matrices defined as $\{\hat{\boldsymbol{R}}_k+\boldsymbol{\Delta}_k : \|\boldsymbol{\Delta}_k\|_{\boldsymbol{Q}_k^{-1}} \le \alpha_k\}$ $(k=1,\ldots,K)$ and $\{\hat{\boldsymbol{R}}_{K+l}+\boldsymbol{\Delta}_{K+l} : \|\boldsymbol{\Delta}_{K+l}\|_{\boldsymbol{Q}_{K+l}^{-1}} \le \alpha_{K+l}\}$ $(l=1,\ldots,L)$. In case of uniform weighting, hence $\boldsymbol{Q}_m=\boldsymbol{I}$, this approach of modeling the uncertainties becomes similar to the non-weighted techniques of [10] and [27].

Next, we follow a worst-case approach to design the downlink beamformers such that the QoS and interference targets at the SUs and PUs, respectively, are met for all mismatched covariance matrices. Taking into account that the true covariances are positive-semidefinite, the robust CR beamforming problem corresponding to (5) can be formulated as

$$\begin{aligned}
\min_{\{\boldsymbol{w}_i\}} \quad & \sum_{i=1}^{K}\|\boldsymbol{w}_i\|^2 \\
\text{s.t.} \quad & \min_{\substack{\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}}\le\alpha_k \\ \hat{\boldsymbol{R}}_k+\bar{\boldsymbol{\Delta}}_k\succeq 0}} \frac{\boldsymbol{w}_k^H(\hat{\boldsymbol{R}}_k+\bar{\boldsymbol{\Delta}}_k)\boldsymbol{w}_k}{\sum_{\substack{i=1\\i\neq k}}^{K}\boldsymbol{w}_i^H(\hat{\boldsymbol{R}}_k+\bar{\boldsymbol{\Delta}}_k)\boldsymbol{w}_i+\sigma_k^2} \ge \gamma_k; \\
& \hspace{20em} k=1,\ldots,K \\
& \max_{\substack{\|\bar{\boldsymbol{\Delta}}_{K+l}\|_{\boldsymbol{Q}_{K+l}^{-1}}\le\alpha_{K+l} \\ \hat{\boldsymbol{R}}_{K+l}+\bar{\boldsymbol{\Delta}}_{K+l}\succeq 0}} \sum_{i=1}^{K}\boldsymbol{w}_i^H(\hat{\boldsymbol{R}}_{K+l}+\bar{\boldsymbol{\Delta}}_{K+l})\boldsymbol{w}_i \le \varepsilon_l; \\
& \hspace{20em} l=1,\ldots,L. \qquad (11)
\end{aligned}$$

The robust modifications of the conventional beamformer in [10] and CR beamformer in [27] are similar to (11) with $\boldsymbol{Q}_m=\boldsymbol{I}$ $(m=1,\ldots,K+L)$, i.e., the Frobenius norm is used instead of the weighted Frobenius norm. However, in contrast to (11), the positive definiteness constraints on the mismatched covariance matrices have been ignored in [10] and [27], which could in fact result in worst-case approaches that are unnecessarily conservative. The robust beamforming solutions proposed in [10] and [27] involve several additional conservative approximations that adversely affect the performance of the beamformers. In particular, the worst-case QoS constraints of (11) are replaced in [10] and [27] by

$$\frac{\min_{\|\bar{\boldsymbol{\Delta}}_k\|\le\alpha_k}\boldsymbol{w}_k^H(\hat{\boldsymbol{R}}_k+\bar{\boldsymbol{\Delta}}_k)\boldsymbol{w}_k}{\sum_{\substack{i=1\\i\neq k}}^{K}\max_{\|\bar{\boldsymbol{\Delta}}_k\|\le\alpha_k}\boldsymbol{w}_i^H(\hat{\boldsymbol{R}}_k+\bar{\boldsymbol{\Delta}}_k)\boldsymbol{w}_i+\sigma_k^2} \ge \gamma_k. \qquad (12)$$

The approximation (12) results in strengthening of the QoS constraints in (11) and, therefore, when the constraints (12) are satisfied, it is guaranteed that the QoS constraints in (11) will also be satisfied. The reverse statement is however not true. Similarly, the modification of the PU interference constraints in [27] is also unnecessarily strict. These constraints are replaced by

$$\sum_{i=1}^{K}\max_{\|\bar{\boldsymbol{\Delta}}_{K+l}\|\le\alpha_{K+l}}\boldsymbol{w}_i^H(\hat{\boldsymbol{R}}_{K+l}+\bar{\boldsymbol{\Delta}}_{K+l})\boldsymbol{w}_i \le \varepsilon_l. \qquad (13)$$

Both approximations may lead to suboptimal beamforming solutions or even infeasibility of the strengthened problem.

Replacing the QoS and PU interference constraints in (11) by (12) and (13), respectively, the authors of [27] reformulate the robust CR beamforming problem as

$$\begin{aligned}
\min_{\{\boldsymbol{w}_i\}} \quad & \sum_{i=1}^{K}\|\boldsymbol{w}_i\|^2 \\
\text{s.t.} \quad & \frac{\boldsymbol{w}_k^H(\hat{\boldsymbol{R}}_k-\alpha_k\boldsymbol{I})\boldsymbol{w}_k}{\sum_{\substack{i=1\\i\neq k}}^{K}\boldsymbol{w}_i^H(\hat{\boldsymbol{R}}_k+\alpha_k\boldsymbol{I})\boldsymbol{w}_i+\sigma_k^2} \ge \gamma_k;\ k=1,\ldots,K \\
& \sum_{i=1}^{K}\boldsymbol{w}_i^H(\hat{\boldsymbol{R}}_{K+l}+\alpha_{K+l}\boldsymbol{I})\boldsymbol{w}_i \le \varepsilon_l;\ l=1,\ldots,L. \qquad (14)
\end{aligned}$$

The latter problem is mathematically similar to (5) and, therefore, can be solved in the same way, i.e., using SDR. For the conventional beamformer case $(L=0)$, problem (14) reduces to the robust modification presented in [10].

An improved robust downlink beamforming approach for the conventional scenario has been presented in [12], that takes into account the positive semidefinitness constraints on the mismatched covariance matrices and avoids the above mentioned conservative modifications of QoS constraints. For the CR case, the authors of [28] extend the idea of [12], however, without including the positive semidefinitness constraints on the mismatched covariance matrices, which, in contrast were considered in our approach [31].

In the next sections, we develop an alternative approach to solve the robust worst-case beamforming problems for both the conventional and CR scenarios. Our technique avoids the aforementioned conservative approximations and, hence, offers a robust beamformer design with improved feasibility and reduced transmitted power. In our approach, we use Lagrange duality, similar to the approaches of [12], [28] and [31], to solve the inner minimizations, appearing in the constraints of (11), exactly, taking the positive semidefinitness of the downlink covariance matrices into account. As we later show in simulations, this results in an improved performance in terms of feasibility and transmitted power.

IV. The Proposed Robust Beamformers

In this section we provide an exact reformulation of the robust downlink beamforming problem (11).

A. The Proposed Conventional Downlink Beamformer

We first consider the conventional beamforming problem and reformulate the QoS constraints in (11) as a separate optimization problem. Introducing the compact matrix notation

$$\boldsymbol{A}_k \triangleq \gamma_k\sum_{i=1,\ i\neq k}^{K}\boldsymbol{w}_i\boldsymbol{w}_i^H-\boldsymbol{w}_k\boldsymbol{w}_k^H \qquad (15)$$

the kth QoS constraint can be written as

$$\min_{\substack{\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}}\le\alpha_k \\ \hat{\boldsymbol{R}}_k+\bar{\boldsymbol{\Delta}}_k\succeq 0}} -\left(\mathrm{Tr}\{\bar{\boldsymbol{\Delta}}_k\boldsymbol{A}_k\}+\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\}+\sigma_k^2\gamma_k\right)\ge 0. \quad (16)$$

The minimization on the left side of (16) corresponds to the following optimization problem:

$$\begin{aligned}\min_{\bar{\boldsymbol{\Delta}}_k} \quad & -\left(\mathrm{Tr}\{\bar{\boldsymbol{\Delta}}_k \boldsymbol{A}_k\} + \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \boldsymbol{A}_k\} + \sigma_k^2 \gamma_k\right) \\ \text{s.t.} \quad & \|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}} \le \alpha_k;\ \hat{\boldsymbol{R}}_k + \bar{\boldsymbol{\Delta}}_k \succeq 0. \end{aligned} \tag{17}$$

For a given matrix $\boldsymbol{A}_k$ in (15), problem (17) is convex in $\bar{\boldsymbol{\Delta}}_k$, and can be replaced by its Lagrange dual. The Lagrange dual function for problem (17) can be written as

$$\begin{aligned} g(\lambda_k, \boldsymbol{Z}_k) &= \inf_{\bar{\boldsymbol{\Delta}}_k} \Big\{ -\mathrm{Tr}\{\bar{\boldsymbol{\Delta}}_k \boldsymbol{A}_k\} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \boldsymbol{A}_k\} - \sigma_k^2 \gamma_k \\ &\qquad + \lambda_k (\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}} - \alpha_k) - \mathrm{Tr}\{(\hat{\boldsymbol{R}}_k + \bar{\boldsymbol{\Delta}}_k)\boldsymbol{Z}_k\} \Big\} \\ &= \inf_{\bar{\boldsymbol{\Delta}}_k} \Big\{ -\mathrm{Tr}\{\bar{\boldsymbol{\Delta}}_k (\boldsymbol{A}_k + \boldsymbol{Z}_k)\} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k (\boldsymbol{A}_k + \boldsymbol{Z}_k)\} \\ &\qquad - \sigma_k^2 \gamma_k + \lambda_k (\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}} - \alpha_k) \Big\} \end{aligned} \tag{18}$$

and, $\lambda_k \ge 0$ and $\boldsymbol{Z}_k \succeq 0$ are the Lagrange variables. To find the infimum in (18), we equate the respective derivative to zero, hence

$$\begin{aligned} 0 &= -\frac{\partial\{\mathrm{Tr}\{\bar{\boldsymbol{\Delta}}_k(\boldsymbol{A}_k + \boldsymbol{Z}_k)\}\}}{\partial \bar{\boldsymbol{\Delta}}_k} \\ &\quad + \lambda_k \frac{\partial \left\{ \sqrt{\mathrm{vec}(\bar{\boldsymbol{\Delta}}_k)^H \boldsymbol{Q}_k^{-1} \mathrm{vec}(\bar{\boldsymbol{\Delta}}_k)} \right\}}{\partial \bar{\boldsymbol{\Delta}}_k} \\ &= -\boldsymbol{A}_k - \boldsymbol{Z}_k + \lambda_k \frac{\mathrm{vec}^{-1}\left(\boldsymbol{Q}_k^{-1}\mathrm{vec}(\bar{\boldsymbol{\Delta}}_k)\right)}{\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}}} \end{aligned} \tag{19}$$

where the function $\mathrm{vec}^{-1}(\boldsymbol{x})$ takes a $K^2 \times 1$ vector $\boldsymbol{x}$ as input and returns a $K \times K$ matrix $\boldsymbol{X}$ such that $\mathrm{vec}(\boldsymbol{X}) = \boldsymbol{x}$. Here, we have also accounted for the fact that the matrices $\boldsymbol{Z}_k$, $\boldsymbol{A}_k$, $\bar{\boldsymbol{\Delta}}_k$, and $\boldsymbol{Q}_k$ are all Hermitian. From (19), we get

$$\lambda_k \frac{\mathrm{vec}^{-1}\left(\boldsymbol{Q}_k^{-1}\mathrm{vec}(\bar{\boldsymbol{\Delta}}_k)\right)}{\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}}} = \boldsymbol{A}_k + \boldsymbol{Z}_k. \tag{20}$$

Taking the $\boldsymbol{Q}_k$-weighted Frobenius norm on both sides, we obtain from (20) the expression for the optimal Lagrange multipler $\lambda_k^\star$ as

$$\|\boldsymbol{A}_k + \boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} = \lambda_k^\star \frac{\sqrt{\mathrm{vec}(\bar{\boldsymbol{\Delta}}_k)^H \boldsymbol{Q}_k^{-1} \mathrm{vec}(\bar{\boldsymbol{\Delta}}_k)}}{\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}}} = \lambda_k^\star. \tag{21}$$

