Chance-Constrained Robust Beamforming for Multi-Cell Coordinated Downlink

Chao Shen*, Tsung-Hui Chang†, Kun-Yu Wang‡, Zhengding Qiu*, and Chong-Yung Chi‡

*Institute of Information Science, Beijing Jiaotong University, Beijing, China 100044

†Department of Elect. and Computer Eng., University of California, Davis, California, USA 95616

‡Institute of Commun. Eng. & Department of Elect. Eng., National Tsing Hua University, Hsinchu, Taiwan 30013

***Abstract*—This paper considers robust multi-cell coordinated beamforming (MCBF) design for downlink wireless systems, in the presence of channel state information (CSI) errors. By assuming that the CSI errors are complex Gaussian distributed, we formulate a chance-constrained robust MCBF design problem which guarantees that the mobile stations can achieve the desired signal-to-interference-plus-noise ratio (SINR) requirements with a high probability. A convex approximation method, based on semidefinite relaxation and tractable probability approximation formulations, is proposed. The goal is to solve the convex approximation formulation in a distributed manner, with only a small amount of information exchange between base stations. To this end, we develop a distributed implementation by applying a convex optimization method, called *weighted variable-penalty alternating direction method of multipliers* (WVP-ADMM), which is numerically more stable and can converge faster than the standard ADMM method. Simulation results are presented to examine the chance-constrained robust MCBF design and the proposed distributed implementation algorithm.**

***Index Terms*—Multicell coordinated beamforming, robust beamforming, chance constraint, outage probability, distributed beamforming.**

I. Introduction

Multi-cell coordinated beamforming (MCBF) design has been of great interest in recent years since it can effectively manage the inter-cell interference (ICI) and improve the throughput of the multi-cell systems; see, e.g., [1], [2]. Most of the existing works assume that the base stations (BSs) have the perfect channel state information (CSI) of the mobile stations (MSs). However, in practical scenarios, it is inevitable to have CSI errors due to, e.g., the imperfect channel estimation and finite rate feedback. In order to provide guaranteed quality of service (QoS) (e.g., the signal-to-interference-plus-noise ratio (SINR)) for the MSs, robust MCBF designs that explicitly account for the CSI errors have been studied. For example, in [3], [4], the CSI errors were modeled as deterministic vectors within a bounded uncertainty region and worst-case robust MCBF designs were investigated.

In this paper, considering the stochastic nature of the CSI errors, we model the CSI errors as complex Gaussian random vectors, and study a chance-constrained robust MCBF design problem. The problem formulation aims to minimize the sum power of all BSs subject to constraints that the SINR requirements of all MSs must be satisfied with a preassigned, usually high probability. However, the associated optimization problem is difficult to handle because the SINR constraints are not convex and, moreover, the probability functions have no tractable expression. Such a chance-constrained robust design problem has only been studied in the single-cell scenario [5]–[7]. In particular, effective convex approximations of the probability constraint were proposed in [7], using the semidefinite relaxation (SDR) technique [8] and the idea of safe tractable approximation [9]. It has been verified in [6] that the presented approximation method outperforms the existing methods. In this paper, we extend the approximation method in [6] to the considered chance-constrained robust MCBF design problem.

Our focus in this paper is on distributed optimization methods, where each BS optimizes only the beamforming vectors for its associated MSs in the serving region, using only local CSI and with a small amount of message exchange between BSs. There have been considerable works for distributed optimization of MCBF designs with perfect CSI; see [1], [10]–[12]. In [3], [4], distributed optimization methods for the worst-case robust MCBF design problem has also been reported. In particular, in [4], the authors proposed the use of the distributed convex optimization method known as alternating direction method of multipliers (ADMM) [13]. It is shown that ADMM can avoid some unboundedness issue occurred in the robust MCBF design and is more numerically stable than the dual decomposition method used in [3], [11]. However, ADMM may converge slowly especially when the problem is ill-conditioned. In this paper, we consider a modified ADMM scheme, called *weighted variable-penalty ADMM algorithm* (WVP-ADMM) [14], which employs weighted augmented penalty terms and thus provides more degrees of freedom to precondition the problem formulation. We show in the paper how the WVP-ADMM can be applied to the chance-constrained robust MCBF design problem in a distributed fashion. Simulation results are presented to demonstrate the effectiveness of the proposed methods.

Notation: $\mathbb{C}^n$ and $\mathbb{R}^n$ ($\mathbb{R}^n_-, \mathbb{R}^n_+$) stand for the sets of n-dimensional complex and real (nonpositive, nonnegative) vectors, respectively. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix, and $\mathbf{0}$ denotes an all-zero vector (matrix) with appropriate dimension. The superscripts $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\dagger$ represent the transpose, Hermitian (conjugate transpose) and pseudo inverse operations, respectively. $\mathrm{diag}(\cdot)$ denotes a diagonal matrix formed from its vector argument. $\Re(a)$ denotes the real part of a complex number a and $\mathrm{Tr}(\mathbf{A})$ represents the trace of matrix $\mathbf{A}$. $\mathbf{A} \succeq \mathbf{0}$ means that matrix $\mathbf{A}$ is positive semidefinite (PSD). $\|\mathbf{a}\|$ denotes the Euclidean norm of vector $\mathbf{a}$, and

978-1-4673-0921-9/12/$31.00 ©2012 IEEE

$\|\mathbf{a}\|_{\mathbf{G}}^2 \triangleq \mathbf{a}^H\mathbf{G}\mathbf{a}$, $\|\mathbf{A}\|_F$ is the Frobenius norm of matrix $\mathbf{A}$. For a variable a_{nmk}, where $n\in\{1,\ldots,N\}$, $m\in\{1,\ldots,M\}$ and $k\in\{1,\ldots,K\}$, $\{a_{nmk}\}_k$ denotes the set containing $a_{nm1},\ldots,a_{nmK}$; while $\{a_{nmk}\}$ denotes the set containing all possible a_{nmk}, i.e., $a_{111},\ldots,a_{11K},a_{121},\ldots,a_{NMK}$.

