DOI:
10.1109/GLOCOM.2016.7841493

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Zhang, X., & Nakhai, M. R. (2017). A distributed algorithm for robust transmission in multicell networks with probabilistic constraints. In *2016 IEEE Global Communications Conference, GLOBECOM 2016 - Proceedings* Article 7841493 Institute of Electrical and Electronics Engineers Inc.. https://doi.org/10.1109/GLOCOM.2016.7841493

A Distributed Algorithm for Robust Transmission in Multicell Networks with Probabilistic Constraints

Xinruo Zhang and Mohammad Reza Nakhai
Centre for Telecommunications Research, King's College London, Strand, UK, WC2R 2LS
E-mail: [xinruo.zhang,reza.nakhai]@kcl.ac.uk

***Abstract*—This paper studies a robust beamforming optimization problem of minimizing total transmit power in a distributed manner in the presence of imperfect channel state information (CSI) in multicell interference networks. Due to the fact that worst-case is a rare occurrence in practical network, this problem is constrained to satisfying a set of signal-to-interference-plus-noise-ratio (SINR) requirements at user terminals with certain SINR outage probabilities. This problem is numerically intractable due to the cross-link coupling effect among base stations (BSs) operating under the same frequency bandwidth and the robust constraints that involve instantaneous CSI uncertainties. The intractable problem is first converted to a semidefinite programming form with linear matrix inequality constraints via Schur complement, S-procedure and semidefinite relaxation technique, and then decomposed into a set of independent subproblems at individual BSs and solved via subgradient iterations with a light inter-BS communication overhead. Simulation results demonstrate the advantage of the proposed strategy in terms of providing larger SINR operational range as compared with recent proposed designs.**

I. Introduction

Intercell interference (ICI) has been considered as a fundamental limiting factor of the system performance for next generation wireless communication network. Recently, multicell coordinated scheduling/ coordinated beamforming (CS/CB), where base stations (BSs) only collaborate at beamforming level for transmission strategies, has shown its promising advantages in terms of ICI mitigation [1]. Although the CS/CB significantly relaxes the backhaul link capacity via avoidance of user terminals (UTs)' data sharing, it still inflicts a considerable signalling overhead due to its need to full channel state information (CSI) and/or a strict CS to secure the quality of service (QoS) for cell-edge UTs. Hence, distributed CS/CB that shares only the key intercell coupling parameters among BSs iteratively via inter-BS communications, has attracted the attention of researchers [2], [4], so that the individual BSs can optimize their transmission strategies independently and globally. Assuming perfect CSI at transmitters, the authors in [4] propose a decentralized iterative algorithm using subgradient method for sum power minimization and max-min signal-to-interference-plus-noise-ratio (SINR) design via limited signaling among BSs in multicell networks. However, the problem in [4] is solved in a multicast manner. On the other hand, the acquired CSI at BSs in the multiuser network is, nevertheless, limited by the channel uncertainties since they may contaminate the CSI at BSs. Hence, the beamforming designs based on the assumption of perfect CSI at BSs can no longer guarantee the desired QoS requirements and may lead to unexpected results to UTs for practical channels. In general, the CSI uncertainties are modeled in two ways: deterministic model that assumes CSI errors to be confined within an uncertainty region [2], [8], [5], and stochastic model [10], [13] that models CSI errors to be statistically unbounded with some known distribution. Under the assumption of bounded CSI perturbations, the authors in [8] propose a distributed algorithm based on the principle of alternating direction method of multipliers (ADMM) technique to minimize the weighted sum power subject to worst-case QoS constraints at UTs with limited backhaul information exchange between BSs. Although the robust design on the basis of deterministic model guarantees the worst-case robustness against CSI uncertainties, it is conservative due to the fact that the worst-case is a rare occurrence in practice and the realistic wireless network can tolerate occasional QoS outages. [10] investigates a beamforming design to jointly coordinate the aggregated transmit power and overall ICI with an outage probability threshold being assigned to each SINR constraint. The design provides robustness against the second order statistical CSI errors and the authors assume that the statistical average of total ICI can be accurately estimated by the UTs and then updated to the local BS. The assumption of statistical channel is, nevertheless unrealistic in practice due to the time-varying nature of wireless communications. Assuming instantaneous CSI errors are Gaussian distributed and employing the Bernstein-type inequality method, the authors in [13] introduce an outage-based robust transmission design to minimize the total transmit power subject to satisfying QoS constraints for UTs above a certain outage probability threshold. This paper proposes a novel probabilistic constrained robust transmission strategy that minimizes overall transmit power while satisfying QoS requirements at a set of outage levels for individual UTs in a distributed manner in the presence of CSI uncertainties to handle instantaneous CSI errors. The results reveal that the proposed transmission strategy outperforms the designs in [8] and [13] in terms of expending SINR operational range.

The rest of this paper is organized as follows. Section II introduces the system model and problem formulation. In Section III, the original problem is first reformulated as a probabilistic constrained optimization problem and then transformed into semidefinite programming (SDP) form with linear matrix inequality (LMI) constraints. Then, the general problem is decomposed and solved via projected subgradient method, followed by the backhaul signaling overhead analysis. Simulation results are analyzed in Section IV. Finally, Section V concludes the paper.

Notations: w, $\mathbf{w}$, $\mathbf{W}$, respectively, present a scalar w, a vector $\mathbf{w}$ and a matrix $\mathbf{W}$. The notations $(.)^H$, tr$(.)$, Pr$(.)$, $\mathbb{N}(.)$, $\mathbb{CN}(.)$, $\mathbb{E}(.)$ and $[.]_{mn}$ denote the complex conjugate transpose operators, the trace operators, the probability operator, the real and complex Gaussian random variables, the expectation value and the mn-th element of a matrix, respectively. $\mathbf{W} \succeq 0$ indicates that $\mathbf{W}$ is a positive semidefinite matrix. The notations vec$(\mathbf{W})$ and diag$(\mathbf{w})$ respectively, represent the vector obtained by stacking the column vectors of $\mathbf{W}$ and the diagonal matrix with vector $\mathbf{w}$ on its main diagonal. The notations $\mathbb{R}^{n\times m}$, $\mathbb{C}^{n\times m}$ and $\mathbb{H}^{n\times m}$ are used for the sets of n-by-m dimensional real matrices, complex matrices and complex Hermitian matrices, respectively.