Next, inserting (20) in the first term of (18), the Lagrange dual function (18) becomes

$$\begin{aligned} g(\lambda_k, \boldsymbol{Z}_k) &= -\lambda_k \frac{\mathrm{Tr}\{\bar{\boldsymbol{\Delta}}_k \mathrm{vec}^{-1}(\boldsymbol{Q}_k^{-1}\mathrm{vec}(\bar{\boldsymbol{\Delta}}_k))\}}{\|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}}} \\ &\quad - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{A}_k + \boldsymbol{Z}_k)\} - \sigma_k^2\gamma_k \\ &\quad + \lambda_k \|\bar{\boldsymbol{\Delta}}_k\|_{\boldsymbol{Q}_k^{-1}} - \lambda_k \alpha_k \\ &= -\mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{A}_k + \boldsymbol{Z}_k)\} - \sigma_k^2\gamma_k - \lambda_k \alpha_k \end{aligned} \tag{22}$$

where we have used the fact that $\mathrm{Tr}\{\boldsymbol{X}\boldsymbol{Y}\} = \mathrm{vec}(\boldsymbol{X}^H)^H \mathrm{vec}(\boldsymbol{Y})$. Inserting the optimal $\lambda_k^\star$ from (21) in (22), we obtain

$$g(\boldsymbol{Z}_k) = -\mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{A}_k + \boldsymbol{Z}_k)\} - \sigma_k^2\gamma_k - \alpha_k \|\boldsymbol{A}_k + \boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} \tag{23}$$

and the dual problem corresponding to (17) can be written as

$$\begin{aligned} \max_{\boldsymbol{Z}_k} \quad & -\alpha_k \|\boldsymbol{A}_k + \boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{Z}_k + \boldsymbol{A}_k)\} - \sigma_k^2\gamma_k \\ \text{s.t.} \quad & \boldsymbol{Z}_k \succeq 0. \end{aligned} \tag{24}$$

Recall that the subproblem (17) is convex and clearly it is bounded below. Furthermore, there always exist error matrices, i.e., $\bar{\boldsymbol{\Delta}}_k = \tilde{\alpha}_k \boldsymbol{I}$ with $0 < \tilde{\alpha}_k < \alpha_k/\sqrt{N}$, that are strictly feasible. Therefore, using ([34]. Th. 1.7.1), we can conclude that strong duality between (17) and (24) holds. This allows us to replace the kth QoS constraint (16) by the equivalent constraint

$$\max_{\boldsymbol{Z}_k \succeq 0} \left(-\alpha_k \|\boldsymbol{A}_k + \boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{Z}_k + \boldsymbol{A}_k)\} - \sigma_k^2\gamma_k \right) \ge 0. \tag{25}$$

We observe that (16) and (25) are satisfied if there exists some $\boldsymbol{Z}_k \succeq 0$ for which

$$-\alpha_k \|\boldsymbol{A}_k + \boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{Z}_k + \boldsymbol{A}_k)\} - \sigma_k^2\gamma_k \ge 0. \tag{26}$$

The QoS constraints (16) can, therefore, also be replaced by (26) for $\boldsymbol{Z}_k \succeq 0$.

Using (26), along with the definition in (6), problem (11), for the case $L = 0$, can be modified as

$$\begin{aligned} \min_{\{\boldsymbol{W}_i, \boldsymbol{Z}_i\}} \quad & \sum_{i=1}^{K} \mathrm{Tr}\{\boldsymbol{W}_i\} \\ \text{s.t.} \quad & -\alpha_k \|\boldsymbol{A}_k + \boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{Z}_k + \boldsymbol{A}_k)\} - \sigma_k^2\gamma_k \ge 0 \\ & \boldsymbol{Z}_k \succeq 0,\ \boldsymbol{W}_k \succeq 0,\ \mathrm{rank}(\boldsymbol{W}_k) = 1;\ k = 1, \ldots, K. \end{aligned} \tag{27}$$

It is proven in Appendix A that $\boldsymbol{Z}_k^\star = \boldsymbol{0}$, $(k = 1, \ldots, K)$ is a solution of (27). To solve the remaining problem, we follow the SDR approach and drop the rank-one constraint, resulting in

$$\begin{aligned} \min_{\{\boldsymbol{W}_i\}} \quad & \sum_{i=1}^{K} \mathrm{Tr}\{\boldsymbol{W}_i\} \\ \text{s.t.} \quad & -\alpha_k \|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \boldsymbol{A}_k\} - \sigma_k^2\gamma_k \ge 0 \\ & \boldsymbol{W}_k \succeq 0;\ k = 1, \ldots, K. \end{aligned} \tag{28}$$

Problem (28) is a convex SDP and can be solved in polynomial time using interior-point algorithms [35], [36]. In Section VI, we discuss the rank of the solutions obtained from (28) and a procedure to find a rank-one approximation from the solution of (28).

B. The Proposed CR Beamformer

Next, we solve the CR beamforming problem. To modify the PU interference constraints in (11), avoiding the conservative approach of (13), we introduce the auxiliary matrix $\boldsymbol{C} \triangleq \sum_{k=1}^{K} \boldsymbol{w}_k \boldsymbol{w}_k^H$. The lth PU interference constraint in (11) can be written as

$$\max_{\substack{\|\bar{\boldsymbol{\Delta}}_{K+l}\|_{\boldsymbol{Q}_{K+l}^{-1}} \leq \alpha_{K+l} \\ \hat{\boldsymbol{R}}_{K+l}+\bar{\boldsymbol{\Delta}}_{K+l} \succeq 0}} \operatorname{Tr}\{\boldsymbol{C}(\hat{\boldsymbol{R}}_{K+l} + \bar{\boldsymbol{\Delta}}_{K+l})\} \leq \varepsilon_l. \tag{29}$$

The maximization on the left side of (29) corresponds to the following optimization problem:

$$\begin{aligned} \min_{\bar{\boldsymbol{\Delta}}_{K+l}} \quad & -\operatorname{Tr}\{\boldsymbol{C}(\hat{\boldsymbol{R}}_{K+l} + \bar{\boldsymbol{\Delta}}_{K+l})\} \\ \text{s.t.} \quad & \|\bar{\boldsymbol{\Delta}}_{K+l}\|_{\boldsymbol{Q}_{K+l}^{-1}} \leq \alpha_{K+l};\ \hat{\boldsymbol{R}}_{K+l} + \bar{\boldsymbol{\Delta}}_{K+l} \succeq 0. \end{aligned} \tag{30}$$

Problem (30) is convex for a given matrix $\boldsymbol{C}$. Following the same steps as used for the QoS constraints, the dual problem corresponding to the problem (30) can be written as

$$\begin{aligned} \max_{\boldsymbol{Z}_{K+l}} \quad & -\alpha_{K+l}\|(\boldsymbol{C} + \boldsymbol{Z}_{K+l})\|_{\boldsymbol{Q}_{K+l}} - \operatorname{Tr}\{\hat{\boldsymbol{R}}_{K+l}(\boldsymbol{C} + \boldsymbol{Z}_{K+l})\} \\ \text{s.t.} \quad & \boldsymbol{Z}_{K+l} \succeq 0. \end{aligned} \tag{31}$$

Note that the problem (30) is convex and bounded below. Furthermore, there always exist error matrices $\bar{\boldsymbol{\Delta}}_{K+l} = \tilde{\alpha}_{K+l}\boldsymbol{I}$ with $0 < \tilde{\alpha}_{K+l} < \alpha_{K+l}/\sqrt{N}$ that are strictly feasible. Therefore, using ([34]. Th. 1.7.1), it can be concluded that strong duality between (30) and (31) holds. Using the fact that $\boldsymbol{C} \succeq 0$ and $\boldsymbol{Z}_{K+l} \succeq 0$, we can define the eigendecomposition of $\boldsymbol{C} + \boldsymbol{Z}_{K+l}$ similar to $\boldsymbol{A}_k + \boldsymbol{Z}_k$, as given in Appendix A. Note that in this case all eigenvalues are non-negative. Then using the arguments similar to the arguments in Appendix A, it can be shown that $\boldsymbol{Z}_{K+l} = \boldsymbol{0}$ solves (31). Thus, the PU interference constraints (29) can be equivalently written as

$$\operatorname{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\} + \alpha_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}} \leq \varepsilon_l. \tag{32}$$

Using (26) and (32), along with the definition in (6), problem (11) can be modified as

$$\begin{aligned} \min_{\{\boldsymbol{W}_i, \boldsymbol{Z}_i\}} \quad & \sum_{i=1}^{K} \operatorname{Tr}\{\boldsymbol{W}_i\} \\ \text{s.t.} \quad & -\alpha_k\|\boldsymbol{A}_k+\boldsymbol{Z}_k\|_{\boldsymbol{Q}_k} - \operatorname{Tr}\{\hat{\boldsymbol{R}}_k(\boldsymbol{Z}_k+\boldsymbol{A}_k)\} - \sigma_k^2\gamma_k \geq 0 \\ & \boldsymbol{Z}_k \succeq 0,\ \boldsymbol{W}_k \succeq 0,\ \operatorname{rank}(\boldsymbol{W}_k) = 1;\ k = 1,\ldots,K \\ & \operatorname{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\} + \alpha_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}} \leq \varepsilon_l;\ l = 1,\ldots,L. \end{aligned} \tag{33}$$

Similar to the problem (27), it can be proven that $\boldsymbol{Z}_k^\star = \boldsymbol{0}$, $(k = 1,\ldots,K)$ at the optimum of (33). Following the SDR approach, we can drop the non-convex rank-one constraints, resulting in

$$\begin{aligned} \min_{\{\boldsymbol{W}_i\}} \quad & \sum_{i=1}^{K} \operatorname{Tr}\{\boldsymbol{W}_i\} \\ \text{s.t.} \quad & -\alpha_k\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k} - \operatorname{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\} - \sigma_k^2\gamma_k \geq 0 \\ & \boldsymbol{W}_k \succeq 0;\ k = 1,\ldots,K \\ & \operatorname{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\} + \alpha_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}} \leq \varepsilon_l;\ l = 1,\ldots,L. \end{aligned} \tag{34}$$

Problem (34) is a convex SDP problem and can be solved in polynomial time using interior-point algorithms [35], [36] (see Section VI for obtaining a rank-one solution from the solution of (34)).

We remark that our analysis in this section and in Appendix A reveals that for the conventional and the cognitive robust downlink problem at optimum the Lagrange multipliers $\boldsymbol{Z}_k$ and $\boldsymbol{Z}_{K+l}$ corresponding to the positive definiteness constraints in (17) and (30), respectively, are zero. This means that at optimum the positive definiteness constraints are inactive. In this sense our analysis rigorously proves the non-trivial result that the positive definiteness constraints can be neglected in the original problem (11) without introducing any conservative approximation.

Comparing the QoS constraints of the robust problem (34) with those of its non-robust counterpart (7), we observe that the term $-\alpha_k\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}$ is added in the robust formulation. Similarly, the term $\alpha_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}$ is added to the PU interference constraints in the robust case. These terms strengthen the constraints and can be viewed as a penalty for making the problem robust. These penalty terms, as we show later in Section VI, may lead to higher rank solutions of the relaxed optimization problem in (34). Further, for colored CSI errors and the weighting matrices $\boldsymbol{Q}_m$ defined according to (9), the errors along the dominant eigenvectors of $\boldsymbol{Q}_m$ are more prominent than the errors along the minor eigenvectors. Correspondingly, the penalty for robustness is also higher or smaller depending on the magnitude of the corresponding eigenvalues.

V. Relation to the Probabilistic Cognitive Beamforming Approach

Next, we derive the downlink beamformers using a probabilistic model and relate the results to the worst-case approach. To this end, we replace the QoS and PU interference constraints in (11) with probabilistic constraints. The robust problem can then be formulated as

$$\begin{aligned} \min_{\{\boldsymbol{W}_i\}} \quad & \sum_{i=1}^{K} \operatorname{Tr}\{\boldsymbol{W}_i\} \\ \text{s.t.} \quad & \Pr\left(-\operatorname{Tr}\{(\hat{\boldsymbol{R}}_k+\boldsymbol{\Delta}_k)\boldsymbol{A}_k\} \geq \sigma_k^2\gamma_k\right) \geq p_k;\ k = 1,\ldots,K \\ & \Pr(\operatorname{Tr}\{(\hat{\boldsymbol{R}}_{K+l} + \boldsymbol{\Delta}_{K+l})\boldsymbol{C}\} \leq \varepsilon_l) \geq p_l;\ l = 1,\ldots,L \\ & \operatorname{rank}(\boldsymbol{W}_k) = 1, \boldsymbol{W}_k \succeq 0 \end{aligned} \tag{35}$$

where $\Pr(\cdot)$ denotes probability taken with respect to $\boldsymbol{\Delta}_m, (m = 1,\ldots,K+L)$ and, p_k and p_l are predefined probability thresholds.