II. Signal model and Problem Statement

Consider a coordinated multi-cell downlink system with N_c cells. Each cell consists of a BS, which is equipped with N_t antennas, and K single-antenna MSs. The N_c BSs are assumed to operate over the same frequency band and communicate with their respective MSs by transmit beamforming. Let BS_n denote the BS in the nth cell, and MS_{nk} denote the kth MS in the nth cell, for $n\in\mathcal{N}_c\triangleq\{1,\ldots,N_c\}$ and $k\in\mathcal{K}\triangleq\{1,\ldots,K\}$. Let $\mathbf{w}_{nk}\in\mathbb{C}^{N_t}$ be the beamforming vector for MS_{nk}, and $\mathbf{h}_{nmk}\in\mathbb{C}^{N_t}$ denote the channel vector from BS_n to MS_{mk}. Moreover, assume that the information signals have the unit power. The signal-to-interference-plus-noise ratio (SINR) of MS_{nk} is given by [1], [4]

$$\begin{aligned}&\mathrm{SINR}_{nk}\left(\{\mathbf{w}_{mk}\}_{m,k},\{\mathbf{h}_{mnk}\}_m\right)\\&=\frac{\left|\mathbf{h}_{nnk}^H\mathbf{w}_{nk}\right|^2}{\sum_{i\neq k}^{K}\left|\mathbf{h}_{nnk}^H\mathbf{w}_{ni}\right|^2+\sum_{m\neq n}^{N_c}\sum_{i=1}^{K}\left|\mathbf{h}_{mnk}^H\mathbf{w}_{mi}\right|^2+\sigma_{nk}^2},\end{aligned}\quad(1)$$

where $\sigma_{nk}^2>0$ denotes the additive noise power at MS_{nk}.

The scenario considered here is that the BSs may not have perfect CSI, due to, e.g., imperfect channel estimation or limited feedback [15]. Specifically, the true channel vector $\mathbf{h}_{nmk}$ can be written as

$$\mathbf{h}_{nmk}=\hat{\mathbf{h}}_{nmk}+\mathbf{e}_{nmk}\ \forall n,m,k,\quad(2)$$

where $\{\hat{\mathbf{h}}_{nmk}\}_{m,k}$ are the channel estimates known to BS_n, and $\mathbf{e}_{nmk}\in\mathbb{C}^{N_t}$ is the unknown CSI error. In this paper, we assume that the CSI errors $\{\mathbf{e}_{nmk}\}$ are complex Gaussian distributed with zero-mean and covariance matrix $\mathbf{C}_{nmk}\succeq\mathbf{0}$, i.e., $\mathbf{e}_{nmk}\sim\mathcal{CN}(\mathbf{0},\mathbf{C}_{nmk})$ for all n,m and k.

Taking into account the CSI errors $\{\mathbf{e}_{nmk}\}$, our goal is to jointly design the beamforming vectors of all coordinated BSs so that each MS can achieve its desired SINR requirement *with a specified probability*. Specifically, we consider the following chance-constrained robust MCBF design

$$\min_{\substack{\mathbf{w}_{nk}\in\mathbb{C}^{N_t}\\\forall n,k}}\ \sum_{n=1}^{N_c}\sum_{k=1}^{K}\|\mathbf{w}_{nk}\|^2\quad(3a)$$

$$\begin{aligned}\text{s.t.}\ &\Pr\left\{\mathrm{SINR}_{nk}\left(\{\mathbf{w}_{mk}\},\{\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk}\}_m\right)\geq\gamma_{nk}\right\}\\&\geq 1-\rho_{nk}\ \forall n\in\mathcal{N}_c,\ k\in\mathcal{K},\end{aligned}\quad(3b)$$

where $\gamma_{nk}>0$ is the target SINR value of MS_{nk}, and $\rho_{nk}\in(0,1)$ is the maximum tolerable SINR outage probability. As seen, formulation (3) guarantees that each MS_{nk} achieves its target SINR value γ_{nk} with probability at least $1-\rho_{nk}$.

Problem (3) is difficult to solve because 1) the SINR constraint $\mathrm{SINR}_{nk}(\{\mathbf{w}_{mk}\},\{\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk}\}_m)\geq\gamma_{nk}$ is nonconvex in $\{\mathbf{w}_{mk}\}$, and 2) the probability function in (3b) has no tractable expression in general. In our previous work [6], we proposed an efficient convex approximation method for handling the chance-constrained beamforming design problem in a single-cell scenario ($N_c=1$). As we show in the next section, the developed method in [6] can also be applied to the robust MCBF problem (3) in spite of more involved formulations.

III. Tractable Convex Approximation for (3)

One of the key steps of the convex approximation method in [6] is to apply the SDR technique [8] to 'linearize' the nonconvex SINR constraints. By replacing $\mathbf{w}_{nk}\mathbf{w}_{nk}^H$ with a general PSD matrix $\mathbf{W}_{nk}$ for all n,k, the resultant rank-relaxed problem, which we call the SDR problem of (3), is given by

$$\min_{\substack{\mathbf{W}_{nk}\succeq\mathbf{0}\\\forall n,k}}\ \sum_{n=1}^{N_c}\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk})\quad(4a)$$

$$\begin{aligned}\text{s.t.}\ &\Pr\Bigg\{(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk})^H\mathbf{B}_{nk}(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk})+\\&\qquad\sum_{m\neq n}^{N_c}(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})^H\mathbf{D}_m(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})\geq\sigma_{nk}^2\Bigg\}\\&\geq 1-\rho_{nk}\quad\forall n\in\mathcal{N}_c,\ k\in\mathcal{K},\end{aligned}\quad(4b)$$

where $\mathbf{B}_{nk}\triangleq\gamma_{nk}^{-1}\mathbf{W}_{nk}-\sum_{i\neq k}^{K}\mathbf{W}_{ni}$ and $\mathbf{D}_m\triangleq-\sum_{i=1}^{K}\mathbf{W}_{mi}$. Note that the objective function of (4) and the arguments in the probability functions are all linear in $\{\mathbf{W}_{nk}\}$.