II. System Model and Problem Formulation

Consider a multicell downlink network with a coordinated cluster of N cells. Each cell consists of a BS equipped with M antennas, transmitting to K active single-antenna UTs over a shared frequency band. Let $\mathcal{N} = \{1, \cdots, N\}$ and $\mathcal{K} = \{1, \cdots, K\}$ denote, respectively, the set of indexes for the BSs and the UTs. Let BS$_i$, $i \in \mathcal{N}$, represent the BS in the i-th cell and UT$_{ik}$, $k \in \mathcal{K}$, denote the k-th UT in the i-th cell. Then, the signal received by UT$_{ik}$ can be expressed as

$$\begin{aligned} z_{ik} = \quad & \mathbf{h}_{iik}^H \mathbf{w}_{ik} s_{ik} + \sum_{\substack{n\neq k,\\ n\in\mathcal{K}}} \mathbf{h}_{iik}^H \mathbf{w}_{in} s_{in} \\ + \quad & \sum_{\substack{j\neq i,\\ j\in\mathcal{N}}} \sum_{m\in\mathcal{K}} \mathbf{h}_{jik}^H \mathbf{w}_{jm} s_{jm} + n_{ik}, \end{aligned} \tag{1}$$

where s_{ik} is the data symbol for UT$_{ik}$, $\mathbf{w}_{ik} \in \mathbb{C}^{M\times 1}$ denotes the associated beamforming vector, $n_{ik} \sim \mathbb{CN}(0, \sigma_n^2)$ is the independently distributed zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise at UT$_{ik}$ and $\mathbf{h}_{ijk} \in \mathbb{C}^{M\times 1}$ indicates the channel vector from BS$_i$ to UT$_{jk}$. The instantaneous channel can be modelled as $\mathbf{h}_{ijk} = \mathbf{R}_{ijk}^{1/2}\mathbf{h}_w$, where the entries of $\mathbf{h}_{ijk}$ are correlated, the entries of $\mathbf{h}_w$ are independent and identically distributed (i.i.d.) ZMCSCG random variables, and $\mathbf{R}_{ijk} \in \mathbb{C}^{M\times M}$ is the channel covariance matrix of UT$_{jk}$, as seen by the i-th BS. Without loss of generality, it is assumed that both the BSs and UTs have the prefect knowledge of $\mathbf{R}_{ijk}$, whereas only partial information of $\mathbf{h}_w$, i.e., $\hat{\mathbf{h}}_w$, is known due to minimum mean square error (MMSE) estimation. Let the MMSE estimation error be denoted as $\mathbf{e}_w = \mathbf{h}_w - \hat{\mathbf{h}}_w$, then the true channel vector $\mathbf{h}_{ijk}$ can be modeled as

$$\begin{aligned} \mathbf{h}_{ijk} \quad & = \mathbf{R}_{ijk}^{1/2}\mathbf{h}_w = \mathbf{R}_{ijk}^{1/2}(\hat{\mathbf{h}}_w + \mathbf{e}_w) \\ & = \hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk} \qquad \forall i, j, k, \end{aligned} \tag{2}$$

where $\hat{\mathbf{h}}_w, \mathbf{e}_w \in \mathbb{C}^{M\times 1}$ are uncorrelated and their entries are i.i.d. ZMCSCG random variables, i.e., $[\hat{\mathbf{h}}_w]_t \sim \mathbb{CN}(0, 1)$ and $[\mathbf{e}_w]_t \sim \mathbb{CN}(0, \sigma_t^2)$ [7]. $\hat{\mathbf{h}}_{ijk}$ denotes the estimated channel vector and $\mathbf{e}_{ijk}$ represents the corresponding CSI error vector. Assuming $\mathbb{E}(|s_{ik}|^2) = 1$, the SINR at UT$_{ik}$ is then given by

$$\text{SINR}_{ik} = \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{K}}} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{j\neq i,\\ j\in\mathcal{N}}} \sum\limits_{m\in\mathcal{K}} |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_n^2}. \tag{3}$$

In order to optimize the overall transmit power while guaranteeing the QoS at the individual UTs in the presence of CSI errors, the following robust transmission strategy is considered

$$\begin{aligned} \min_{\mathbf{w}_{ik}, \forall i,k} \quad & \sum_{i\in\mathcal{N}} \sum_{k\in\mathcal{K}} \|\mathbf{w}_{ik}\|^2 \\ \text{s.t.} \quad & \text{SINR}_{ik} \geq \gamma_{ik}, \quad \forall i, k, \end{aligned} \tag{4}$$

where γ_{ik} is the requested target SINR by UT$_{ik}$.

III. Outage Based Distributed Optimization

In this section, we start by introducing slack variables $\{p_{ijk}\}_{i,j,k} \in \mathbb{R}$ to (4) to account for the coupling effects among the multicells, as

$$\begin{aligned} \min_{\mathbf{w}_{ik}, \forall i,k} \quad & \sum_{i\in\mathcal{N}} \sum_{k\in\mathcal{K}} \|\mathbf{w}_{ik}\|^2 \qquad (5) \\ \text{s.t.} \quad & \frac{|\left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{K}}} |\left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{N}}} p_{lik} + \sigma_n^2} \\ & \geq \gamma_{ik}, \quad \forall i, k, \\ & p_{ijk} \geq \sum_{m\in\mathcal{K}} |\left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right)^H \mathbf{w}_{im}|^2, \ \forall i, j\neq i, k, \end{aligned}$$

where p_{ijk} indicates the ICI from BS$_i$ to UT$_{jk}$.