We assume that the real-valued diagonal and complex-valued upper or lower triangle elements of $\boldsymbol{\Delta}_m$ are zero-mean random with the covariance matrices as defined in (8). We define new real-valued random variables

$$y_k \triangleq -\operatorname{Tr}\{(\hat{\boldsymbol{R}}_k + \boldsymbol{\Delta}_k)\boldsymbol{A}_k\} \tag{36}$$

$$y_{K+l} \triangleq \operatorname{Tr}\{(\hat{\boldsymbol{R}}_{K+l} + \boldsymbol{\Delta}_{K+l})\boldsymbol{C}\}. \tag{37}$$

Then

$$\mathrm{E}\{y_k\} = \mathrm{E}\{-\operatorname{Tr}\{(\hat{\boldsymbol{R}}_k + \boldsymbol{\Delta}_k)\boldsymbol{A}_k\}\} = -\operatorname{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\} \tag{38}$$

$$\mathrm{E}\{y_{K+l}\} = \mathrm{E}\{\operatorname{Tr}\{(\hat{\boldsymbol{R}}_{K+l} + \boldsymbol{\Delta}_{K+l})\boldsymbol{C}\}\} = \operatorname{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}. \tag{39}$$

The variance of y_k can be computed as

$$\begin{aligned}
&\mathrm{E}\{\mathrm{Tr}\{\boldsymbol{\Delta}_k\boldsymbol{A}_k\}\mathrm{Tr}\{\boldsymbol{\Delta}_k\boldsymbol{A}_k\}^*\}\\
&\quad= \mathrm{E}\left\{\mathrm{vec}\left(\boldsymbol{A}_k^H\right)^H \mathrm{vec}\left(\boldsymbol{\Delta}_k\right)\mathrm{vec}(\boldsymbol{\Delta}_k)^H\mathrm{vec}\left(\boldsymbol{A}_k^H\right)\right\}\\
&\quad= \mathrm{vec}\left(\boldsymbol{A}_k^H\right)^H \mathrm{E}\{\mathrm{vec}(\boldsymbol{\Delta}_k)\mathrm{vec}(\boldsymbol{\Delta}_k)^H\}\mathrm{vec}\left(\boldsymbol{A}_k^H\right)\\
&\quad= \eta_k\mathrm{vec}\left(\boldsymbol{A}_k^H\right)^H \boldsymbol{Q}_k\mathrm{vec}\left(\boldsymbol{A}_k^H\right)\\
&\quad= \eta_k\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}^2.
\end{aligned} \tag{40}$$

Similarly, the variance of y_{K+l} can be computed as

$$\mathrm{E}\{\mathrm{Tr}\{\boldsymbol{\Delta}_{K+l}\boldsymbol{C}\}\mathrm{Tr}\{\boldsymbol{\Delta}_{K+l}\boldsymbol{C}\}^*\} = \eta_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}^2. \tag{41}$$

Suppose that the random variables y_k and y_{K+l} have probability density functions $f_{Y_k}(y)$ and $f_{Y_{K+l}}(y)$, respectively. Then, we can rewrite

$$\begin{aligned}
&\Pr\left(-\mathrm{Tr}\{(\hat{\boldsymbol{R}}_k+\boldsymbol{\Delta}_k)\boldsymbol{A}_k\} \geq \sigma_k^2\gamma_k\right)\\
&\quad= \int_{\sigma_k^2\gamma_k}^{\infty} f_{Y_k}(y)\mathrm{d}y\\
&\quad= \sqrt{2\eta_k}\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}\int_{d_k}^{\infty} f_{Y_k}(\sqrt{2\eta_k}\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}y\\
&\qquad - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\})\mathrm{d}y\\
&\quad\triangleq G_k(d_k)
\end{aligned} \tag{42}$$

and

$$\begin{aligned}
&\Pr(\mathrm{Tr}\{(\hat{\boldsymbol{R}}_{K+l}+\boldsymbol{\Delta}_{K+l})\boldsymbol{C}\} \leq \varepsilon_l)\\
&\quad= \int_{-\infty}^{\varepsilon_l} f_{Y_{K+l}}(y)\mathrm{d}y\\
&\quad= \sqrt{2\eta_{K+l}}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}\int_{-\infty}^{d_{K+l}} f_{Y_{K+l}}(\sqrt{2\eta_{K+l}}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}y\\
&\qquad + \mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\})\mathrm{d}y\\
&\quad\triangleq G_{K+l}(d_{K+l})
\end{aligned} \tag{43}$$

where

$$d_k \triangleq \frac{\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\}+\sigma_k^2\gamma_k}{\sqrt{2\eta_k}\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}};\quad d_{K+l} \triangleq \frac{\varepsilon_l - \mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}}{\sqrt{2\eta_{K+l}}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}}. \tag{44}$$

Using (42) and (43), the probabilistic constraints in (35) can be modified as

$$d_k \geq G_k^{-1}(p_k) \tag{45}$$
$$d_{K+l} \geq G_{K+l}^{-1}(p_l). \tag{46}$$

In the following we investigate, as an example, the special case, that the real-valued diagonal and complex-valued upper or lower triangle elements of $\boldsymbol{\Delta}_m, (m = 1,\ldots,K+L)$ are zero-mean jointly Gaussian random variables. Then $G_k(d_k)$ and $G_{K+l}(d_{K+l})$ take the form (47) and (48), respectively [see below]. Using the fact that $\mathrm{erf}(-x) = -\mathrm{erf}(x)$, the QoS and PU interference constraints in (35) become

$$\mathrm{erf}\left(\frac{-\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\}-\sigma_k^2\gamma_k}{\sqrt{2\eta_k}\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}}\right) \geq 2p_k - 1 \tag{49}$$
$$\mathrm{erf}\left(\frac{\varepsilon_l-\mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}}{\sqrt{2\eta_{K+l}}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}}\right) \geq 2p_l - 1. \tag{50}$$

After simple manipulations, we can modify (49) and (50) as

$$-\rho_k\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\} - \sigma_k^2\gamma_k \geq 0 \tag{51}$$
$$\mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\} + \rho_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}} \leq \varepsilon_l \tag{52}$$

where $\rho_k \triangleq \sqrt{2\eta_k}\mathrm{erf}^{-1}(2p_k-1), (k=1,\ldots,K)$ and $\rho_{K+l} \triangleq \sqrt{2\eta_{K+l}}\mathrm{erf}^{-1}(2p_l-1), (l=1,\ldots,L)$.

Using (51) and (52), and applying SDR, (35) can be rewritten as

$$\begin{aligned}
\min_{\{\boldsymbol{W}_i\}}\quad & \sum_{i=1}^{K}\mathrm{Tr}\{\boldsymbol{W}_i\}\\
\text{s.t.}\quad & -\rho_k\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\} - \sigma_k^2\gamma_k \geq 0\\
& \mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\} + \rho_{K+l}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}} \leq \varepsilon_l;\ l=1,\ldots,L\\
& \boldsymbol{W}_k \succeq 0;\ k=1,\ldots,K.
\end{aligned} \tag{53}$$

We observe that the outage probability based problem (53) is mathematically equivalent to the worst-case formulation (34) for proper choices of α_k and α_{K+l}. The norm-bound coefficients α_k and α_{K+l} in (34), are replaced by the constants ρ_k and ρ_{K+l} in (53), that reflect the outage probability, respectively.

$$G_k(d_k) = \begin{cases} \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{-\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\}-\sigma_k^2\gamma_k}{\sqrt{2\eta_k}\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}}\right), & \text{if } \sigma_k^2\gamma_k \leq -\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\}\\ \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{\sigma_k^2\gamma_k+\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\}}{\sqrt{2\eta_k}\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}}\right), & \text{if } \sigma_k^2\gamma_k \geq -\mathrm{Tr}\{\hat{\boldsymbol{R}}_k\boldsymbol{A}_k\} \end{cases} \tag{47}$$

$$G_{K+l}(d_{K+l}) = \begin{cases} \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{\mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}-\varepsilon_l}{\sqrt{2\eta_{K+l}}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}}\right), & \text{if } \varepsilon_l \leq \mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}\\ \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\varepsilon_l-\mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}}{\sqrt{2\eta_{K+l}}\|\boldsymbol{C}\|_{\boldsymbol{Q}_{K+l}}}\right), & \text{if } \varepsilon_l \geq \mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l}\boldsymbol{C}\}. \end{cases} \tag{48}$$

VI. Rank Properties of the Resulting SDP

The solutions to the SDR-based problems are not rank-one in general. In such cases, so-called *randomization techniques* [37] can be used to obtain an approximate solution of the original problem from the solution of the relaxed problem. However, in the case of (28), (34), and (53), standard randomization techniques cannot be directly applied. In the randomization approach the candidate vectors are scaled up to meet the constraints but simple scaling in (28), (34), and (53) generally results in the violation of the QoS and PU interference constraints. Nevertheless, our simulation results show that we generally obtain rank-one solutions for $\boldsymbol{W}_k$ and, consequently, the weight vectors $\boldsymbol{w}_k$ can be retrieved exactly from the principal eigenvectors of $\boldsymbol{W}_k$. However, there exist specific scenarios with highly symmetric channels or very large channel uncertainties in which higher rank solutions are obtained. When a higher rank solution is obtained, the corresponding rank-one solution can be approximated using the principal eigenvectors of $\boldsymbol{W}_k$. The resulting solution, however, will not guarantee that all the constraints are satisfied. We remark that based on the results of this paper and [31], an efficient iterative robust beamforming technique can be applied that extends the popular power iteration methods [2]–[4], to compute feasible rank-one solutions (when convergent). For details we refer to [32].

In the following we provide a simple example and illustrate under which conditions such higher-rank solutions may be obtained. Consider the case of $N = 3$, $K = 2$, and $L = 0$ with covariance matrices

$$\hat{\boldsymbol{R}}_1 = \mathrm{diag}\{r_1, r_2, 0\}, \quad \hat{\boldsymbol{R}}_2 = \mathrm{diag}\{0, 0, r_3\} \tag{54}$$

where $\mathrm{diag}\{\cdot\}$ denotes a diagonal matrix. In Appendix B, it is shown that in this case the solution has the diagonal structure

$$\boldsymbol{W}_1 = \mathrm{diag}\{a_{11}, a_{22}, a_{33}\}, \quad \boldsymbol{W}_2 = \mathrm{diag}\{b_{11}, b_{22}, b_{33}\} \tag{55}$$

and either a_{11} or a_{22}, and b_{33} must be non-zero for the QoS constraints to be satisfied. The solution (55) is generally of higher rank if either of a_{33}, b_{11} and b_{22} are non-zero. Further it is shown that if these values are zero, which, according to Appendix B, is, e.g., the case for $\gamma_1\gamma_2 \leq 1$, then the solution has the form

$$\boldsymbol{W}_1^\star = \mathrm{diag}\{\mu^\star c, (1-\mu^\star)c, 0\}, \quad \boldsymbol{W}_2^\star = \mathrm{diag}\{0, 0, b_{33}\} \tag{56}$$

where the values of b_{33} and c depend on the parameters of (34) and $0 \leq \mu^\star \leq 1$. We also show in Appendix B that larger values of the uncertainty thresholds α_1 and α_2 favor a higher rank solution while smaller values result in a rank-one solution. This observation is intuitively clear as the norm term $-\alpha_k\|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k}$, with $\boldsymbol{A}_k$ defined in (15), becomes dominant in the QoS constraints in (34) for increased α_k and this term favors the higher rank solutions.

In Appendix C, we show that the simple class of diagonal channel covariances in (54), for which higher rank solutions are obtained, can be extended to more general, however still highly symmetric, channel cases.

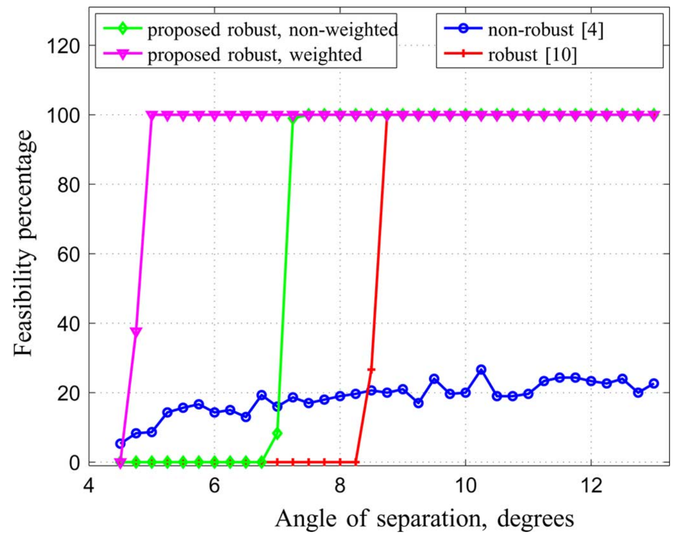

Fig. 1. Feasibility percentage versus angular separation ϕ.