The second step of our convex approximation method is to use a conservative, but computational tractable (convex) formulation to approximate the probability function in (4b). To illustrate this, let us define the normalized CSI errors as

$$\mathbf{v}_{mnk}=\mathbf{C}_{mnk}^{-1/2}\mathbf{e}_{mnk}\sim\mathcal{CN}(\mathbf{0},\mathbf{I}_{N_t})\quad(5)$$

for all $m,n\in\mathcal{N}_c$, and $k\in\mathcal{K}$, where $\mathbf{C}_{mnk}^{1/2}\succeq\mathbf{0}$ is a PSD square root of $\mathbf{C}_{mnk}$ Further define the following notations

$$\begin{aligned}\mathbf{Q}_{nnk}&\triangleq\mathbf{C}_{nnk}^{1/2}\mathbf{B}_{nk}\mathbf{C}_{nnk}^{1/2}, &\mathbf{Q}_{mnk}&\triangleq\mathbf{C}_{mnk}^{1/2}\mathbf{D}_m\mathbf{C}_{mnk}^{1/2},\\\mathbf{u}_{nnk}&\triangleq\mathbf{C}_{nnk}^{1/2}\mathbf{B}_{nk}\hat{\mathbf{h}}_{nnk}, &\mathbf{u}_{mnk}&\triangleq\mathbf{C}_{mnk}^{1/2}\mathbf{D}_m\hat{\mathbf{h}}_{mnk},\\c_{nnk}&\triangleq\hat{\mathbf{h}}_{nnk}^H\mathbf{B}_{nk}\hat{\mathbf{h}}_{nnk}, &c_{mnk}&\triangleq\hat{\mathbf{h}}_{mnk}^H\mathbf{D}_m\hat{\mathbf{h}}_{mnk},\end{aligned}$$

for all n,k and $m\neq n$. Then we can express each of the probability functions in (4b) as

$$\begin{aligned}\Pr\Bigg\{&\sum_{m=1}^{N_c}(\mathbf{v}_{mnk}^H\mathbf{Q}_{mnk}\mathbf{v}_{mnk}+\\&2\Re(\mathbf{v}_{mnk}^H\mathbf{u}_{mnk})+c_{mnk})\geq\sigma_{nk}^2\Bigg\}\geq 1-\rho_{nk},\end{aligned}\quad(6)$$

which is a probability inequality of a quadratic form of complex Gaussian random variables. Since $\mathbf{Q}_{mnk}\not\succeq\mathbf{0}$, the above probability function does not have a closed-form expression in general. In [6], by applying a Bernstein-type inequality [16], we showed that a probability inequality like

(6) can actually be safely approximated by tractable convex formulations. Specifically, consider the following inequalities

$$\sqrt{\sum_{m=1}^{N_c}(\|\mathbf{Q}_{mnk}\|_F^2+2\|\mathbf{u}_{mnk}\|^2)} \leq \frac{1}{\sqrt{2\delta_{nk}}}\left(\sum_{m=1}^{N_c}\Big[\mathrm{Tr}(\mathbf{Q}_{mnk})+c_{mnk}\Big]-\delta_{nk}x_{nk}-\sigma_{nk}^2\right), \tag{7a}$$

$$x_{nk}\geq 0,\ x_{nk}\mathbf{I}_{N_t}+\mathbf{Q}_{mnk}\succeq\mathbf{0}, m=1,\ldots,N_c, \tag{7b}$$

where $\delta_{nk}\triangleq-\ln(\rho_{nk})$ and x_{nk} is a slack variable. Note that the inequalities in (7) are all convex. Then, by following the derivations in [6], one can show that (7) is a sufficient condition for (6), and therefore, the former can be used as a conservative approximation of the latter. By substituting (7) into (4), we obtain the following convex problem as an approximation to the robust MCBF problem in (3):

$$\min_{\substack{\mathbf{W}_{nk}\succeq\mathbf{0},x_{nk}\geq 0,\\ \forall n,k}} \sum_{n=1}^{N_c}\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk}) \tag{8a}$$

$$\text{s.t. } \sqrt{\sum_{m=1}^{N_c}(\|\mathbf{Q}_{mnk}\|_F^2+2\|\mathbf{u}_{mnk}\|^2)} \leq \frac{1}{\sqrt{2\delta_{nk}}} \cdot\left(\sum_{m=1}^{N_c}\Big[\mathrm{Tr}(\mathbf{Q}_{mnk})+c_{mnk}\Big]-\delta_{nk}x_{nk}-\sigma_{nk}^2\right)\forall n,k, \tag{8b}$$

$$x_{nk}\mathbf{I}_{N_t}+\mathbf{Q}_{mnk}\succeq\mathbf{0}\ \forall m,n,k. \tag{8c}$$

The resultant SDR problem (8) does not necessarily yield rank-one solutions of $\mathbf{W}_{nk}$. If the obtained optimal $\mathbf{W}_{nk}$ is of rank one for all n,k, then one can obtain a set of rank-one beamforming solutions by decomposition of $\mathbf{W}_{nk}=\mathbf{w}_{nk}\mathbf{w}_{nk}^H$ for all n,k. One can easily show that the obtained $\{\mathbf{w}_{nk}\}_{n,k}$ is feasible and is an conservative approximate solution to the original problem (3). In the cases that the obtained optimal $\mathbf{W}_{nk}$ is not of rank one, rank-one approximation methods [8] can be used for obtaining a rank-one approximate solution. It is worthwhile to mention that per-BS or per-antenna power constraints can be readily taken into account in our proposed method. Furthermore, the presented convex approximation method outperforms the existing methods, as we showed by simulations under the single-cell scenario in [6].

IV. Proposed Distributed Optimization Method

As mentioned in the introduction, our focus in this paper is on the distributed optimization methods. In particular, for solving (8), each BS$_n$ optimizes its own variables $\{\mathbf{W}_{nk}\}_k$ and $\{x_{nk}\}_k$, by using local CSI $\{\hat{\mathbf{h}}_{nmk},\mathbf{C}_{nmk}\}_{m,k}$ only and some information exchange with other BSs. However, such a distributed design is challenging because, in (8b) and (8c), the variables of all BSs are nontrivially coupled with each other. In this section, we first reformulate (8) into a compact form with linear coupled constraints, followed by applying a distributed optimization method, called WVP-ADMM [14], to solve (8) in a decentralized fashion.