A. *Chance-constrained Optimization of problem in* (4)

In the sequel, the optimization problem in (5) is first reformulated with chance-constrained settings, as

$$\min_{\mathbf{w}_{ik}, \forall i,k} \sum_{i\in\mathcal{N}} \sum_{k\in\mathcal{K}} \|\mathbf{w}_{ik}\|^2 \tag{6}$$

$$\text{s.t.} \quad \Pr\Bigg(\frac{|\left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{K}}} |\left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{N}}} p_{lik} + \sigma_n^2} \geq \gamma_{ik}\Bigg) \geq 1 - \rho_{ik}, \quad \forall i, k, \tag{7}$$

$$\Pr\left(\sum_{m\in\mathcal{K}} |\left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right)^H \mathbf{w}_{im}|^2 \leq p_{ijk}\right) \geq 1 - \rho_{ik}, \quad \forall i, j\neq i, k, \tag{8}$$

where $\rho_{ik} \in (0, 1)$ is the maximum SINR outage probability and $1 - \rho_{ik}$ indicates that the individual UTs is guaranteed to achieve its target SINR with probability of $1 - \rho_{ik}$ at the least. Let the rank-one positive semidefinite matrix be defined as $\mathbf{W}_{ik} = \mathbf{w}_{ik}\mathbf{w}_{ik}^H$, we can expand the set of constraints (7) and (8), respectively, as

$$\Pr\left(\text{tr}(-\mathbf{B}_{ik}\boldsymbol{\Delta}_{iik}) \leq \Theta + \text{tr}(\mathbf{B}_{ik}\mathbf{e}_{iik}\mathbf{e}_{iik}^H)\right) \geq 1 - \rho_{ik}, \tag{9}$$

$$\Pr\left(\text{tr}(\mathbf{Q}_{ijk}\mathbf{\Delta}_{ijk}) \leq \Upsilon - \text{tr}(\mathbf{Q}_{ijk}\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H)\right) \geq 1 - \rho_{ik}, \quad (10)$$

where

$$\begin{cases} \mathbf{B}_{ik} = \gamma_{ik}^{-1}\mathbf{W}_{ik} - \sum_{\substack{n\neq k,\\ n\in\mathcal{K}}} \mathbf{W}_{in}, \\ \mathbf{\Delta}_{iik} = \hat{\mathbf{h}}_{iik}\mathbf{e}_{iik}^H + \mathbf{e}_{iik}\hat{\mathbf{h}}_{iik}^H, \\ \Theta = \text{tr}(\mathbf{B}_{ik}\hat{\mathbf{h}}_{iik}\hat{\mathbf{h}}_{iik}^H) - \sum_{\substack{l\neq i,\\ l\in\mathcal{N}}} p_{lik} - \sigma_n^2, \end{cases} \quad (11)$$

$$\begin{cases} \mathbf{Q}_{ijk} = \sum_{m\in\mathcal{K}} \mathbf{W}_{im}, \\ \mathbf{\Delta}_{ijk} = \hat{\mathbf{h}}_{ijk}\mathbf{e}_{ijk}^H + \mathbf{e}_{ijk}\hat{\mathbf{h}}_{ijk}^H, \\ \Upsilon = p_{ijk} - \text{tr}(\mathbf{Q}_{ijk}\hat{\mathbf{h}}_{ijk}\hat{\mathbf{h}}_{ijk}^H). \end{cases} \quad (12)$$

In order to deal with the unknown terms that involve $\mathbf{e}_{iik}\mathbf{e}_{iik}^H$ and $\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H$, we introduce slack variables $\pi_1, \pi_2 \in \mathbb{R}$ and further assume that the summation of error variance of each entry of $\mathbf{e}_{ijk}$ lies within a hyper-spherical region with radius of d_e, i.e., $\|\mathbf{e}_{ijk}\|^2 = \sum_{t=1}^{M} |[\mathbf{e}_{ijk}]_t|^2 \leq d_e^2$. Due to the fact that in practice, the entries of $\mathbf{e}_{ijk},\ \forall i,j,k$ are unbounded random variables, the constraints $\|\mathbf{e}_{ijk}\|^2 \leq d_e^2$ naturally indicate that the CSI errors lie within the hyper-spherical uncertainty region with a certain probability. Therefore, the radius of uncertainty region d_e should be carefully chosen in accordance with the predefined outage probability, i.e., d_e is a function of ρ_{ik}. Hence, the problem in (6) can be reformulated as

$$\begin{aligned} \min_{\mathbf{W}_{ik}\succeq 0, \forall i,k} \quad & \sum_{i\in\mathcal{N}}\sum_{k\in\mathcal{K}} \text{tr}(\mathbf{W}_{ik}) \\ \text{s.t.} \quad & \Pr\left(\text{tr}(-\mathbf{B}_{ik}\mathbf{\Delta}_{iik}) \leq \Theta + \pi_1\right) \geq 1 - \rho_{ik}, \\ & \Pr\left(\text{tr}(\mathbf{Q}_{ijk}\mathbf{\Delta}_{ijk}) \leq \Upsilon + \pi_2\right) \geq 1 - \rho_{ik}, \\ & \text{tr}(\mathbf{B}_{ik}\mathbf{e}_{iik}\mathbf{e}_{iik}^H) \geq \pi_1, \quad \forall i,k, \\ & -\text{tr}(\mathbf{Q}_{ijk}\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H) \geq \pi_2, \ \forall i, j \neq i, k, \\ & \|\mathbf{e}_{ijk}\|^2 \leq d_e^2(\rho_{ik}), \quad \forall i,j,k, \\ & \text{rank}\left(\mathbf{W}_{ik}\right) = 1, \quad \forall i,k. \end{aligned} \quad (13)$$

The problem in (13) is numerically intractable since the inclusion of estimation uncertainties in SINR constraints naturally lead to an infinite number of convex sets. In the sequel, following the similar principles as in [10], we first equivalently convert the first two probabilistic constraints of the problems in (13) into more convenient forms through the following Lemma.

Lemma 1. *Let $\mathbf{\Delta} \in \mathbb{C}^{M\times M}$ be a Hermitian random matrix with each ZMCSCG element being characterized as $[\mathbf{\Delta}]_{cd} \sim \mathbb{CN}(0, \sigma_{cd}^2)$. Then, for any Hermitian matrix $\mathbf{A}$, $\mathbf{A} \in \mathbb{C}^{M\times M}$,*

$$tr(\mathbf{A}\mathbf{\Delta}) \sim \mathbb{N}(0, \|\mathcal{D}_{\mathbf{\Delta}} vec(\mathbf{A})\|^2),$$
$$tr(\mathbf{A}\mathbf{\Delta}) = \|\mathcal{D}_{\mathbf{\Delta}} vec(\mathbf{A})\| U,\ U \sim \mathbb{N}(0,1),$$

where $\mathcal{D}_{\mathbf{\Delta}} = diag(vec({\mathbf{\Sigma}_{\mathbf{\Delta}}}^H))$ and $\mathbf{\Sigma}_{\mathbf{\Delta}}$ denotes a real-valued $M \times M$ matrix with each entry $[\mathbf{\Sigma}_{\mathbf{\Delta}}]_{cd} = \sigma_{cd}$.