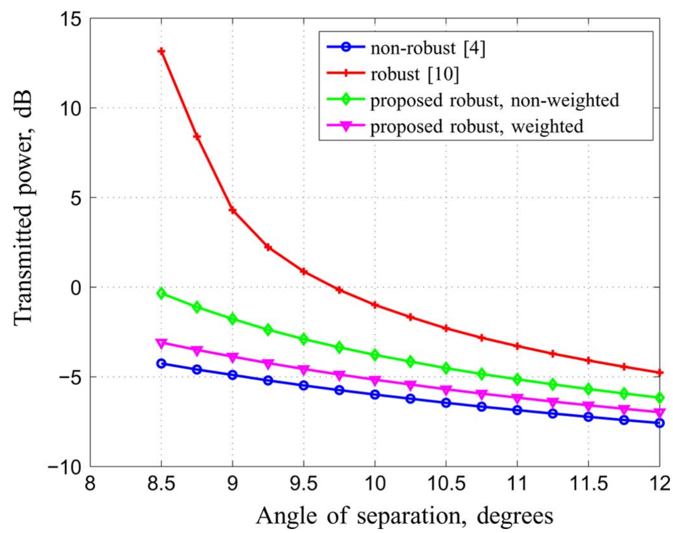

Fig. 2. Total transmitted power versus angular separation ϕ.

VII. Simulation Results

In this section, we present simulation results for both the conventional and the cognitive beamforming approaches. First, in Figs. 1–7, we compare the performance of the proposed robust approach for the conventional downlink beamforming to the non-robust and robust approaches of [4] and [10], respectively. In Figs. 8–11, we compare the proposed robust CR beamforming approaches with the non-robust and robust CR methods of [27].

For the conventional beamforming, we first consider the same scenario as in [4]. The base station is equipped with a uniform linear array of $N = 6$ sensors spaced half a wavelength apart and $K = 3$ single-antenna users are assumed. One of the users is located at $\theta_1 = 20^\circ$ relative to the array broadside, while the other two are located at $\theta_{2,3} = 20^\circ \pm \phi$, and ϕ is varied from 6° to 12°. The users are assumed to be surrounded by a large number of local scatterers corresponding to an angular spread of $\sigma_\theta = 5^\circ$, as seen from the base station. The channel covariance matrices $\boldsymbol{R}_k$, $k = 1, \ldots, K$ are calculated in the same way

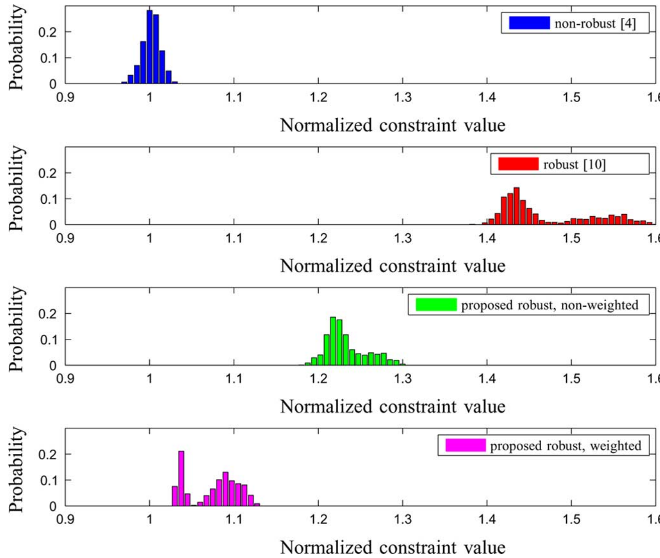

Fig. 3. Histogram of normalized QoS constraints ($\phi = 9.5^\circ$).

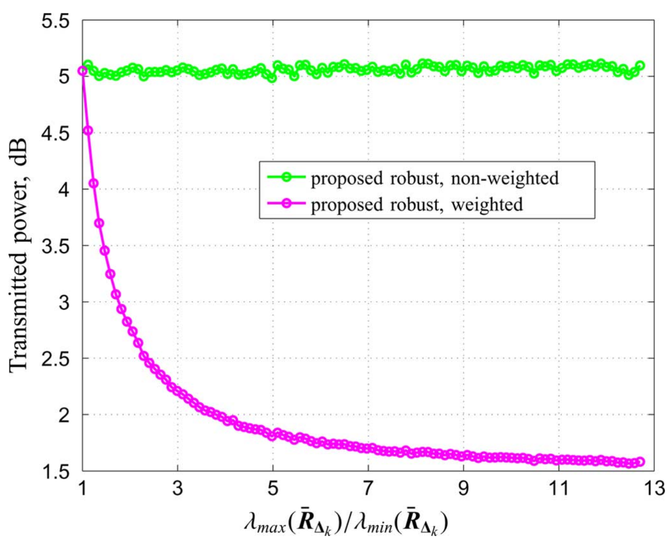

Fig. 4. Total transmitted power versus the condition number of the error covariance matrix.

as in [4]. Throughout the simulations, the sensor noise at the receivers is assumed to be additive white Gaussian with variances $\sigma_k^2 = 0.1$, $(k = 1, \ldots, K)$. To model the mismatch matrices $\boldsymbol{\Delta}_k$ $(k = 1, \ldots, K)$, we considered the practically important case that the channel covariance matrices are estimated in the presence of colored noise, e.g., resulting from interference from neighboring base stations, such that the errors have a particular spatial signature. In this case the weighted Frobenius norm shall essentially whiten the estimation errors. In our simulations, we generate errors that are randomly distributed within ellipsoids with orientations and shapes corresponding to the eigenvectors and eigenvalues of interference-plus-noise covariance matrices from the neighboring cells, respectively. Let $\boldsymbol{R}_{k,\mathrm{I}}$ denote the interference-plus-noise covariance matrix present during the estimation of $\boldsymbol{R}_k$ with eigendecomposition $\boldsymbol{R}_{k,\mathrm{I}} = \boldsymbol{U}_{k,\mathrm{I}}\boldsymbol{\Gamma}_{k,\mathrm{I}}\boldsymbol{U}_{k,\mathrm{I}}^H$. Then we model the mismatch matrix as $\boldsymbol{\Delta}_k = \boldsymbol{U}_{k,\mathrm{I}}\boldsymbol{\Gamma}_{k,\mathrm{I}}^{1/2}\boldsymbol{\Upsilon}_k\boldsymbol{\Gamma}_{k,\mathrm{I}}^{1/2}\boldsymbol{U}_{k,\mathrm{I}}^H$ where $\boldsymbol{\Upsilon}_k$ denotes a random matrix with $\|\boldsymbol{\Upsilon}_k\| \leq \alpha_k$ and $\boldsymbol{R}_k = \hat{\boldsymbol{R}}_k + \boldsymbol{\Delta}_k \succeq 0$. Choosing the

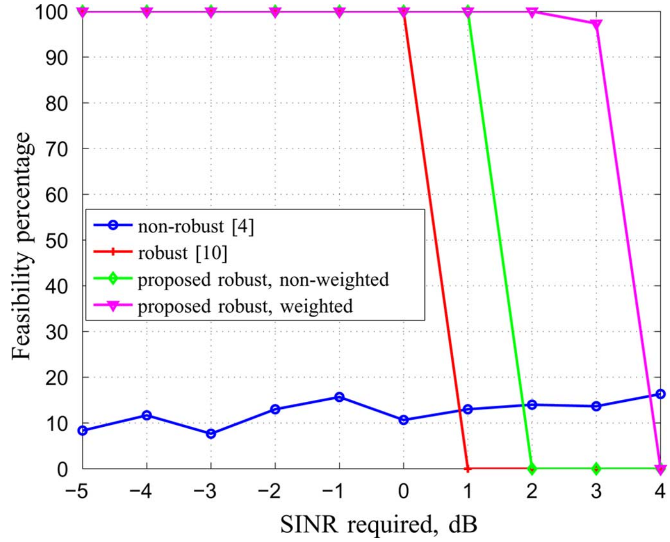

Fig. 5. Feasibility percentage versus required SINR γ.

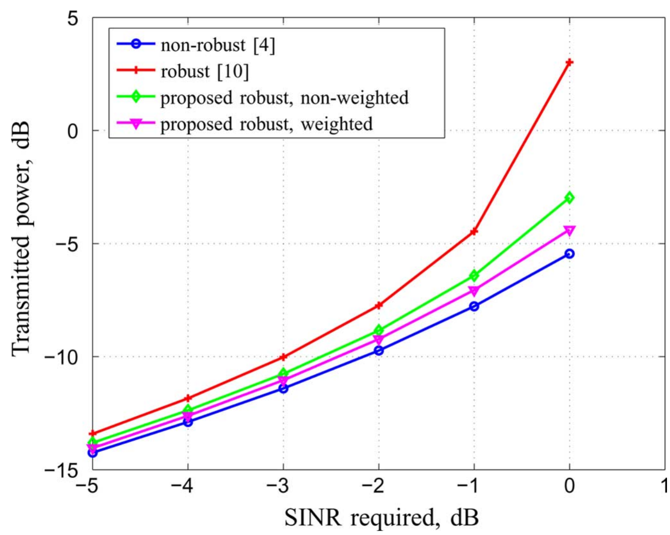

Fig. 6. Total transmitted power versus required SINR γ.

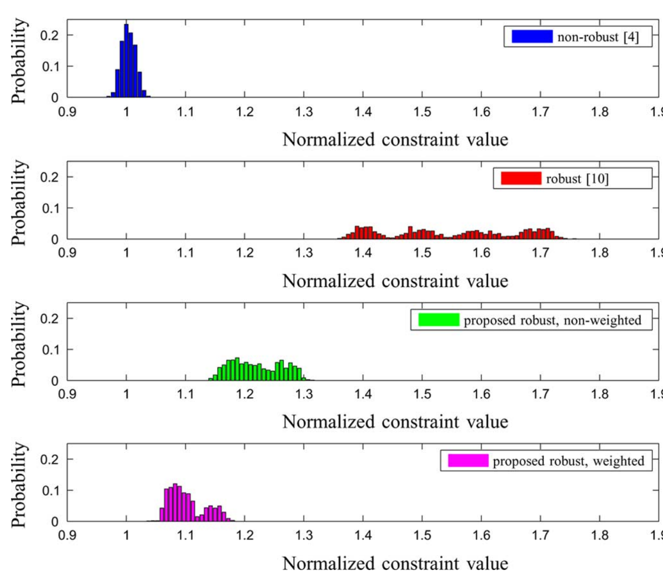

Fig. 7. Histogram of normalized QoS constraints ($\gamma = 0$ dB).

weighting matrix $\boldsymbol{Q}_k = \bar{\boldsymbol{R}}_{\boldsymbol{\Delta}_k}^T \otimes \bar{\boldsymbol{R}}_{\boldsymbol{\Delta}_k}$ with $\bar{\boldsymbol{R}}_{\boldsymbol{\Delta}_k} = E[\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^H]$, where $\otimes$ denotes the Kronecker matrix product, it can be shown that $\|\boldsymbol{\Delta}_k\|_{\boldsymbol{Q}_k^{-1}} = \sqrt{\mathrm{Tr}\{\boldsymbol{\Gamma}_{k,\mathrm{I}}^{-1/2} \boldsymbol{U}_{k,\mathrm{I}}^H \boldsymbol{\Delta}_k \bar{\boldsymbol{R}}_{\Delta_k}^{-1} \boldsymbol{\Delta}_k \boldsymbol{U}_{k,\mathrm{I}} \boldsymbol{\Gamma}_{k,\mathrm{I}}^{-1/2}\}} = \|\boldsymbol{\Upsilon}_k\| \leq \alpha_k$, where we made use of the property $\mathrm{vec}(\boldsymbol{AXB}) = (\boldsymbol{B}^T \otimes \boldsymbol{A})\mathrm{vec}(\boldsymbol{X})$ for arbitrary matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{X}$ of comfortable dimensions. For a fair performance comparison of the robust schemes with both weighted and non-weighted Frobenius norms, we chose for the non-weighted robust scheme a bound $\alpha_{k,\mathrm{n-w}} = \sqrt{\lambda_{\boldsymbol{\Delta},max}}$, which follows from the observation that $\|\boldsymbol{\Delta}_k\|_{\boldsymbol{Q}_k^{-1}} \leq \alpha_k$ implies $\|\boldsymbol{\Delta}_k\|_{\boldsymbol{I}} \leq \sqrt{\lambda_{\boldsymbol{\Delta},max}}\alpha_k = \alpha_{k,\mathrm{n-w}}$, where $\lambda_{\boldsymbol{\Delta},max}$ is the maximum eigenvalue of $\boldsymbol{Q}_k$. In our first simulation example the colored interference-plus-noise covariance matrices $\boldsymbol{R}_{k,\mathrm{I}}$ are composed of an interference part according to the covariance model [4], with angles 50°, 22° and −10°, and an additive white noise term with variance 0.1.