A. Problem Reformulation

To reveal the intrinsic coupled structures of (8b) and (8c), let us introduce some slack variables $\{a_{nmk},b_{nmk}\}_{m,k}$ and p_n for each BS$_n$, and rewrite the constraints of (8) as

$$\begin{cases}\mathrm{Tr}(\mathbf{Q}_{nmk})+c_{nmk}=a_{nmk}\ \forall n,m,k, & (9a)\\ \sqrt{\|\mathbf{Q}_{nmk}\|_F^2+2\|\mathbf{u}_{nmk}\|^2}\leq b_{nmk}\ \forall n,m,k, & (9b)\\ \|[b_{1nk},\ldots,b_{N_cnk}]^T\|\leq \frac{1}{\sqrt{2\delta_{nk}}}\left(\sum_{m=1}^{N_c}a_{mnk}-\delta_{nk}x_{nk}-\sigma_{nk}^2\right)\forall n,k, & (9c)\\ x_{mk}\geq 0,\ x_{mk}\mathbf{I}_{N_t}+\mathbf{Q}_{nmk}\succeq\mathbf{0}\ \forall n,m,k, & (9d)\\ \mathbf{W}_{nk}\succeq\mathbf{0},\ p_n=\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk})\ \forall n,k. & (9e)\end{cases}$$

It can be verified that constraints in (9a) to (9d) are equivalent to (8b) and (8c). By (9), we observe that the slack variables $\{a_{mnk},b_{mnk}\}_{n,k}$ and $\{x_{mk}\}_k$ with different m are coupled in (9c) and (9d). A commonly used way for handling this coupling issue is to introduce 'local' versions of the coupled variables for each BS$_n$. Define the following vectors

$$\mathbf{y}_n\triangleq\left[\mathbf{a}_n^T,\mathbf{b}_n^T,\mathbf{x}_n^T\right]^T\in\mathbb{R}^{(5N_c-4)K}, \tag{10a}$$

$$\mathbf{a}_n\triangleq\left[\{a_{nmk}^{(n)}\}_{m\neq n,k},\ \{a_{mnk}^{(n)}\}_{m\neq n,k}\right]\in\mathbb{R}_-^{2(N_c-1)K}, \tag{10b}$$

$$\mathbf{b}_n\triangleq\left[\{b_{nmk}^{(n)}\}_{m\neq n,k},\ \{b_{mnk}^{(n)}\}_{m\neq n,k}\right]\in\mathbb{R}_+^{2(N_c-1)K}, \tag{10c}$$

$$\mathbf{x}_n\triangleq\left[\{x_{1k}^{(n)}\}_k,\ \ldots,\ \{x_{N_ck}^{(n)}\}_k\ \right]^T\in\mathbb{R}_+^{N_cK}, \tag{10d}$$

for all $n\in\mathcal{N}_c$, where the superscript $^{(n)}$ denotes that the *local variables* are maintained by BS$_n$. Then, it holds true that

$$a_{nmk}^{(n)}=a_{nmk},\ b_{nmk}^{(n)}=b_{nmk},\ x_{mk}^{(n)}=x_{mk}\ \forall n,m,k, \tag{11}$$

where the original $\{a_{mnk},b_{mnk}\}_{n,m\neq n,k}$ and $\{x_{mk}\}_{m,k}$, by contrast, are referred to as the *public variables*. Let

$$\mathbf{z}\triangleq\left[\{a_{nmk}\}_{n,m\neq n,k},\{b_{nmk}\}_{n,m\neq n,k},\{x_{mk}\}_{m,k}\right]^T. \tag{12}$$

Then (11) can be written in a compact form as

$$\mathbf{A}_n\mathbf{z}=\mathbf{y}_n\ \forall n, \tag{13}$$

where $\mathbf{A}_n\in\{0,1\}^{(5N_c-4)K\times N_cK(2N_c-1)}$.

With the vectors defined in (10) and the linear equality (13), we can decompose the constraints in (9) into the following N_c convex subsets:

$$\mathcal{C}_n\triangleq\left\{\left(p_n,\{\mathbf{W}_{nk},\,a_{nnk},\,b_{nnk}\}_k,\,\mathbf{y}_n\right)\middle|\ \begin{array}{l}\mathrm{Tr}(\mathbf{Q}_{nmk})+c_{nmk}=a_{nmk}^{(n)}\ \forall m,k,\\ \sqrt{\|\mathbf{Q}_{nmk}\|_F^2+2\|\mathbf{u}_{nmk}\|^2}\leq b_{nmk}^{(n)}\ \forall m,k,\\ \left\|[b_{1nk}^{(n)},\ldots,b_{N_cnk}^{(n)}]^T\right\|\leq\\ \quad\frac{1}{\sqrt{2\delta_{nk}}}\left(\sum_{m=1}^{N_c}a_{mnk}^{(n)}-\delta_{nk}x_{nk}^{(n)}-\sigma_{nk}^2\right)\forall k,\\ x_{mk}^{(n)}\geq 0,\ x_{mk}^{(n)}\mathbf{I}_{N_t}+\mathbf{Q}_{nmk}\succeq\mathbf{0}\ \forall m,k,\\ \mathbf{W}_{nk}\succeq\mathbf{0}\ \forall k,\ p_n=\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk})\end{array}\right\}, \tag{14}$$

for $n=1,\ldots,N_c$. As a result, we end up with the following compact formulation for problem (8):

$$\min_{\{p_n\},\{\mathbf{W}_{nk}\},\{\mathbf{y}_n\},\mathbf{z}} \sum_{n=1}^{N_c}p_n \tag{15a}$$

$$\text{s.t. } \left(p_n,\{\mathbf{W}_{nk},\,a_{nnk},b_{nnk}\}_k,\mathbf{y}_n\right)\in\mathcal{C}_n\ \forall n, \tag{15b}$$

$$\mathbf{A}_n\mathbf{z}=\mathbf{y}_n\ \forall n. \tag{15c}$$

From (15c), we see that the variables of different BSs are coupled linearly. One approach for solving (15) in a distributed manner is to apply the so called ADMM [13]. In our recent work [4], we used ADMM for distributed optimization of a worst-case robust MCBF problem. However, our experiences in simulations reveal that the ADMM algorithm for (15) suffers from serious numerical issues. More precisely, since large-scale fading is considered in the simulations, the coefficients of the channel vectors $\hat{\mathbf{h}}_{nmk}$ for different n, m, k may have very different numerical scales (i.e., large dynamic range, depending on the distances between the MSs and BSs). As a result, the problem can be ill-conditioned and ADMM converge very slowly.