Proof. See [10]. □

By applying Lemma 1 and the cumulative distribution function (CDF) of a standard normal distribution, i.e., $\phi(u) = \Pr(U \leq u) = \frac{1}{2}[1 + \text{erf}(\frac{u}{\sqrt{2}})]$, where $U \sim \mathbb{N}(0,1)$, the first and the second probabilistic constraints in problem (13), respectively, can be expressed as follows

$$\begin{aligned} & \Pr\left(\text{tr}(-\mathbf{B}_{ik}\mathbf{\Delta}_{iik}) \leq \Theta + \pi_1\right) \\ = \ & \Pr\left(\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\text{vec}(-\mathbf{B}_{ik})\| U \leq \Theta + \pi_1\right) \\ = \ & \Pr\left(U \leq \frac{\Theta + \pi_1}{\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\text{vec}(-\mathbf{B}_{ik})\|}\right) \\ = \ & \frac{1}{2}[1 + \text{erf}\left(\frac{\Theta + \pi_1}{\sqrt{2}\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\text{vec}(-\mathbf{B}_{ik})\|}\right)] \geq 1 - \rho_{ik}, \end{aligned} \quad (14)$$

$$\begin{aligned} & \Pr\ \left(\text{tr}(\mathbf{Q}_{ijk}\mathbf{\Delta}_{ijk}) \leq \Upsilon + \pi_2\right) \\ = \ & \Pr\left(U \leq \frac{\Upsilon + \pi_2}{\|\mathcal{D}_{\mathbf{\Delta}_{ijk}}\text{vec}(\mathbf{Q}_{ijk})\|}\right) \\ = \ & \frac{1}{2}[1 + \text{erf}\left(\frac{\Upsilon + \pi_2}{\sqrt{2}\|\mathcal{D}_{\mathbf{\Delta}_{ijk}}\text{vec}(\mathbf{Q}_{ijk})\|}\right)] \geq 1 - \rho_{ik}, \end{aligned} \quad (15)$$

which are equivalent to the following expressions, respectively,

$$\sqrt{2}\text{erf}^{-1}(1 - 2\rho_{ik})\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\text{vec}(-\mathbf{B}_{ik})\| \leq \Theta + \pi_1, \quad (16)$$

$$\sqrt{2}\text{erf}^{-1}(1 - 2\rho_{ik})\|\mathcal{D}_{\mathbf{\Delta}_{ijk}}\text{vec}(\mathbf{Q}_{ijk})\| \leq \Upsilon + \pi_2. \quad (17)$$

Then we can transform the first two probabilistic constraints in (13) into tractable forms using the following Lemma.

Lemma 2. *The following second order cone constraint on x*

$$\|Ax + b\| \leq e^T x + d$$

is equivalent to the following LMI form [6]

$$\begin{bmatrix} (e^T x + d)\mathbf{I} & Ax + b \\ (Ax + b)^T & e^T x + d \end{bmatrix} \succeq 0.$$

By applying Lemma 2 to (16) and (17), the first two probabilistic constraints in (13) can be reformulated as LMI forms, respectively, as

$$\begin{bmatrix} \frac{\Theta+\pi_1}{\sqrt{2}\text{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}}\text{vec}(-\mathbf{B}_{ik}) \\ \text{vec}^H(-\mathbf{B}_{ik})\mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\Theta+\pi_1}{\sqrt{2}\text{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0, \quad (18)$$

$$\begin{bmatrix} \frac{\Upsilon+\pi_2}{\sqrt{2}\text{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}}\text{vec}(\mathbf{Q}_{ijk}) \\ \text{vec}^H(\mathbf{Q}_{ijk})\mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{\Upsilon+\pi_2}{\sqrt{2}\text{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0. \quad (19)$$

However, the problem in (13) is still numerically intractable as terms that involve $\mathbf{e}_{iik}\mathbf{e}_{iik}^H$ and $\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H$ is unknown to the BSs. Thus, following the similar principles as in [8], we overcome the problem of intractability via the following Lemma.

Lemma 3. *(S-procedure [3]) The implication $\mathbf{e}^H\mathbf{A}_1\mathbf{e} + 2\Re(\mathbf{b}_1^H\mathbf{e}) + d_1 \leq 0 \Rightarrow \mathbf{e}^H\mathbf{A}_2\mathbf{e} + 2\Re(\mathbf{b}_2^H\mathbf{e}) + d_2 \leq 0$, where $\mathbf{A}_i \in \mathbb{H}^{M\times M}$, $\mathbf{b}_i \in \mathbb{C}^M$, $d_i \in \mathbb{R}$ and $\mathbf{e} \in \mathbb{C}^{M\times 1}$, holds if and only if there exists a $\mu \geq 0$ such that*

$$\begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & d_2 \end{bmatrix} \preceq \mu \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & d_1 \end{bmatrix}.$$

To apply Lemma 3, we first expand the third, fourth and fifth constraints in (13) in their equivalent quadratic forms of $\mathbf{e}_{iik}$ and $\mathbf{e}_{ijk}$, respectively, as

$$\begin{cases} \mathbf{e}_{iik}^H \mathbf{I}_M \mathbf{e}_{iik} - d_e^2 \leq 0, \\ -\mathbf{e}_{iik}^H \mathbf{B}_{ik} \mathbf{e}_{iik} + \pi_1 \leq 0, \ \forall i,k, \end{cases} \tag{20}$$

$$\begin{cases} \mathbf{e}_{ijk}^H \mathbf{I}_M \mathbf{e}_{ijk} - d_e^2 \leq 0, \\ \mathbf{e}_{ijk}^H \mathbf{Q}_{ijk} \mathbf{e}_{ijk} + \pi_2 \leq 0, \ \forall i, j \neq i, k. \end{cases} \tag{21}$$