We assume that $\gamma_k = \gamma$ and $\alpha_k = \alpha = 0.2$, for all $k = 1, \ldots, K$. A total number of 500 Monte-Carlo runs has been used. In all simulations on conventional beamforming, the proposed robust technique is compared to the robust technique of [10]. As a benchmark we also display the results for the non-robust technique of [4].

In Fig. 1, we show the feasibility percentage, i.e., the percentage of channel realizations for which the different schemes under consideration yield feasible solutions. A beamforming solution is considered as feasible, if it satisfies all the constraints in (5) (for $L = 0$) for the true covariance matrices. The simulations are performed for $\gamma = 2$ dB. We observe that the proposed robust approach with both weighted and non-weighted Frobenius norms outperforms the non-robust and robust approaches of [4] and [10], respectively, in terms of feasibility percentage. Note, that the feasibility of the non-robust approach is highly dependent on the structure of the error matrices $\boldsymbol{\Delta}_k$ and can change with the variations in the considered error distribution. Fig. 2 displays the transmitted power of all the techniques for the same scenario. For better comparison, here we consider only the cases for which the robust scheme of [10] yields feasible solutions. We remark that in all these instances the proposed approach is always feasible. Our robust approach is more power efficient compared to the robust approach of [10]. Recall that the proposed approach avoids the conservative approximations of [10] and, therefore, this result is in line with the expectations. From Fig. 2, the non-robust approach of [4] appears to be more power efficient but, as can be observed from Fig. 1, it does not always provide a feasible solution.

For further comparison, we plot the histogram of the achieved normalized QoS in Fig. 3 for $\phi = 9.5°$. We define the normalized QoS as

$$\zeta_k = \frac{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k}{\sum_{\substack{i=1 \\ i \neq k}}^{K} \gamma_k \boldsymbol{w}_i^H \boldsymbol{R}_k \boldsymbol{w}_i + \gamma_k \sigma_k^2}; \; k = 1, \ldots, K. \tag{57}$$

Due to the normalization (57), a value greater than one in Fig. 3 corresponds to satisfied QoS constraints. The proposed approach and the robust approach of [10] both satisfy all the constraints, as expected, but in the case of the non-robust approach of [4], a considerable number of constraints is not satisfied.

In order to further illustrate the benefits of using the weighted Frobenius norm, we show in Fig. 4 the transmitted power obtained for elliptic uncertainty sets with increasing eccentricity. We compare the performance of our weighted and non-weighted robust schemes. We fix the noise variance to 0.5 and vary the power of the interference part in the noise and interference matrix $\boldsymbol{R}_{k,\mathrm{I}}$ from 0.01 to 5, in order to obtain error covariance matrices with increasing eigenvalue spread.

Next, we consider the case with $N = 6$ sensors at the base station and $K = 4$ single-antenna users fixed at locations 10°, 40°, 55°, and 70° relative to the array broadside, while the QoS target γ is varied from −6 dB to 4 dB.

Here, we assume the angular spread of $\sigma_\theta = 10°$ and $\alpha_k = \alpha = 0.25$, for all $k = 1, \ldots, K$. Similar as in the previous example, colored interference-plus-noise covariance matrices $\boldsymbol{R}_{k,I}$ corresponding to angles 60°, 60°, 40°, and 43° are assumed with an angular spread of $\sigma_\theta = 5°$. The variance of the interfering signals is 0.25 and the variance of the noise is 0.1. We plot the feasibility percentage of all the schemes in Fig. 5 and observe that the proposed robust approach outperforms both the non-robust and robust approaches of [4] and [10], respectively, in terms of the feasibility percentage. In Fig. 6, we plot the corresponding transmitted power for all these techniques for the cases when the robust technique of [10] is feasible. The proposed approach remains more power efficient compared to the robust approach of [10] for this scenario as well. The non-robust approach of [4] exhibits reduced transmitted power but, as mentioned above, it cannot guarantee a feasible solution. In order to elaborate on this issue further, we plot the histogram of the normalized QoS constraints in Fig. 7 for $\gamma = 0$ dB. We observe from Fig. 7 that the non-robust approach does not satisfy all the constraints while all the constraints are satisfied for both the robust approaches.

In Figs. 8–11, we present simulation results for the CR beamformer. To reveal that in our proposed robust approach performance gains are obtained irrespective of the introduced weighting, we consider the case that the CSI errors are white and that the uncertainty sets are modeled by the non-weighted Frobenius norm. We compare the proposed robust approach to the non-robust and robust techniques of [27]. We assume $N = 6$, $K = 4$, and $L = 3$. The channel covariance matrices for the SUs and PUs, $\boldsymbol{R}_k$ and $\boldsymbol{R}_{K+l}$, are generated in the same way as in the case of the conventional beamformer above. The SUs are located at 10°, 40°, 55°, and 70° relative to the array broadside, while the PUs are located at 25°, 83°, and 86° relative to the array broadside. The users are assumed to be surrounded by a large number of local scatterers corresponding to an angular spread of $\sigma_\theta = 10°$ in the case of SUs and $\sigma_\theta = 10°$ for the PUs, as seen from the base station. The error matrices $\boldsymbol{\Delta}_k$ and $\boldsymbol{\Delta}_{K+l}$ are randomly generated in a sphere centered at zero with radii α_k and α_{K+l}, respectively. The resulting error matrices are then added to the covariance matrices to obtain the estimated covariance matrices $\hat{\boldsymbol{R}}_k$ and $\hat{\boldsymbol{R}}_{K+l}$, respectively. If $\hat{\boldsymbol{R}}_k$ or $\hat{\boldsymbol{R}}_{K+l}$ contain negative eigenvalues, then the corresponding eigenvalues are replaced by 0. We assume that $\gamma_k = \gamma$, $\alpha_k = \alpha_{K+l} = \alpha = 0.25$ and $\varepsilon_l = \varepsilon = 5$ dB, for all $k = 1, \ldots, K$ and $l = 1, \ldots, L$. A total number of 500 Monte-Carlo runs has been performed.

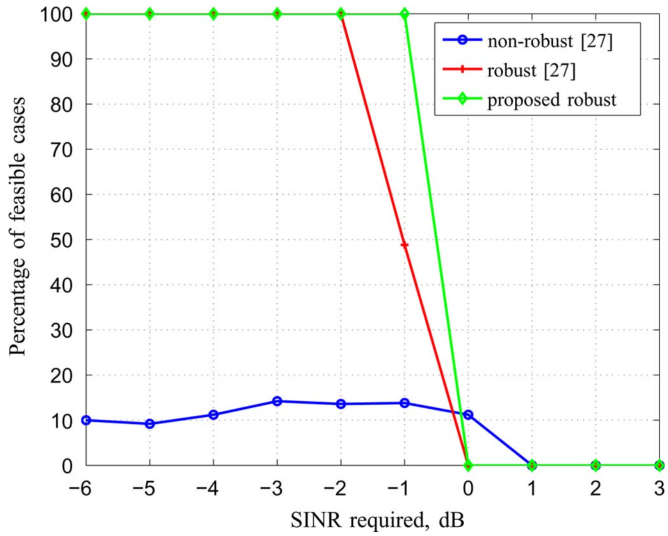

Fig. 8. Feasibility percentage of the CR beamforming schemes.

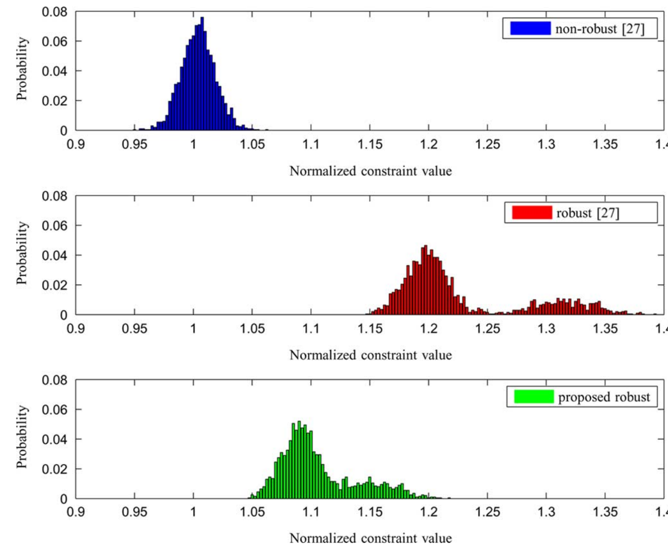

Fig. 10. Histogram of normalized QoS constraints ($\gamma = -2$ dB).

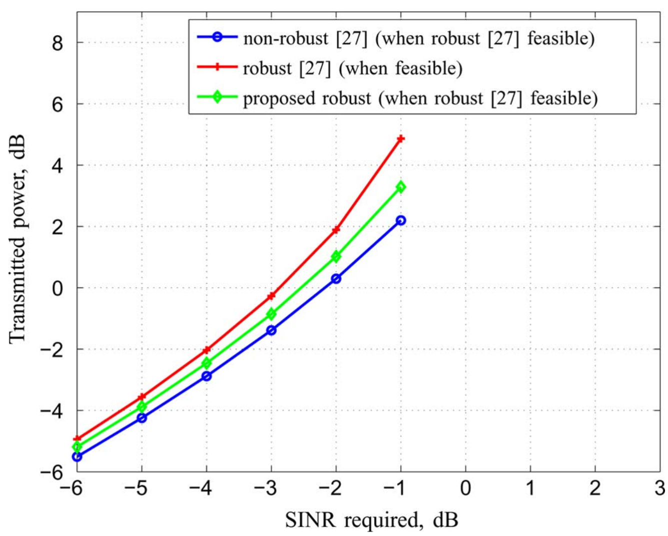

Fig. 9. Total transmitted power versus SINR required by SUs.

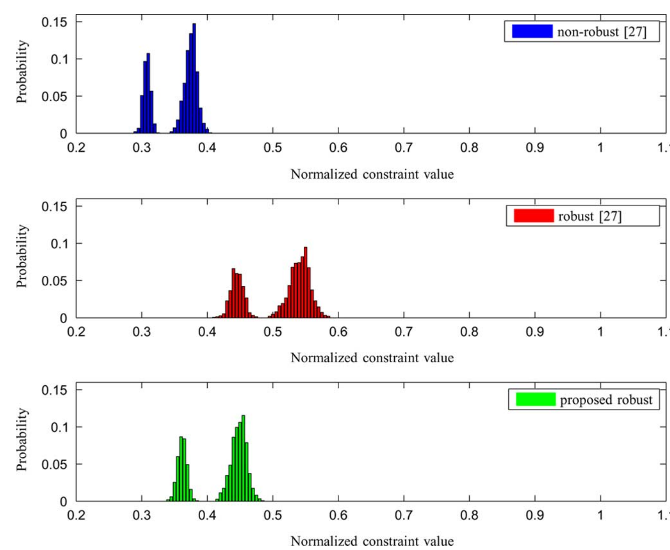

Fig. 11. Histogram of normalized PU interference constraints ($\gamma = -2$ dB).

In Fig. 8, we plot the feasibility percentage of all the schemes for targets γ varying from -6 dB to 3 dB. Similar to the conventional beamforming case, a solution is considered as feasible, if it satisfies all the constraints in (5) for the true covariance matrices. We observe that the percentage of feasible runs for both the robust approaches decreases monotonously as the target SINR increases. The proposed robust approach shows an increased feasibility percentage as compared to both the robust and non-robust approaches of [27]. It is important to note that similar to the case of the conventional beamforming problem, the feasibility percentage of the non-robust CR approach is highly dependent on the considered error distribution for matrices $\mathbf{\Delta}_k$ and $\mathbf{\Delta}_{K+l}$. Fig. 9 displays the corresponding transmitted power versus the SINR required by the SUs for the three schemes under consideration. Similar to the conventional beamforming case, we consider only those cases for which the robust scheme of [27] yields feasible solutions. We remark that in all these cases, the proposed approach is always feasible. Our method is more power efficient compared to the robust approach of [27]. The non-robust approach of [27] appears to be more power efficient, but the solutions it provides are not feasible when CSI errors are taken into account.