B. Distributed Optimization by WVP-ADMM

As a variant of ADMM, WVP-ADMM [14] avoids the ill-condition issue by using weighted variable penalties. To illustrate WVP-ADMM, let us consider a general problem formulation as follows

$$\min_{\substack{\tilde{\mathbf{z}}\in\mathcal{Z},\tilde{\mathbf{y}}_n\in\mathcal{Y}_n,\\ n=1,\ldots,N_c}} \; f(\tilde{\mathbf{z}}) + \sum_{n=1}^{N_c} g_n(\tilde{\mathbf{y}}_n) \tag{16a}$$

$$\text{s.t.} \quad \tilde{\mathbf{A}}_n\tilde{\mathbf{z}} = \tilde{\mathbf{B}}_n\tilde{\mathbf{y}}_n,\; n=1,\ldots,N_c, \tag{16b}$$

where $f: \mathbb{R}^N \mapsto \mathbb{R}$, $g_n: \mathbb{R}^M \mapsto \mathbb{R}$, $\tilde{\mathbf{A}}_n \in \mathbb{R}^{L\times N}$, $\tilde{\mathbf{B}} \in \mathbb{R}^{L\times M}$, and $\mathcal{Z}$, $\mathcal{Y}_n$ are nonempty convex sets. According to [14], WVP-ADMM for solving (16) is given in Algorithm 1.

Note that when $\mathbf{G}_n(t) \propto \mathbf{I}_L$ for all n, t, WVP-ADMM boils down to the standard ADMM [13]. The weighting matrices $\{\mathbf{G}_n(t)\}$ provide additional degrees of freedom to cancel out or precondition the matrices $\tilde{\mathbf{A}}_n$ such that an ill-conditioned dual optimization (Step 5) can be avoided.

Now let us apply the WVP-ADMM algorithm to the SDR MCBF problem (15) with the following correspondence:

$$f \equiv 0, \quad g_n \equiv p_n, \quad \tilde{\mathbf{z}} \equiv \mathbf{z}, \quad \tilde{\mathbf{y}}_n \equiv [p_n, \mathbf{y}_n^T]^T, \tag{17a}$$

$$\tilde{\mathbf{A}}_n \equiv \mathbf{A}_n, \qquad \tilde{\mathbf{B}}_n \equiv [\mathbf{0}, \mathbf{I}_{5(N_c-4)K}], \tag{17b}$$

$$\mathcal{Z} \equiv \mathbb{R}^{N_cK(2N_c+1)},\; \mathcal{Y}_n \equiv \bar{\mathcal{C}}_n, \tag{17c}$$

where $\bar{\mathcal{C}}_n = \{(p_n, \mathbf{y}_n) | (p_n, \{\mathbf{W}_{nk}, a_{nnk}, b_{nnk}\}_k, \mathbf{y}_n) \in \mathcal{C}_n\}$. Hence, the corresponding optimization steps for solving (15) follow those of Algorithm 1.

Note that the weighting matrix $\{\mathbf{G}_n(t)\}$ should be well adjusted such that the condition in step 7 of Algorithm 1 holds. In this paper, we propose the following simple strategy:

$$\mathbf{G}_n(t) = \beta_n(t)\mathbf{T}_n^2,\; \mathbf{T}_n = \mathrm{diag}(\mathbf{A}_n\mathbf{s}), \tag{18}$$

where $\mathbf{s} \in \mathbb{R}_+^{N_cK(2N_c-1)}$ should be selected to rescale the coupled variables so that the nominal values are balanced; and $\beta_n(t) > 0$ satisfies

$$\frac{\beta_n(t)}{1+\eta(t)} \le \beta_n(t+1) \le (1+\eta(t))\beta_n(t) \tag{19}$$

Algorithm 1 WVP-ADMM [14]

1: Given $\varrho \in (0, \frac{1+\sqrt{5}}{2})$, a non-negative sequence $\{\eta(t)\}$ satisfying $\sum_{t=0}^{\infty}\eta(t) < \infty$, $\tilde{\mathbf{z}}(0) \in \mathcal{Z}$, and $\mathbf{G}(0) \succeq \mathbf{0}$ and $\boldsymbol{\lambda}_n(0) \in \mathbb{R}^L$, $n = 1, \ldots, N_c$.
2: Set $t = 0$.
3: **repeat**
4: $\tilde{\mathbf{y}}_n(t+1) :=$ $\arg\min_{\tilde{\mathbf{y}}_n\in\mathcal{Y}_n} g_n(\tilde{\mathbf{y}}_n) + \frac{1}{2}\|\tilde{\mathbf{A}}_n\tilde{\mathbf{z}}(t) - \tilde{\mathbf{B}}_n\tilde{\mathbf{y}}_n + \boldsymbol{\lambda}_n(t)\|^2_{\mathbf{G}_n(t)}$, for all $n = 1, \ldots, N_c$.
5: $\tilde{\mathbf{z}}(t+1) :=$ $\arg\min_{\tilde{\mathbf{z}}\in\mathcal{Z}} f(\tilde{\mathbf{z}}) + \frac{1}{2}\sum_{n=1}^{N_c}\|\tilde{\mathbf{A}}_n\tilde{\mathbf{z}} - \tilde{\mathbf{B}}_n\tilde{\mathbf{y}}_n(t+1) + \boldsymbol{\lambda}_n(t)\|^2_{\mathbf{G}_n(t)}$.
6: $\boldsymbol{\lambda}_n(t+1) := \boldsymbol{\lambda}_n(t) + \varrho\big(\tilde{\mathbf{A}}_n\tilde{\mathbf{z}}(t+1) - \tilde{\mathbf{B}}_n\tilde{\mathbf{y}}_n(t+1)\big)$ for all $n = 1, \ldots, N_c$.
7: Adjust $\mathbf{G}_n(t)$ such that $$\tfrac{1}{1+\eta(t)}\mathbf{G}_n(t) \preceq \mathbf{G}_n(t+1) \preceq (1+\eta(t))\mathbf{G}_n(t)$$ for all $n = 1, \ldots, N_c$.
8: $t := t+1$;
9: **until** the predefined stopping criterion is satisfied.