Then, we can rewrite the constraints (20) and (21) in terms of LMI constraints as

$$\begin{aligned} &\begin{bmatrix} \mathbf{B}_{ik} + \mu_{ik}\mathbf{I}_M & 0 \\ 0 & -\pi_1 - \mu_{ik} d_e^2 \end{bmatrix} \succeq 0, \\ &\mu_{ik} \geq 0, \ \forall i,k, \\ &\begin{bmatrix} -\mathbf{Q}_{ijk} + \mu_{ijk}\mathbf{I}_M & 0 \\ 0 & -\pi_2 - \mu_{ijk} d_e^2 \end{bmatrix} \succeq 0, \\ &\mu_{ijk} \geq 0, \ \forall i, j \neq i, k, \end{aligned} \tag{22}$$

where the set of auxiliary parameters $\mu_{ik} \geq 0$ and $\mu_{ijk} \geq 0$ appear as a result of the application of Lemma 3. Finally, combining (18), (19) with (22) and relaxing the set of non-convex rank-one constraints via standard semidefinite relaxation (SDR) approach, the problem in (13) can be reformulated as SDP form with LMI constraints, as

$$\begin{aligned} \min_{\mathbf{W}_{ik} \succeq 0, \forall i,k} \ & \sum_{i \in \mathcal{N}} \sum_{k \in \mathcal{K}} \mathrm{tr}(\mathbf{W}_{ik}) \qquad (23) \\ \text{s.t.} \ & \begin{bmatrix} \frac{\Theta + \pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}} \mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik}) \mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\Theta + \pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} \mathbf{B}_{ik} + \mu_{ik}\mathbf{I}_M & 0 \\ 0 & -\pi_1 - \mu_{ik} d_e^2 \end{bmatrix} \succeq 0, \\ & \mu_{ik} \geq 0, \ \forall i,k, \\ & \begin{bmatrix} \frac{\Upsilon + \pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}} \mathrm{vec}(\mathbf{Q}_{ijk}) \\ \mathrm{vec}^H(\mathbf{Q}_{ijk}) \mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{\Upsilon + \pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} -\mathbf{Q}_{ijk} + \mu_{ijk}\mathbf{I}_M & 0 \\ 0 & -\pi_2 - \mu_{ijk} d_e^2 \end{bmatrix} \succeq 0, \\ & \mu_{ijk} \geq 0, \ \forall i, j \neq i, k, \end{aligned}$$

The problem in (23) can now be optimally solved in a centralized fashion. In case that the rank of optimal solutions to (23) are greater than one, a similar randomization method to [4] can be adopted to approximate the feasible rank-one solution. In the next section, the problem in (23) will be decomposed via primal decomposition [12].

B. *Distributed Optimization of problem in* (23)

Let the global intercell coupling variables $\mathbf{p} \in \mathbb{R}^{N(N-1)K \times 1}$ be defined as

$$\mathbf{p} = \left[p_{121}, p_{122}, ..., p_{12K}, ..., p_{N11}, ..., p_{NN-1K} \right]^T. \tag{24}$$

Then we use direction vector $\mathbf{d}_{iik}$ and $\mathbf{d}_{ijk} \in \{0,1\}^{N(N-1)K \times 1}$ to extract $\sum_{l \neq i, l \in \mathcal{N}} p_{lik}$ and p_{ijk} from $\mathbf{p}$, respectively, as

$$\begin{cases} \sum_{l \neq i, l \in \mathcal{N}} p_{lik} &= \mathbf{d}_{iik}^T \mathbf{p}, \quad \forall k, \\ p_{ijk} &= \mathbf{d}_{ijk}^T \mathbf{p}, \quad \forall j \neq i, k. \end{cases} \tag{25}$$

Consequently, for any given $\mathbf{p}$, we can decompose the problem in (23) into N sub-problems at each BS i, as

$$\begin{aligned} \min_{\mathbf{W}_{ik} \succeq 0, \forall k} \quad & f_i(\mathbf{W}_{ik}, \mathbf{p}_i) \triangleq \sum_{k \in \mathcal{K}} \mathrm{tr}(\mathbf{W}_{ik}) \qquad (26) \\ \text{s.t.} \quad & \mathbf{T}_{ik} = \mathbf{T}'_{ik} - (\mathbf{d}_{iik}^T \mathbf{p}) \mathbf{I}_{(M^2+1)} \succeq 0, \\ & \mathbf{E}_{ik} = \begin{bmatrix} \mathbf{B}_{ik} + \mu_{ik}\mathbf{I}_M & 0 \\ 0 & -\pi_1 - \mu_{ik} d_e^2 \end{bmatrix} \succeq 0, \\ & \mu_{ik} \geq 0, \ \forall i,k, \\ & \mathbf{T}_{ijk} = \mathbf{T}'_{ijk} + (\mathbf{d}_{ijk}^T \mathbf{p}) \mathbf{I}_{(M^2+1)} \succeq 0, \\ & \mathbf{E}_{ijk} = \begin{bmatrix} \mu_{ijk}\mathbf{I}_M - \mathbf{Q}_{ijk} & 0 \\ 0 & -\pi_2 - \mu_{ijk} d_e^2 \end{bmatrix} \succeq 0, \\ & \mu_{ijk} \geq 0, \ \forall i, j \neq i, k, \end{aligned}$$

$$\mathbf{T}'_{ik} = \begin{bmatrix} \frac{\mathrm{tr}(\mathbf{B}_{ik} \hat{\mathbf{h}}_{iik} \hat{\mathbf{h}}_{iik}^H) - \sigma_n^2 + \pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}} \mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik}) \mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\mathrm{tr}(\mathbf{B}_{ik} \hat{\mathbf{h}}_{iik} \hat{\mathbf{h}}_{iik}^H) - \sigma_n^2 + \pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix},$$

$$\mathbf{T}'_{ijk} = \begin{bmatrix} \frac{-\mathrm{tr}(\mathbf{Q}_{ijk} \hat{\mathbf{h}}_{ijk} \hat{\mathbf{h}}_{ijk}^H) + \pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}} \mathrm{vec}(\mathbf{Q}_{ijk}) \\ \mathrm{vec}^H(\mathbf{Q}_{ijk}) \mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{-\mathrm{tr}(\mathbf{Q}_{ijk} \hat{\mathbf{h}}_{ijk} \hat{\mathbf{h}}_{ijk}^H) + \pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix}.$$