For further insight, we plot the histogram of the achieved normalized QoS and normalized PU interference power in Figs. 10 and 11, respectively for $\gamma = -2$ dB. The normalized QoS is defined as in (57), while the normalized PU interference power is defined as

$$\xi_l = \frac{1}{\varepsilon_l} \sum_{i=1}^{K} \boldsymbol{w}_i^H \boldsymbol{R}_{K+l} \boldsymbol{w}_i; \ l = 1, \ldots, L. \tag{58}$$

From (57), a value greater than one in Fig. 10 implies that the corresponding QoS constraint is satisfied. Similarly, from (58), in Fig. 11, a value less than one corresponds to satisfied PU interference constraint. The proposed robust approach and the robust approach of [27] both satisfy all the constraints, as expected, but for the non-robust approach of [27], a considerable number of constraints is left unsatisfied.

VIII. Conclusion

In this paper we have considered the problem of worst-case robust multiuser downlink beamforming using second order covariance-based CSI. In addition we have treated the robust downlink beamforming problem in a CR scenario with additional constraints for limiting the interference leaked to the PUs. We have derived exact reformulations of worst-case QoS and PU constraints using Lagrange duality. We have then converted the resulting problems into convex SDP problems by applying SDR. The final problem formulations have additional terms in the QoS and PU constraints that show the penalty being paid for achieving the robustness. Additionally, we have shown that solving the aforementioned downlink beamforming problems using the probabilistic constraints leads to the same solutions. Although generally rank-one solutions are obtained from the resulting SDPs, we have discussed some special cases under which higher rank solutions are obtained. We have verified the improvements in terms of the transmitted power and feasibility of the problem in the simulations.

Appendix A
Proof of $\boldsymbol{Z}_k^\star = \mathbf{0}$

Let the optimal matrices solving (27) be given by $(\{\boldsymbol{W}_i^\star\}_{i=1}^K, \{\boldsymbol{Z}_i^\star\}_{i=1}^K)$. In order to show that $\boldsymbol{Z}_k^\star = \mathbf{0}$ is optimal for (27), we define $\boldsymbol{Z}_k \triangleq \boldsymbol{Z}_k^{||} + \boldsymbol{Z}_k^{\perp}$ where matrix $\boldsymbol{Z}_k^{||}$ is in the space spanned by the eigenvectors of $\boldsymbol{A}_k$, and $\boldsymbol{Z}_k^{\perp}$ is in the null space of $\boldsymbol{A}_k$. We then define the eigendecomposition of $\boldsymbol{A}_k + \boldsymbol{Z}_k$ as

$$\boldsymbol{A}_k + \boldsymbol{Z}_k = \gamma_k \sum_{\substack{i=1\\ i\neq k}}^{K} \boldsymbol{W}_i - \boldsymbol{W}_k + \boldsymbol{Z}_k = \boldsymbol{U}_k \boldsymbol{\Lambda}_k \boldsymbol{U}_k^H \tag{59}$$

where

$$\boldsymbol{U}_k \triangleq [\boldsymbol{U}_{k+}, \boldsymbol{U}_{k-}, \boldsymbol{U}_{k\perp}];\ \boldsymbol{\Lambda}_k \triangleq \begin{bmatrix} \boldsymbol{\Lambda}_{k+} & 0 & 0 \\ 0 & -\boldsymbol{\Lambda}_{k-} & 0 \\ 0 & 0 & \boldsymbol{\Lambda}_{k\perp} \end{bmatrix}, \tag{60}$$

the matrices $\boldsymbol{\Lambda}_{k+} \succeq 0$ and $\boldsymbol{\Lambda}_{k-} \succeq 0$ contain the positive and absolute values of the negative eigenvalues of $\boldsymbol{A}_k + \boldsymbol{Z}_k^{||}$ on the main diagonal. The matrices containing the corresponding eigenvectors are denoted by $\boldsymbol{U}_{k+}$ and $\boldsymbol{U}_{k-}$, respectively. Similarly, the matrices $\boldsymbol{\Lambda}_{k\perp} \succeq 0$ and $\boldsymbol{U}_{k\perp}$ contain the eigenvalues and corresponding eigenvectors of $\boldsymbol{Z}_k^{\perp}$, respectively.

In the following we want to prove by contradiction. Assuming first that $\boldsymbol{Z}_k^\star = 0$, we can reformulate the first constraint in (27) as

$$\begin{aligned}
&-\alpha_k \|\boldsymbol{A}_k\|_{\boldsymbol{Q}_k} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \boldsymbol{A}_k\} - \sigma_k^2 \gamma_k \\
&= -\alpha_k \sqrt{\mathrm{vec}(\boldsymbol{A}_k)^H \boldsymbol{Q}_k \mathrm{vec}(\boldsymbol{A}_k)} - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \boldsymbol{A}_k\} - \sigma_k^2 \gamma_k \\
&= -\alpha_k \sqrt{\mathrm{vec}\left(\boldsymbol{U}_k \boldsymbol{\Lambda}_k \boldsymbol{U}_k^H\right)^H \boldsymbol{Q}_k \mathrm{vec}\left(\boldsymbol{U}_k \boldsymbol{\Lambda}_k \boldsymbol{U}_k^H\right)} \\
&\quad - \mathrm{Tr}\left\{\hat{\boldsymbol{R}}_k \boldsymbol{U}_k \boldsymbol{\Lambda}_k \boldsymbol{U}_k^H\right\} - \sigma_k^2 \gamma_k \\
&= -\alpha_k \sqrt{\mathrm{vec}(\boldsymbol{\Lambda}_k)^H (\boldsymbol{U}_k^* \otimes \boldsymbol{U}_k)^H \boldsymbol{Q}_k (\boldsymbol{U}_k^* \otimes \boldsymbol{U}_k) \mathrm{vec}(\boldsymbol{\Lambda}_k)} \\
&\quad - \mathrm{Tr}\left\{\hat{\boldsymbol{R}}_k \boldsymbol{U}_{k+} \boldsymbol{\Lambda}_{k+} \boldsymbol{U}_{k+}^H\right\} + \mathrm{Tr}\left\{\hat{\boldsymbol{R}}_k \boldsymbol{U}_{k-} \boldsymbol{\Lambda}_{k-} \boldsymbol{U}_{k-}^H\right\} \\
&\quad - \mathrm{Tr}\left\{\hat{\boldsymbol{R}}_k \boldsymbol{U}_{k\perp} \boldsymbol{\Lambda}_{k\perp} \boldsymbol{U}_{k\perp}^H\right\} - \sigma_k^2 \gamma_k \geq 0.
\end{aligned} \tag{61}$$

Now we consider that $\boldsymbol{Z}_k^\star \succeq 0$. If at optimum $\boldsymbol{Z}_k^\star$ has any rank-one component in the direction of a column of $\boldsymbol{U}_{k+}$, then the corresponding $\boldsymbol{\Lambda}_{k+}^\star$ would have increased entries as compared to the case where $\boldsymbol{Z}_k^\star = 0$. The non-zero component in $\boldsymbol{Z}_k^\star$ would thus result in a loosened QoS constraint, as can be observed from the last inequality in (61). This contradicts optimality as we could then always find a solution with a reduced cost by scaling down $\boldsymbol{W}_k^\star$ without violating other constraints. This argument naturally extends to the case of higher rank components in the space spanned by $\boldsymbol{U}_{k+}$. Therefore, we conclude that $\boldsymbol{Z}_k^\star$ has no components in the space spanned by $\boldsymbol{U}_{k+}$.

Next, if an arbitrary rank-one component of $\boldsymbol{Z}_k^\star$ in the direction of a column of $\boldsymbol{U}_{k-}$ is nonzero at optimum, then we could always find a different matrix $\boldsymbol{Z}_k^{\star\prime}$ similar as $\boldsymbol{Z}_k^\star$, but with that particular eigenvalue of $\boldsymbol{Z}_k^\star$ set to zero. In this case we could find an optimum counterpart $\boldsymbol{W}_k^{\star\prime}$ similar to $\boldsymbol{W}_k^\star$ where the corresponding component is reduced accordingly such that $\boldsymbol{Z}_k^{\star\prime} - \boldsymbol{W}_k^{\star\prime} = \boldsymbol{Z}_k^\star - \boldsymbol{W}_k^\star$. The resulting $\boldsymbol{W}_k^{\star\prime}$ would yield a reduced objective function as $\mathrm{Tr}\{\boldsymbol{W}_k^{\star\prime}\} < \mathrm{Tr}\{\boldsymbol{W}_k^\star\}$ without violating other constraints in (27). This contradicts to the optimality assumption. This clearly can be extended for higher rank components and we conclude that $\boldsymbol{Z}_k^\star$ has no components in the space spanned by $\boldsymbol{U}_{k-}$.

From the last inequality in (61), it can be observed that any component of $\boldsymbol{Z}_k^\star$ orthogonal to $\boldsymbol{A}_k$, i.e., $\boldsymbol{Z}_k^{\perp}$, would only make the first and fourth terms more negative. This implies that at optimum all entries of $\boldsymbol{\Lambda}_{k\perp}$ must be zero as otherwise these could be reduced, resulting in a loosened QoS constraint. Again, this contradicts optimality as we could always find a solution with a reduced cost by scaling down $\boldsymbol{W}_k^\star$ without violating other constraints. □

Appendix B
Example With Higher Rank Solution

In this Appendix, we show that, for the covariance matrices in (54), there exist conditions under which the solution of (28) has higher rank. For simplicity, we consider the non-weighted Frobenius norm to model the error matrices, i.e., $\boldsymbol{Q}_k = \boldsymbol{I}$. In this case, the inequality QoS constraints for the first and second users in (28), respectively, become

$$-\alpha_1 \|-\boldsymbol{W}_1 + \gamma_1 \boldsymbol{W}_2\| + \mathrm{Tr}\{\hat{\boldsymbol{R}}_1 \boldsymbol{W}_1\} - \gamma_1 \mathrm{Tr}\{\hat{\boldsymbol{R}}_1 \boldsymbol{W}_2\} - \sigma_1^2 \gamma_1 \geq 0 \tag{62}$$

$$-\alpha_2 \|-\boldsymbol{W}_2 + \gamma_2 \boldsymbol{W}_1\| + \mathrm{Tr}\{\hat{\boldsymbol{R}}_2 \boldsymbol{W}_2\} - \gamma_2 \mathrm{Tr}\{\hat{\boldsymbol{R}}_2 \boldsymbol{W}_1\} - \sigma_2^2 \gamma_2 \geq 0. \tag{63}$$

Let the solution of (28) be

$$\boldsymbol{W}_1 = \begin{pmatrix} a_{11} & a_{21}^* & a_{31}^* \\ a_{21} & a_{22} & a_{32}^* \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \boldsymbol{W}_2 = \begin{pmatrix} b_{11} & b_{21}^* & b_{31}^* \\ b_{21} & b_{22} & b_{32}^* \\ b_{31} & b_{32} & b_{33} \end{pmatrix}. \tag{64}$$

Then the cost function in (28) is given by

$$\mathrm{Tr}\{\boldsymbol{W}_1 + \boldsymbol{W}_2\} = a_{11} + a_{22} + a_{33} + b_{11} + b_{22} + b_{33} \tag{65}$$

and the terms in the constraint (62) become

$$\begin{aligned}
\mathrm{Tr}\{\hat{\boldsymbol{R}}_1 \boldsymbol{W}_1\} &= r_1 a_{11} + r_2 a_{22}, \\
\mathrm{Tr}\{\hat{\boldsymbol{R}}_1 \boldsymbol{W}_2\} &= r_1 b_{11} + r_2 b_{22},
\end{aligned} \tag{66}$$

and

$$\|-\boldsymbol{W}_1+\gamma_1\boldsymbol{W}_2\| = \left\| \begin{matrix} -a_{11}+\gamma_1 b_{11} & -a_{21}^*+\gamma_1 b_{21}^* & -a_{31}^*+\gamma_1 b_{31}^* \\ -a_{21}+\gamma_1 b_{21} & -a_{22}+\gamma_1 b_{22} & -a_{32}^*+\gamma_1 b_{32}^* \\ -a_{31}+\gamma_1 b_{31} & -a_{32}+\gamma_1 b_{32} & -a_{33}+\gamma_1 b_{33} \end{matrix} \right\|. \quad (67)$$