for all t. Substituting (18) into the penalty term in Step 4 yields

$$\begin{aligned}
&\|\mathbf{A}_n\mathbf{z}(t) - \mathbf{y}_n + \boldsymbol{\lambda}_n(t)\|^2_{\mathbf{G}_n(t)} \\
&\quad = \beta_n(t)\|\mathbf{T}_n\big(\mathbf{A}_n\mathbf{z}(t) - \mathbf{y}_n + \boldsymbol{\lambda}_n(t)\big)\|^2 \\
&\quad = \beta_n(t)\|\mathbf{A}_n\,\mathrm{diag}(\mathbf{s})\mathbf{z}(t) - \mathbf{T}_n\mathbf{y}_n + \mathbf{T}_n\boldsymbol{\lambda}_n(t)\|^2 \\
&\quad = \beta_n(t)\|\mathbf{A}_n\bar{\mathbf{z}}(t) - \bar{\mathbf{y}}_n + \bar{\boldsymbol{\lambda}}_n(t)\|^2,
\end{aligned}$$

where the second equality is due to the fact that $\mathbf{A}_n\,\mathrm{diag}(\mathbf{s}) = \mathrm{diag}(\mathbf{A}_n\mathbf{s})\mathbf{A}_n$ by exploiting the structure of $\mathbf{A}_n$, and

$$\bar{\mathbf{z}}(t) = \mathrm{diag}(\mathbf{s})\mathbf{z}(t),\; \bar{\mathbf{y}}_n = \mathbf{T}_n\mathbf{y}_n,\; \bar{\boldsymbol{\lambda}}_n(t) = \mathbf{T}_n\boldsymbol{\lambda}_n(t) \tag{20}$$

are the weighted scaled variables.

By (17), (18) and (20), we can rewrite the WVP-ADMM Steps 4 to 6 for solving (15) as follows:

$$\begin{aligned}
&\bar{\mathbf{y}}_n(t+1) := \\
&\quad \arg\min \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk}) + \frac{\beta_n(t)}{2}\left\|\mathbf{A}_n\bar{\mathbf{z}}(t) - \bar{\mathbf{y}}_n + \bar{\boldsymbol{\lambda}}_n(t)\right\|^2 \\
&\quad\quad \text{s.t.} \quad \big(\{\mathbf{W}_{nk}, a_{nnk}, b_{nnk}\}_k, \mathbf{T}_n^{-1}\bar{\mathbf{y}}_n\big) \in \mathcal{C}_n,
\end{aligned}$$

$$\text{for all } n = 1, \ldots, N_c, \tag{21}$$

$$\bar{\mathbf{z}}(t+1) := \bar{\mathbf{A}}^{\dagger}(\bar{\mathbf{y}}(t+1) - \bar{\boldsymbol{\lambda}}(t)), \tag{22}$$

$$\bar{\boldsymbol{\lambda}}_n(t+1) := \bar{\boldsymbol{\lambda}}_n(t) + \varrho\left(\mathbf{A}_n\bar{\mathbf{z}}(t+1) - \bar{\mathbf{y}}_n(t+1)\right)\; \forall n \in \mathcal{N}_c, \tag{23}$$

where $\bar{\mathbf{y}}(t+1) = [(\bar{\mathbf{y}}_1(t+1))^T, \ldots, (\bar{\mathbf{y}}_{N_c}(t+1))^T]^T$, $\bar{\boldsymbol{\lambda}}(t+1) = [(\bar{\boldsymbol{\lambda}}_1(t+1))^T, \ldots, (\bar{\boldsymbol{\lambda}}_{N_c}(t+1))^T]^T$ and $\bar{\mathbf{A}} = [\mathbf{A}_1^T, \ldots, \mathbf{A}_{N_c}^T]^T$.

It is important to note that both the optimization steps in (21) and (23) can be independently computed by each BS_n, using only local CSI. However, for updating (23), each BS_n has to know $\bar{\mathbf{z}}(t+1)$. In general, this can be obtained by information exchange of $\bar{\mathbf{y}}_n(t+1) - \bar{\boldsymbol{\lambda}}_n(t)$ with the other BSs so that all BSs can compute $\bar{\mathbf{z}}(t+1)$ by (22) on its own. We summarize the obtained distributed algorithm for solving the robust MCBF design (8) in Algorithm 2.

Algorithm 2 **Distributed Optimization for Solving** (8)

Choose $\varrho \in (0, \frac{1+\sqrt{5}}{2})$, $\eta(0) > 0$, $\beta_n(0) > 0$ for all n, and the vector $\mathbf{s}$ in (18). Set $t = 0$.
Initialize $\mathbf{z}(0)$ which is known to all BSs, and initialize $\boldsymbol{\lambda}_n(0)$ for all n.
repeat
    Each BS_n updates the local variable $\bar{\mathbf{y}}_n(t+1)$ by (21);
    Each BS_n exchanges $\bar{\mathbf{y}}_n(t+1) - \bar{\boldsymbol{\lambda}}_n(t)$ with the other BSs.
    Each BS_n updates the public variables $\bar{\mathbf{z}}(t+1)$ by (22);
    Each BS_n updates the dual variables $\bar{\boldsymbol{\lambda}}_n(t+1)$ by (23);
    Each BS_n updates the penalty coefficient β_n by (19);
    $t := t+1$;
until the predefined stopping criterion is satisfied.