where $\mathbf{p}_i \in \mathbb{R}^{NK \times 1}$,$\forall i, j \neq i$ is a real-valued vector that contains only the local intercell coupling variables at the i-th BS, i.e., $\sum_{l \neq i, l \in \mathcal{N}} p_{lik}, \ \forall k$ and $p_{ijk}, \ \forall j \neq i, k$. The function $f_i(\mathbf{W}_{ik}, \mathbf{p}_i) = \sum_{k \in \mathcal{K}} \mathrm{tr}(\mathbf{W}_{ik})$ in (26) indicates the dependence of f_i on $\mathbf{p}_i$. Since the optimal solution $\mathbf{w}_{ik}^*$ is obtained as a function of $\mathbf{p}$, we introduce an algorithm to iteratively coordinates $\mathbf{p}$ and $\mathbf{w}_{ik}, \forall i,k$, at their globally optimal settings of $\mathbf{p}^*$ and $\mathbf{w}_{ik}^*$, respectively, to minimize the total power consumption in the multicell network. Let $\lambda_{ik}, \lambda_{ijk} \in \mathbb{H}^{(M^2+1) \times (M^2+1)}$, $\alpha_{ik}, \alpha_{ijk} \in \mathbb{H}^{(M+1) \times (M+1)}$ and $\beta_{ik}, \beta_{ijk} \in \mathbb{R}$ be defined as the Lagrange multipliers, then we can express the Lagrangian of the i-th subproblem in (26) as $L_i = \sum_{k \in \mathcal{K}} \mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k \in \mathcal{K}} \mathrm{tr}\left(\lambda_{ik} \mathbf{T}_{ik}\right) - \sum_{l \neq i, l \in \mathcal{N}} \sum_{k \in \mathcal{K}} \mathrm{tr}\left(\lambda_{ijk} \mathbf{T}_{ijk}\right) - \sum_{k \in \mathcal{K}} \mathrm{tr}\left(\alpha_{ik} \mathbf{E}_{ik}\right) - \sum_{l \neq i, l \in \mathcal{N}} \sum_{k \in \mathcal{K}} \mathrm{tr}\left(\alpha_{ijk} \mathbf{E}_{ijk}\right) - \beta_{ik}\mu_{ik} - \beta_{ijk}\mu_{ijk}$. Since the problem in (26) is convex, strong duality holds [3] and the dual function is given by $\ell_i(\mathbf{p}) = \inf_{\mathbf{W}_{ik} \succeq 0} L_i = \Xi\left(\left\{ \lambda_{ik}^*, \alpha_{ik}^*, \beta_{ik}^* \right\}_k, \left\{ \lambda_{ijk}^*, \alpha_{ijk}^*, \beta_{ijk}^* \right\}_{j \neq i, k} \right) + \left(\sum_{k \in \mathcal{K}} \mathrm{tr}(\lambda_{ik}\mathbf{I}) \mathbf{d}_{iik}^T - \sum_{l \neq i, l \in \mathcal{N}} \sum_{k \in \mathcal{K}} \mathrm{tr}(\lambda_{ijk}\mathbf{I}) \mathbf{d}_{iik}^T \right) \mathbf{p}$, where $\Xi\left(\{\lambda_{ik}^*,\alpha_{ik}^*,\beta_{ik}^*\}_k,\left\{\lambda_{ijk}^*,\alpha_{ijk}^*,\beta_{ijk}^*\right\}_{j\neq i,k}\right) = \inf_{\mathbf{W}_{ik}\succeq 0}\sum_{k\in\mathcal{K}}\mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{\substack{l\neq i,\\ l\in\mathcal{N}}}\sum_{k\in\mathcal{K}}\mathrm{tr}\left(\alpha_{ijk}\mathbf{E}_{ijk}\right) - \beta_{ik}\mu_{ik} - \beta_{ijk}\mu_{ijk} - \sum_{k\in\mathcal{K}}\mathrm{tr}\big(\lambda_{ik}\mathbf{T}'_{ik}\big) - \sum_{\substack{l\neq i,\\ l\in\mathcal{N}}}\sum_{k\in\mathcal{K}}\mathrm{tr}\big(\lambda_{ijk}\mathbf{T}'_{ijk}\big) - \sum_{k\in\mathcal{K}}\mathrm{tr}\left(\alpha_{ik}\mathbf{E}_{ik}\right)$. Then we can write

$$\begin{aligned} f_i^*(\mathbf{W}_{ik}^*,\mathbf{p}_i) = f_i^*(\mathbf{p}_i) = \ell_i^*(\mathbf{p}) = \mathbf{g}_i\mathbf{p} \\ +\Xi\left(\{\lambda_{ik}^*,\alpha_{ik}^*,\beta_{ik}^*\}_k,\left\{\lambda_{ijk}^*,\alpha_{ijk}^*,\beta_{ijk}^*\right\}_{j\neq i,k}\right), \end{aligned} \tag{27}$$

where

$$\mathbf{g}_i = \sum_{k\in\mathcal{K}}\mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{N}}}\sum_{k\in\mathcal{K}}\mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{d}_{ijk}^T. \tag{28}$$

It can be easily concluded from (27) that for any given $\hat{\mathbf{p}}$,

$$\ell_i^*(\hat{\mathbf{p}}) \geq \ell_i^*(\mathbf{p}) + \mathbf{g}_i(\hat{\mathbf{p}} - \mathbf{p}). \tag{29}$$

Therefore, $\mathbf{g}_i \in \mathbb{R}^{1\times N(N-1)K}$ is the subgradient vector of $\ell_i^*(\mathbf{p})$ and $f_i^*(\mathbf{p}_i)$ obtained for the i-th subproblem [12]. Following the similar steps of analysis as for the i-th subproblem in (26), one can easily calculate the global subgradient $\sum_{i\in\mathcal{N}} f_i^*(\mathbf{p}_i)$, obtained for the general problem in (23) at a given $\mathbf{p}$, as $\mathbf{g} = \sum_{i\in\mathcal{N}}\sum_{k\in\mathcal{K}}\mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{d}_{iik}^T - \sum_{i\in\mathcal{N}}\sum_{\substack{j\neq i,\\ j\in\mathcal{N}}}\sum_{k\in\mathcal{K}}\mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{d}_{ijk}^T = \sum_{i\in\mathcal{N}}\mathbf{g}_i$.Then, by sharing the subgradient vector $\mathbf{g}_i$ with other BSs via inter-BS communications, each BS i can compute the global subgradient $\mathbf{g}$ locally and updates the global intercell coupling vector $\mathbf{p}$ as

$$\mathbf{p}^{[t+1]} = \left[\mathbf{p}^{[t]} - \frac{\alpha\mathbf{g}^{[t]^T}}{\sqrt{t}\left\|\mathbf{g}^{[t]}\right\|}\right]^+, \tag{30}$$