The off-diagonal entries of $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ do not affect $\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_1\}$, $\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_2\}$ and the cost function (65). However, these entries make the norm term (67) larger, which is required to be as small as possible. Similarly, it can also be shown for the second QoS constraint (63) that the off-diagonal entries of $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ do not affect $\mathrm{Tr}\{\hat{\boldsymbol{R}}_2\boldsymbol{W}_1\}$ and $\mathrm{Tr}\{\hat{\boldsymbol{R}}_2\boldsymbol{W}_2\}$, but make the norm term $\|-\boldsymbol{W}_2+\gamma_2\boldsymbol{W}_1\|$ larger. Therefore, at the optimum, the off-diagonal entries of matrices $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ must compensate each other, to reduce the terms (67) and $\|-\boldsymbol{W}_2+\gamma_2\boldsymbol{W}_1\|$. This implies that, if possible, we should have

$$-a_{ij}+\gamma_1 b_{ij}=0; \quad i\neq j \quad (68)$$
$$-b_{ij}+\gamma_2 a_{ij}=0; \quad i\neq j. \quad (69)$$

One solution of the (68) and (69) is $\gamma_1=\gamma_2=1$ with $a_{ij}=b_{ij}$. For general γ_1 and γ_2, both (68) and (69) are satisfied only if $a_{ij}=b_{ij}=0$ for $i\neq j$. Therefore, the solution has the form

$$\boldsymbol{W}_1=\mathrm{diag}\{a_{11},a_{22},a_{33}\}, \quad \boldsymbol{W}_2=\mathrm{diag}\{b_{11},b_{22},b_{33}\}. \quad (70)$$

The QoS constraints (62) and (63), respectively, are then modified as

$$\begin{aligned} &-\alpha_1\|-\boldsymbol{W}_1+\gamma_1\boldsymbol{W}_2\|+\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_1\}-\gamma_1\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_2\}-\sigma_1^2\gamma_1 \\ &\quad=-\alpha_1\sqrt{(a_{11}-\gamma_1 b_{11})^2+(a_{22}-\gamma_1 b_{22})^2+(a_{33}-\gamma_1 b_{33})^2} \\ &\qquad+r_1(a_{11}-\gamma_1 b_{11})+r_2(a_{22}-\gamma_1 b_{22})-\sigma_1^2\gamma_1\geq 0 \quad (71) \\ &-\alpha_2\|-\boldsymbol{W}_2+\gamma_2\boldsymbol{W}_1\|+\mathrm{Tr}\{\hat{\boldsymbol{R}}_2\boldsymbol{W}_2\}-\gamma_2\mathrm{Tr}\{\hat{\boldsymbol{R}}_2\boldsymbol{W}_1\}-\sigma_2^2\gamma_2 \\ &\quad=-\alpha_2\sqrt{(b_{11}-\gamma_2 a_{11})^2+(b_{22}-\gamma_2 a_{22})^2+(b_{33}-\gamma_2 a_{33})^2} \\ &\qquad+r_3(b_{33}-\gamma_2 a_{33})-\sigma_2^2\gamma_2\geq 0. \quad (72) \end{aligned}$$

From (71), it is clear that when (28) is feasible, either a_{11} or a_{22} must be positive as these are the only positive terms on the left side of (71). Similarly from (72), the term b_{33} must also be positive. Thus, for the case when, at optimum, either of the terms $a_{33}^\star$, $b_{11}^\star$, and $b_{22}^\star$ become non-zero, a higher rank solution is obtained for (28).

Let us further investigate the case when $a_{33}^\star=0$, $b_{11}^\star=0$, and $b_{22}^\star=0$. We first prove that for $\gamma_1\gamma_2\leq 1$, we have $a_{33}^\star=0$, $b_{11}^\star=0$, and $b_{22}^\star=0$ at the optimum. Towards this aim, let us assume that, at the optimum, $a_{33}^\star\neq 0$, $b_{11}^\star\neq 0$ and $b_{22}^\star\neq 0$. If we choose $\bar{b}_{11}=b_{11}^\star-\delta_b$ and $\bar{a}_{11}=a_{11}^\star-\gamma_1\delta_b$, for some $\delta_b>0$, and replace $a_{11}^\star$ by $\bar{a}_{11}$ and $b_{11}^\star$ by $\bar{b}_{11}$, then the value of (71) remains unchanged while the cost function (65) decreases. Next, we note that

$$\begin{aligned} \bar{b}_{11}-\gamma_2\bar{a}_{11}&=b_{11}^\star-\delta_b-\gamma_2 a_{11}^\star+\gamma_1\gamma_2\delta_b \\ &=b_{11}^\star-\gamma_2 a_{11}^\star+(\gamma_1\gamma_2-1)\delta_b. \end{aligned} \quad (73)$$

If we have $\gamma_1\gamma_2\leq 1$, then $\bar{b}_{11}-\gamma_2\bar{a}_{11}\leq b_{11}^\star-\gamma_2 a_{11}^\star$. From (72) we observe that $\boldsymbol{W}_1=\mathrm{diag}\{\bar{a}_{11},a_{22}^\star,a_{33}^\star\}$ and $\boldsymbol{W}_2=\mathrm{diag}\{\bar{b}_{11},b_{22}^\star,b_{33}^\star\}$ denote a feasible point with lower value of cost function (65) as compared to $\boldsymbol{W}_1=\mathrm{diag}\{a_{11}^\star,a_{22}^\star,a_{33}^\star\}$ and $\boldsymbol{W}_2=\mathrm{diag}\{b_{11}^\star,b_{22}^\star,b_{33}^\star\}$. This contradicts optimality and, therefore, we can conclude that at optimum $b_{11}^\star=0$ for $\gamma_1\gamma_2\leq 1$. Using similar arguments, it can easily be shown that both $b_{22}^\star=0$ and $a_{33}^\star=0$ at the optimum in this scenario. Therefore, for $\gamma_1\gamma_2\leq 1$, the solution has the form

$$\boldsymbol{W}_1=\mathrm{diag}\{a_{11},a_{22},0\}, \quad \boldsymbol{W}_2=\mathrm{diag}\{0,0,b_{33}\}. \quad (74)$$

We next derive conditions under which, for $a_{33}^\star=b_{11}^\star=b_{22}^\star=0$, we still obtain higher rank solutions. From (74) we can see that $\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_2\}=0$. Let us assume that at optimum $\mathrm{Tr}\{\boldsymbol{W}_1\}=c$ where c is a constant. Then, we can introduce the parameterization $a_{11}+a_{22}=c$ such that

$$a_{11}=\mu c, \quad a_{22}=(1-\mu)c; \quad 0\leq\mu\leq 1. \quad (75)$$

From (74), it is obvious that only $\mu=0$ or 1 corresponds to a rank-one solution. The value of μ can affect both the norm term and the term $\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_1\}$. Using (74) and the parameterization (75), the QoS constraint (62) modifies as

$$\begin{aligned} &-\alpha_1\|-\boldsymbol{W}_1+\gamma_1\boldsymbol{W}_2\|+\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_1\}-\gamma_1\mathrm{Tr}\{\hat{\boldsymbol{R}}_1\boldsymbol{W}_2\}-\sigma_1^2\gamma_1 \\ &\quad=-\alpha_1\sqrt{(2\mu^2+1-2\mu)c^2+\gamma_1^2 b_{33}^2}+((r_1-r_2)\mu+r_2)c-\sigma_1^2\gamma_1 \\ &\quad\geq 0. \end{aligned} \quad (76)$$

Similarly, the QoS constraint (63) becomes

$$\begin{aligned} &-\alpha_1\|-\boldsymbol{W}_2+\gamma_2\boldsymbol{W}_1\|+\mathrm{Tr}\{\hat{\boldsymbol{R}}_2\boldsymbol{W}_2\}-\gamma_2\mathrm{Tr}\{\hat{\boldsymbol{R}}_2\boldsymbol{W}_1\}-\sigma_2^2\gamma_2 \\ &\quad=-\alpha_2\sqrt{\gamma_2^2(2\mu^2+1-2\mu)c^2+b_{33}^2}+r_3 b_{33}-\sigma_2^2\gamma_2\geq 0. \end{aligned} \quad (77)$$

It can easily be shown that at optimum both the QoS constraints (62) and (63) will be satisfied at equality. Therefore, setting (76) and (77) equal to zero we get the optimal $b_{33}^\star$ and $c^\star$, respectively, as

$$b_{33}^\star=\sqrt{\frac{\left(((r_1-r_2)\mu+r_2)c-\sigma_1^2\gamma_1\right)^2-\alpha_1^2(2\mu^2+1-2\mu)}{\alpha_1^2\gamma_1^2}} \quad (78)$$

$$c^\star=\sqrt{\frac{r_3^2 b_{33}^2+\sigma_2^4\gamma_2^2-2r_3 b_{33}\sigma_2^2\gamma_2-\alpha_2^2 b_{33}^2}{\alpha_2^2\gamma_2^2(2\mu^2+1-2\mu)}}. \quad (79)$$

Next, inserting (79) in (76) and (78) in (77), and then setting both (76) and (77) to zero, respectively, we get the decoupled equations as

$$-\alpha_1\sqrt{(2\mu^2+1-2\mu)c^{\star 2}+\gamma_1^2 b_{33}^{\star 2}}+((r_1-r_2)\mu+r_2)c^\star-\sigma_1^2\gamma_1=0 \quad (80)$$

$$-\alpha_2\sqrt{\gamma_2^2(2\mu^2+1-2\mu)c^{\star 2}+b_{33}^{\star 2}}+r_3 b_{33}^\star-\sigma_2^2\gamma_2=0. \quad (81)$$

In order to find the optimal $\mu^\star$, we numerically solve both (80) and (81) by varying μ from 0 to 1, and then choose the μ for which the transmitted power $c+b_{33}$ is minimized. In Fig. 12, we plot the value of the cost function for different values of α (we consider $\alpha_1=\alpha_2=\alpha$) as μ is varied from 0 to 1. For this plot we use $r_1=1.2$, $r_2=1$, $r_3=1$, $\gamma_1=\gamma_2=1$ and $\sigma_1^2=\sigma_2^2=1$. The minimum for each curve is marked with a circle. As can be observed from Fig. 12, the optimal parameter $\mu^\star$ varies from 1 (rank-one solution) towards 0.5 (higher rank solution) as the uncertainty parameter α is increased. This implies that for larger values of error, higher rank solutions will be favored. In Fig. 13, the value of r_1 is varied from 1 to 1.8 with $\alpha=0.4$. In this case the optimal $\mu^\star$ varies from 0.5 towards 1 as r_1 is increased, implying that as the covariance matrices become less symmetric, the rank-one solution is favored.

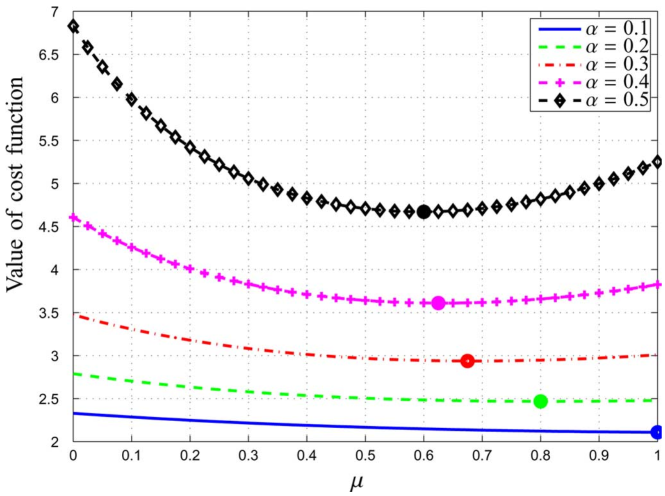

Fig. 12. Value of cost function vs μ with varying α.

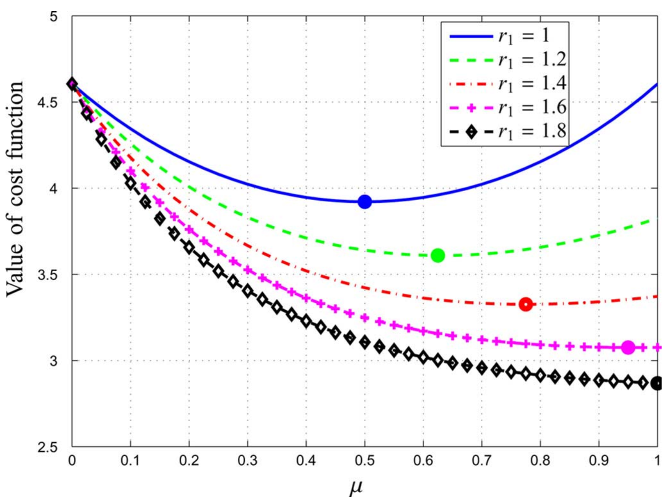

Fig. 13. Value of cost function vs μ with varying r_1.