V. Simulation Results and Discussions

In this section, we present some simulation results to demonstrate the effectiveness of the proposed chance-constrained robust MCBF design and distributed optimization method. A hexagonal layout with 3 cells and 2 MSs per cell is considered. Each BS is equipped with 4 antennas and the inter-BS distance is set to 500 m. The MSs are uniformly located in the triangular region formed by the three BSs and have a minimum distance of 35 m to their respective BSs. We follow the simulation setting in [4], considering both large-scale and small-scale channel fadings. Each BS is assumed to be able to accurately track the large-scale fadings while having CSI errors for the small-scale components. The CSI errors are modeled as zero mean, spatially i.i.d. complex Gaussian random variables (i.e., $\mathbf{C}_{nmk} \triangleq \epsilon^2 \mathbf{I}_{N_t}$ for all n, m and k). The channel estimates $\{\hat{\mathbf{h}}_{nmk}\}$ are also generated following i.i.d. complex Gaussian distribution. The SINR target values and the outage probabilities of all MSs are set the same, i.e., $\gamma_{nk} \triangleq \gamma$, $\rho_{nk} \triangleq \rho$ for all n, k. The CVX [17] is used to handle the centralized problem (8) and the subproblem (21). For Algorithm 2, we choose $\varrho = 1$, $\mathbf{s} = [s_a \mathbf{1}^T_{2(N_c-1)K}, s_b \mathbf{1}^T_{2(N_c-1)K}, s_x \mathbf{1}^T_{N_c K}]^T$ with $s_a = 8000$, $s_b = s_x = 80,000$, $\mathbf{z}(0) = \mathbf{0}$, $\boldsymbol{\lambda}(0) = \mathbf{0}$, $\beta_n = \beta$ for all n, $\beta(0) = 10^{-5}$ and $\beta(t+1) := \min(\beta(t)(t+5)/t,\ 0.02)$.

A. *Performance of Convex Approximation Formulation* (8)

Figure 1 displays the histograms of the achieved SINR values of the non-robust design [1] and the (centralized) robust MCBF design using the approximation formulation in (8), for $\gamma = 10$ dB, $\rho = 10\%$ and $\epsilon^2 = 0.002$. The beamforming solutions are obtained under a set of randomly generated channel estimates $\{\hat{\mathbf{h}}_{nmk}\}$, and the histograms are plotted by testing over 10,000 sets of randomly generated CSI errors. We can observe from Fig. 1 that, for more than half of the tested cases (54.89%), the non-robust design cannot achieve the desired SINR value. In contrast, the proposed approximation formulation (8) for robust MCBF can achieve the desired SINR value for most of the cases, and has only a 0.98% outage probability. Note that the achieved outage value is in fact far smaller than the desired probability 10%, owing to the approximation formulation in (8) is conservative in nature.

Figure 2 shows the average transmission power of the robust beamforming solution obtained by (8) versus the target SINR

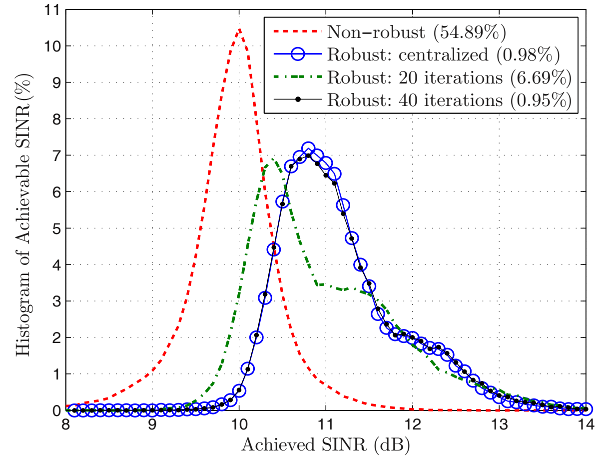

Fig. 1: Histogram of achieved SINR values, for $\rho = 10\%$ and $\gamma = 10$dB.

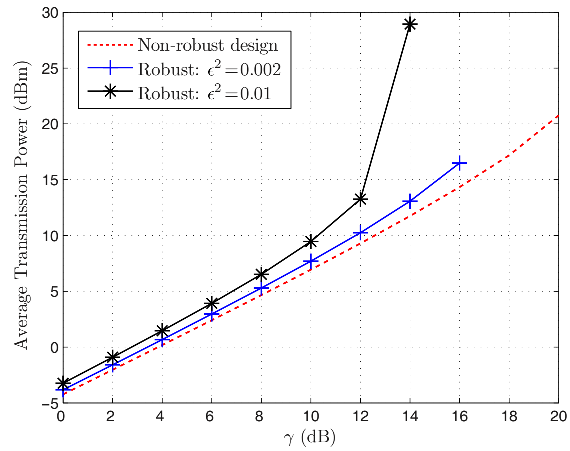

Fig. 2: Average transmission power vs. target SINR for $\rho = 10\%$

γ, for $\rho = 10\%$, and $\epsilon^2 = 0.002, 0.01$. From 10,000 sets of randomly generated channel estimates, we pick up all channel realizations for which (8) is feasible under the setting of $\gamma = 14$ dB and $\epsilon^2 = 0.01$, and the results in Fig. 2 are obtained by averaging over these 288 feasible channel realizations. As expected, the required minimum transmission power increases for larger target SINR values or larger CSI error variances.

Figure 3 presents both the average transmission power and problem feasibility rate versus the target SINR satisfaction probability $1 - \rho$, for $\epsilon^2 = 0.002$. Ten thousand channel realizations are tested. The average transmission powers are obtained by averaging over the channel realizations for which (8) is feasible for $\rho = 0.05$. We can observe that a larger transmission power is required to achieve a more strict outage performance, and the problem feasibility rate also decreases.