where $[.]^+$ indicates the projection onto nonnegative orthant, t represents the iteration index and $\alpha > 0$ is the step size. The steps of solving the problem in (4) are summarized in Algorithm 1. Furthermore, the Algorithm 1 is guaranteed to converge to the optimal solution of (4) provided a proper selection of step size α and the iteration number can be limited at the cost of sub-optimal solutions in order to reduce the signalling overhead [12]. In each iteration of our proposed strategy, the major information that the i-th BS needs to be exchanged with the other $N-1$ BSs is the subgradient $\mathbf{g}_i$ that contains NK non-zero real-valued entries, i.e., $\mathrm{tr}(\lambda_{ik}^*\mathbf{I}), \forall k$ and $\mathrm{tr}(\lambda_{ijk}^*\mathbf{I}), \forall k, j \neq i$. Thus, the total signaling overhead among all the BSs in each iteration for Algorithm 1 is $O(N^2K(N-1))$, which is same as ADMM approach in [8].

Algorithm 1 Distributed Algorithm for Solving (4) at individual BSs

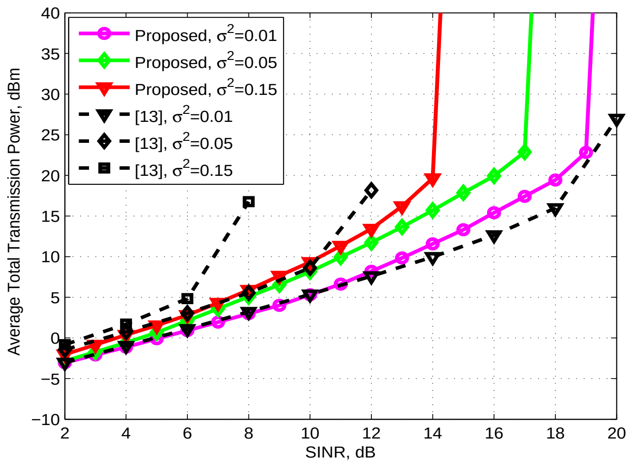
Initialize: iteration index $t = 0$, global intercell coupling vector $\mathbf{p}(0) \in \mathbb{R}^{KN(N-1)\times 1}$;
while the result of problem in (26) is not converged **do**
  each BS locally solves its own sub-problem (26);
  each BS calculates its local subgradient $\mathbf{g}_i$ using (28);
  each BS exchanges $\mathbf{g}_i$ via inter-BS communications;
  each BS locally calculates the global subgradient based on the exchanged information, as $\mathbf{g} = \sum_{i=1}^N \mathbf{g}_i$;
  each BS update the global intercell coupling vector $\mathbf{p}$ according to (30);
  increment the iteration number $t = t + 1$ in (30);
end while
if $\mathbf{W}_{ik}^*$ is rank-one **then**
  The optimal $\mathbf{w}_{ik}$ is the eigenvector of $\mathbf{W}_{ik}^*$;
else
  Apply the standard Gaussian randomization method [4] to approximate rank-one $\mathbf{w}_{ik}$ solutions;
end if

IV. Simulation Results

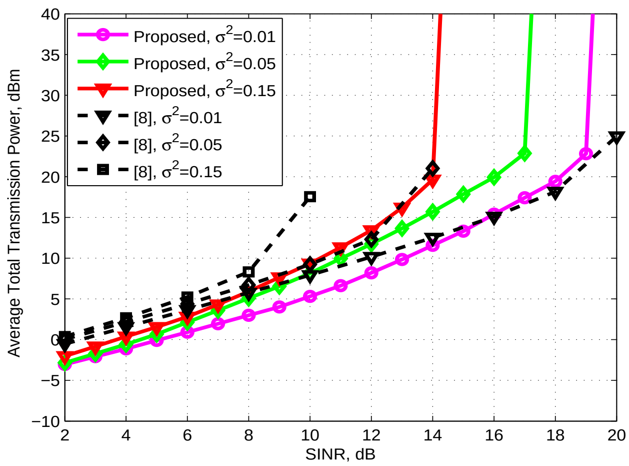
In this paper, a 3-cell cellular network is considered, where 2 UTs are randomly scheduled in the vicinity of the boundaries in each cell. The distance between two neighboring BSs is 500 m and each BS equips with 6 antennas. Similar to [9], the (m,n)-th element of $\mathbf{R}_{ijk}$ is modeled as $[\mathbf{R}_{ijk}]_{mn} = e^{j\frac{2\pi\delta}{\lambda}[(n-m)sin\theta_{ijk}]}e^{-2\left[\frac{\pi\delta\sigma}{\lambda}(n-m)cos\theta_{ijk}\right]^2}$, $m,n \in [1,M]$, where $\delta = \lambda/2$ is the spacing between two adjacent antenna elements, λ is the carrier wavelength, $\sigma = 2^\circ$ is angular offset standard deviation and θ_{ijk} is the angle of departure for UT_{jk} with respect to the broadside of the antenna of BS_i. To account for the path loss, shadowing and fading, we scaled the channel vector $\hat{\mathbf{h}}_{ijk}$ and its corresponding estimation error $\mathbf{e}_{ijk}$ by $G_aL_{ijk}\sigma_F^2e^{-0.5\frac{(\sigma_s ln10)^2}{100}}$, where $G_a = 15$ dBi is array antenna gain, $L_{ijk} = 34.53 + 38\log_{10}(\ell)$ represents the path loss model over a distance of ℓ m between BS_i and UT_{jk}, σ_F^2 is the variance of the complex Gaussian fading coefficient and $\sigma_s = 10$ dB is log-normal shadowing standard deviation. Equal SINR targets $\gamma_{ik} = \gamma$ and equal SINR outage probability $\rho_{ik} = \rho$ are used for all UTs. Without loss of generality, it is further assumed that each entry of estimation error $\mathbf{e}_w$ has the same variance $\sigma_t^2 = \sigma^2$, i.e., $[\mathbf{e}_w]_t \sim \mathbb{CN}(0,\sigma^2)$. In the sequel, we illustrate a connection between the radius of uncertainty region d_e and the outage probability ρ as follows. Since $\mathbf{e}_{ijk} \in \mathbb{C}^{M\times 1}$ consists of M ZMCSCG random variables, which is equivalent to $2M$ real normal random variables, i.e., $[\mathbf{e}_{ijk}]_t = \Re\{[\mathbf{e}_{ijk}]_t\} + \Im\{[\mathbf{e}_{ijk}]_t\}$, where $\Re\{[\mathbf{e}_{ijk}]_t\} = \frac{\sigma_t}{\sqrt{2}}U$, $\Im\{[\mathbf{e}_{ijk}]_t\} = \frac{\sigma_t}{\sqrt{2}}U$, $U \sim \mathbb{N}(0,1)$, then, we can write