APPENDIX C
GENERALIZATION OF SOLUTIONS

For simplicity, we consider the case that the non-weighted Frobenius norm is used to model the error matrices. Let us assume that $\boldsymbol{W}_k^\star$ $(k = 1, \ldots, K)$ is the solution of the problem (34) with covariance matrices $\hat{\boldsymbol{R}}_k$ $(k = 1, \ldots, K)$. Keeping the remaining parameters of problem (34) unchanged, we introduce a new set of covariance matrices

$$\tilde{\boldsymbol{R}}_k = \boldsymbol{U}^H \hat{\boldsymbol{R}}_k \boldsymbol{U};\ k = 1, \ldots, K \tag{82}$$

where $\boldsymbol{U}$ is a unitary matrix. The QoS and PU interference constraints of (34), respectively, become

$$-\alpha_k \|\boldsymbol{A}_k\| - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \boldsymbol{U} \boldsymbol{A}_k \boldsymbol{U}^H\} - \sigma_k^2 \gamma_k \geq 0 \tag{83}$$

$$\mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l} \boldsymbol{U} \boldsymbol{C} \boldsymbol{U}^H\} + \alpha_{K+l} \|\boldsymbol{C}\| \leq \varepsilon_l. \tag{84}$$

Defining the matrices

$$\tilde{\boldsymbol{W}}_k \triangleq \boldsymbol{U} \boldsymbol{W}_k \boldsymbol{U}^H; \quad \tilde{\boldsymbol{A}}_k \triangleq \boldsymbol{U} \boldsymbol{A}_k \boldsymbol{U}^H; \quad \tilde{\boldsymbol{C}} \triangleq \boldsymbol{U} \boldsymbol{C} \boldsymbol{U}^H, \tag{85}$$

we obtain that $\|\tilde{\boldsymbol{A}}_k\| = \|\boldsymbol{A}_k\|$ and $\|\tilde{\boldsymbol{C}}\| = \|\boldsymbol{C}\|$. Also, the cost function remains unchanged as $\sum_{i=1}^K \mathrm{Tr}\{\tilde{\boldsymbol{W}}_i\} = \sum_{i=1}^K \mathrm{Tr}\{\boldsymbol{W}_i\}$. We can modify the QoS constraints (83) and the PU interference constraints (84) as

$$-\alpha_k \|\tilde{\boldsymbol{A}}_k\| - \mathrm{Tr}\{\hat{\boldsymbol{R}}_k \tilde{\boldsymbol{A}}_k\} - \sigma_k^2 \gamma_k \geq 0 \tag{86}$$

$$\mathrm{Tr}\{\hat{\boldsymbol{R}}_{K+l} \tilde{\boldsymbol{C}}\} + \alpha_{K+l} \|\tilde{\boldsymbol{C}}\| \leq \varepsilon_l. \tag{87}$$

This means that the solution for the problem with the modified covariance matrices (82) is given by $\boldsymbol{U}^H \boldsymbol{W}_k^\star \boldsymbol{U}$.

ACKNOWLEDGMENT

We dedicate this paper to our advisor, colleague, and friend Prof. Alex B. Gershman, who unfortunately passed away during the completion of this paper.

REFERENCES

[1] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.

[2] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

[3] 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.

[4] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Annu. l Allerton Conf. Commun., Control, Comput.*, Sept. 1999, pp. 987–996.

[5] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[6] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?," *IEEE Commun. Mag.*, vol. 42, pp. 54–59, Oct. 2004.

[7] M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, pp. 714–724, Dec. 2007.

[8] A. Mutapcic, S.-J. Kim, and S. Boyd, "A tractable method for robust downlink beamforming in wireless communications," in *Proc. ACSSC*, Pacific Grove, CA, USA, Nov. 2007, pp. 1224–1228.

[9] N. Vucic and H. Boche, "Downlink precoding for multiuser MISO systems with imperfect channel knowledge," in *Proc. ICASSP*, Las Vegas, NV, USA, Mar. 2008, pp. 3121–3124.

[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, Aug. 2001, ch. 18.

[11] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust downlink power control in wireless cellular systems," *EURASIP J. Wireless Commun. Netw.*, no. 2, pp. 261–272, Dec. 2004.

[12] I. Wajid, Y. C. Eldar, and A. B. Gershman, "Robust downlink beamforming using covariance channel state information," in *Proc. ICASSP*, Taipei, Taiwan, Apr. 2009, pp. 2285–2288.

[13] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "A robust linear receiver for uplink multi-user MIMO systems based on probability-constrained optimization and second-order cone programming," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Jul. 2004, pp. 153–157.

[14] B. K. Chalise and A. Czylwik, "Robust uplink beamforming based upon minimum outage probability criterion," in *Proc. Global Telecommun. Conf. (IEEE GLOBECOM)*, Dec. 2004, vol. 6, pp. 3974–3978.

[15] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust adaptive beamforming using probability-constrained optimization," in *Proc. IEEE Workshop Stat. Signal Process.*, July 2005, pp. 934–939.

[16] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman, "Robust downlink beamforming based on outage probability specifications," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3498–3503, Oct. 2007.

[17] S. A. Vorobyov, Y. Rong, and A. B. Gershman, "Robust minimum variance adaptive beamformers and multiuser MIMO receivers: From the worst-case to probabilistically constrained Designs," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2006, vol. 5, p. V.

[18] S. A. Vorobyov, H. Chen, and A. B. Gershman, "On the relationship between robust minimum variance beamformers with probabilistic and worst-case distortionless response constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5719–5724, Nov. 2008.

[19] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femto networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[20] D. Oh, H. Lee, and Y. Lee, "Power control and beamforming for femtocells in the presence of channel uncertainty," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2545–2554, Jun. 2011.

[21] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[22] G. Scutari and D. P. Palomar, "MIMO cognitive radio: A game theoretical approach," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 761–780, Feb. 2010.
[23] M. H. Islam, Y.-C. Liang, and A. T. Hoang, "Joint power control and beamforming for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2415–2419, Jul. 2008.
[24] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint beamforming and power control for multiple access channels in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 38–51, Jan. 2008.
[25] X. Fu, J. Wang, and S. Li, "Joint power management and beamforming for base stations in cognitive radio systems," in *Proc. 6th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Sep. 2009, pp. 403–407.
[26] M. Pesavento, D. Ciochina, and A. B. Gershman, "Iterative dual downlink beamforming for cognitive radio networks," presented at the CROWNCOM, Cannes, France, Jun. 2010.
[27] K. Cumanan, R. Krishna, V. Sharma, and S. Lambotharan, "Robust interference control techniques for multiuser cognitive radios using worst-case performance optimization," in *Proc. 42nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2008, pp. 378–382.
[28] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust downlink beamforming in multiuser MISO cognitive radio networks with imperfect channel-state information," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2852–2860, Jul. 2010.
[29] G. Zheng, K. K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, pp. 4871–4881, Dec. 2009.
[30] M. F. Hanif, P. J. Smith, and M.-S. Alouini, "SINR balancing in the downlink of cognitive radio networks with imperfect channel knowledge," presented at the CROWNCOM, Cannes, France, Jun. 2010.
[31] I. Wajid, M. Pesavento, Y. C. Eldar, and A. B. Gershman, "Robust downlink beamforming for cognitive radio networks," presented at the IEEE GLOBECOM, Miami, FL, USA, Dec. 2010.
[32] I. Wajid, H. Nikolaeva, and M. Pesavento, "Iterative algorithms for robust downlink beamforming in cognitive radio networks," presented at the CROWNCOM, Osaka, Japan, Jun. 2011.
[33] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.
[34] A. Nemirovski, Lectures on Modern Convex Optimization, Class Notes Georgia Inst. Technol., Atlanta, GA, USA, 2005.
[35] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11–12, pp. 625–653, Aug. 1999.
[36] M. Grant, S. Boyd, and Y. Ye, "CVX: MATLAB software for disciplined convex programming," Jun. 2009 [Online]. Available: http://www.stanford.edu/boyd/cvx/V.1.0RC3, Web page and software.
[37] S. Zhang, "Quadratic maximization and semidefinite relaxation," *Math. Program.*, ser. A, vol. 87, pp. 453–465, 2000.

Imran Wajid received the B.Eng. degree in information technology from the Hamdard University Karachi, Islamabad Campus, Pakistan in 2002, and the M.S. degree in electronic engineering from the Muhammad Ali Jinnah University, Islamabad, Pakistan in 2005. He is currently pursuing the Ph.D. degree in the Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt, Germany. His research interests lie in the areas of wireless communication, signal processing, convex optimization, and robust algorithms. He has been awarded the Higher Education Commission of Pakistan (HEC)/German Academic Exchange Service (DAAD) scholarship for his Ph.D. studies. He also received Ministry of Science and Technology (MOST) scholarship during his Bachelor's Studies.

Marius Pesavento (M'00) received the Dipl.-Ing. and M.Eng. degree from Ruhr-Universität Bochum, Germany, and McMaster University, Hamilton, ON, Canada, in 1999 and 2000, respectively, and the Dr.-Ing. degree in electrical engineering from Ruhr-Universität Bochum, Germany, in 2005. Between 2005 and 2007, he was a Research Engineer at FAG Industrial Services GmbH, Aachen, Germany. From 2007 to 2009, he was the Director of the Signal Processing Section at mimoOn GmbH, Duisburg, Germany. In 2010, he became Assistant Professor and in 2013 Full Professor at the Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt, Darmstadt, Germany, where he is the head of the Communication Systems Group. His research interests are in the area of robust signal processing and adaptive beamforming, high-resolution sensor array processing, resource allocation and transceiver design for multiantenna, multiuser and multicarrier wireless communication systems, cooperative communication techniques in multiway relay networks, distributed space-time coding, convex optimization and mixed-integer nonliner programming for signal processing and communications, statistical signal processing, spectral analysis, parameter estimation, and detection theory. Dr. Pesavento was a recipient of the 2003 ITG/VDE Best Paper Award, the 2005 Young Author Best Paper Award of the IEEE Signal Processing Society, and the 2010 Best Paper Award of the CROWNCOM conference. He is a member of the Editorial board of the *EURASIP Signal Processing Journal* and the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society (SPS).

Yonina C. Eldar (S'98–M'02–SM'07–F'12) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering both from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2002.

From January 2002 to July 2002, she was a Postdoctoral Fellow at the Digital Signal Processing Group at MIT. She is currently a Professor in the Department of Electrical Engineering at the Technion Israel Institute of Technology, Haifa and holds the The Edwards Chair in Engineering. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT and a Visiting Professor at Stanford University, Stanford, CA. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, optimization methods, and their applications to biology and optics.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in electrical engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002 to 2005, she was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005 the Andre and Bella Meyer Lectureship, in 2007 the Henry Taub Prize for Excellence in Research, in 2008 the Hershel Rich Innovation Award, the Award for Women with Distinguished Contributions, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion Outstanding Lecture Award, in 2009 the Technion's Award for Excellence in Teaching, in 2010 the Michael Bruno Memorial Award from the Rothschild Foundation, and in 2011 the Weizmann Prize for Exact Sciences. In 2012 she was elected to the Young Israel Academy of Science and to the Israel Committee for Higher Education, and elected an IEEE Fellow. In 2013 she received the Technion's Award for Excellence in Teaching, and the Hershel Rich Innovation Award. She received several best paper awards together with her research students and colleagues. She is a Signal Processing Society Distinguished Lecturer, a member of the IEEE Bio Imaging Signal Processing technical committee, an Associate Editor for the *SIAM Journal on Imaging Sciences*, and Editor in Chief of *Foundations and Trends in Signal Processing*. In the past, she was a member of the IEEE Signal Processing Theory and Methods technical committee, and served as an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the *EURASIP Journal of Signal Processing*, and the *SIAM Journal on Matrix Analysis and Applications*. She is co-author of the books "Compressed Sensing" and "Convex Optimization Methods in Signal Processing and Communications" both published by Cambridge University Press.

Dana Ciochina (S'09) received her Dipl.-Ing. degree from University 'Politehnica' Bucharest in 2008. Between 2007 and 2009 she was with the Department of Access Technologies in Orange Romania. In 2009 she joined the Communication Systems Group at Technische Universität Darmstadt, where she is currently pursuing her Ph.D. She was awarded with scholarships from Infineon Technologies and Orange Romania during her graduate studies and co-authored the paper that received the Best Paper Award of the CROWNCOM conference in 2010. Her research interests are in the area of robust signal processing, adaptive beamforming, resource allocation and optimization techniques for signal processing and wireless communication.