B. *Performance of Proposed Distributed Optimization Method*

We here examine the performance of the proposed distributed optimization algorithm (Algorithm 2). In Fig. 4, we present four typical convergence curves of Algorithm 2 for different parameter settings. The outage probability is set to 10%. The normalized power accuracy is defined as $\|\frac{p(t)}{p^\star} - 1\|$, where $p(t) = \sum_n \sum_k \mathrm{Tr}(\mathbf{W}_{nk}(t))$, in which $\{\mathbf{W}_{nk}(t)\}$ are obtained by (21), and $p^\star$ is the centralized optimal value of (8). We

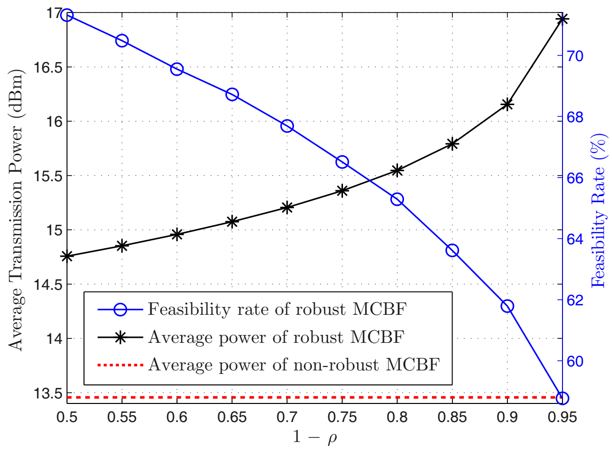

Fig. 3: Average transmission power and feasibility rate vs. satisfaction probability for $\gamma = 10$ dB and $\epsilon^2 = 0.01$.

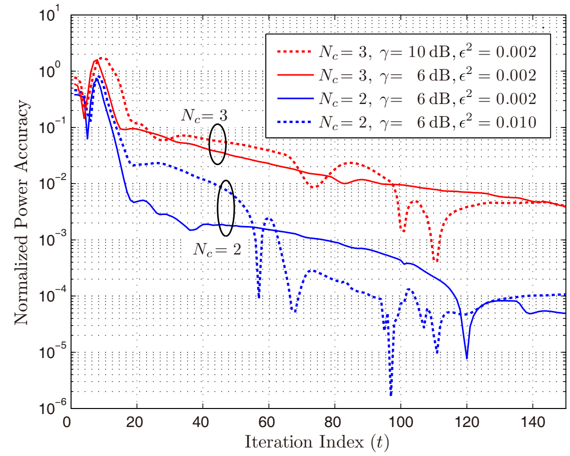

Fig. 4: Typical convergence curves of Algorithm 2.

can see from this figure that Algorithm 2 can achieve a 10% power accuracy within 20 iterations for $N_c = 2$ and within 40 iterations for $N_c = 3$. It is anticipated that more iterations are required when N_c increases.

We are also interested in how the number of iterations affects the achieved outage probability. Let us recall Figure 1, where the histograms of achieved SINR values of Algorithm 2 are also shown. One can see that Algorithm 2 with 40 iterations exhibits almost the same SINR distribution as the centralized solution, and achieves a 0.95% outage probability. Note that the achieved outage probability is far smaller than the target outage probability (which is 10%) owing to the conservativeness of (8). Interestingly, as seen from Figure 1, Algorithm 2 with 20 iterations achieves a 6.69% outage probability which also meets the desired outage performance. This result implies an advisable early termination criterion that may achieve a better trade-off between the outage performance and the number of iterations, thus reducing the communication overhead between BSs.

In summary, we have developed a convex approximation formulation (in (8)) and a distributed optimization method (Algorithm 2) for the chance-constrained robust MCBF design problem (3). The proposed distributed optimization method is based on WVP-ADMM by which one can precondition the problem and thus improve the convergence behavior of the ADMM algorithm. The presented simulation results have shown that the convex approximation formulation (8) can provide guaranteed SINR outage performance for the MSs, and that the proposed Algorithm 2 can yield solutions satisfying the SINR outage requirement in a few tens of iterations.

Acknowledgment: This work is supported in part by the Discipline Construction and Postgraduate Education Project of Beijing Municipal Education Commission, and in part by the NSC, R.O.C., Grant No. NSC-99-2221-E007-052-MY3. The authors wish to thank Prof. Bing-Sheng He and Dr. Min Tao for useful discussions.

References

[1] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[2] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[3] A. Tajer, N. Prasad, and X.-D. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 235–251, Jan. 2011.

[4] C. Shen, T.-H. Chang, K.-Y. Wang, Z. Qiu, and C.-Y. Chi, "Distributed robust multicell coordinated beamforming with imperfect CSI: An ADMM approach," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2988–3003, June 2012.

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

[6] K.-Y. Wang, T.-H. Chang, W.-K. Ma, A. M.-C. So, and C.-Y. Chi, "Probabilistic SINR constrained robust transmit beamforming: A Bernstein-type inequality based conservative approach," in *Proc. IEEE ICASSP*, May 2011, pp. 3080–3083.

[7] K.-Y. Wang, A. Man-Cho So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *ArXiv e-prints*, Aug. 2011.

[8] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, pp. 20–34, May 2010.

[9] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton and Oxford: Princeton University Press, 2009.

[10] H. Pennanen, A. Tölli, and M. Latva-aho, "Decentralized coordinated downlink beamforming via primal decomposition," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 647–650, Nov. 2011.

[11] A. Tölli, H. Pennanen, and P. Komulainen, "Decentralized minimum power multi-cell beamforming with limited backhaul signaling," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 570–580, Feb. 2011.

[12] D. H. N. Nguyen and T. Le-Ngoc, "Multiuser downlink beamforming in multicell wireless systems: A game theoretical approach," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 3326–3338, July 2011.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[14] B.-S. He, S.-L. Wang, and H. Yang, "A modified variable-penalty alternating directions method for monotone variational inequalities," *J. Comput. Math.*, vol. 21, no. 4, pp. 495–504, Apr. 2003.

[15] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[16] I. Bechar, "A Bernstein-type inequality for stochastic processes of quadratic forms of Gaussian variables," *ArXiv e-prints*, Sept. 2009.

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