$$\begin{aligned} \|\mathbf{e}_{ijk}\|^2 &= \sum_{t=1}^{M}|[\mathbf{e}_{ijk}]_t|^2 = \sum_{t=1}^{M}(\Re([\mathbf{e}_{ijk}]_t)^2 + \Im([\mathbf{e}_{ijk}]_t)^2) \\ &= \sum_{t=1}^{2M}\frac{\sigma_t^2}{2}U^2 = \frac{\sigma^2}{2}\sum_{t=1}^{2M}U^2 \leq d_e^2(\rho). \end{aligned}$$

Then according to the definition of the CDF of chi-square distribution [14], the CDF of $\Pr(\sum_{t=1}^{2M}U^2) \leq \frac{2d_e^2}{\sigma^2}$ can be expressed as $\psi_{\chi^2_{2M}}(\frac{2d_e^2}{\sigma^2}) = 1-\rho$, which indicates the probability of $1-\rho$ that a hyper-spherically bounded uncertainty region holds for radius $d_e = \sqrt{\frac{\sigma^2 \psi^{-1}_{\chi^2_{2M}}(1-\rho)}{2}}$, where $\psi^{-1}_{\chi^2_{2M}}(.)$ is the inverse CDF of a standard chi-square distribution with $2M$ degrees of freedom. All of the system designs in this paper are efficiently simulated and averaged via the existing solvers, e.g., CVX [11]. The results are presented in comparison with the relevant literature, e.g., the worst-case robust design against bounded error in [8], the chance-constrained robust design against instantaneous CSI error in [13].



(a)



(b)

Fig. 1: Comparison of total transmit power with $\rho = 0.3$ for the proposed strategy and a) chance-constrained design in [13], b) ADMM approach in [8].

Fig. 1 presents the performance comparison of total transmit power for the proposed strategy for instantaneous CSI error variances of $\rho = 0.3$ against chance-constrained design in [13] and ADMM approach in [8]. One can conclude from the figure that the proposed strategy performs overwhelmingly better than the designs in [8] and [13] in terms of expanding SINR operational range for the observed error variance except for the case of $\sigma^2 = 0.01$. This confirms the improved resilience against instantaneous CSI uncertainties of the proposed strategy. In the case of $\sigma^2 = 0.01$, the proposed strategy requires less transmit power as compared with the conservative worst-case design in [8] for low and medium SINR requirements and closely follows the chance-constrained design in [13] up to medium target SINR.

V. Conclusion

This paper proposes a probabilistic constrained robust beamforming for minimizing the overall transmit power in multicell interference networks in the presence of imperfect CSI. The problem is constrained to SINR requirements and provides robustness against the instantaneous CSI uncertainties with different SINR outage levels at individual UTs. We first convert this numerically intractable problem to a SDP form with LMI constraints via Schur complement, S-procedure and SDR technique. Then the general problem is decomposed into a set of parallel subproblems to be solved at individual BSs via subgradient iterations with a light backhaul signaling overhead. Our simulation results confirm the advantages of the proposed strategy in terms of providing larger SINR operational range as compared to recent introduced designs.

References

[1] J.Lee, *et al.*, "Coordinated multipoint transmission and reception in LTE-advanced systems," in *IEEE Communications Magazine*,vol.50, no.11, pp.44-50, Nov. 2012.

[2] Pennanen, H., Tolli, A., Latva-aho, M., "Decentralized Robust Beamforming for Coordinated Multi-Cell MISO Networks," in *IEEE Signal Processing Letters*, vol.21, no.3, pp.334,338, March 2014.

[3] S.Boyd and L.Vandenberghe, "Convex Optimization." Cambridge, UK: Cambridge University Press, 2004.

[4] Z.Xiang, M.Tao and X.Wang, "Coordinated Beamforming Design in Multicell Multicast Networks," in *IEEE Transactions on Wireless Communications*, vol. 12, pp. 12-21, Jan 2013.

[5] A.Tajer, N. Prasad and X. Wang, "Robust Linear Precoder Design for Multi-cell Downlink Transmission," in *IEEE Transactions on Signal Processing*, Vol. 59, No. 1, pp. 235–251, Jan. 2011.

[6] H. Hindi, "A Tutorial on Convex Optimization," in *IEEE American Control Conference 2004*, vol. 4, pp. 3252-3265, July. 2014.

[7] L.Musavian, *et al.*, "Effect of Channel Uncertainty on the Mutual Information of MIMO Fading Channels," in *IEEE Transactions on Vehicular Technology*, vol.56, no.5, pp. 2798-2806, Sept. 2007.

[8] 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," in *IEEE Transactions on Signal Processing*, Feb 2012.

[9] T.A.Le and M.R.Nakhai, "Coordinated Beamforming Using Semidefinite Programming," in *IEEE ICC 2012*, pp. 3790-3794, Jun 2012.

[10] S. Nasseri, M. R. Nakhai, "Robust Interference Management via Outage-Constrained Downlink Beamforming in Multicell Networks," in *IEEE GLOBECOM 2013*, pp. 3470-3475, Dec. 2013.

[11] M.Grant and S.Boyd, "CVX: Matlab Software for Disciplined Convex Programming, Version 2.0 (Beta)," http://cvxr.com/cvx/, Mar 2013.

[12] D.P.Palomar and M.Chiang, "A Tutorial on Decomposition Methods for Network Utility Maximization," in *IEEE Journal On Selected Areas In Communications*, vol. 24, no. 8, pp. 1439-1451, Aug 2006.

[13] C. Shen, T. Chang, K. Wang, Z. Qiu and C. Chi, "Chance-constrained robust beamforming for multi-cell coordinated downlink," in *IEEE GlobeCom 2012*, pp. 4957-4962, Dec. 2012.

[14] Lancaster, H. O. and Seneta, E., "Chi-Square Distribution" , in *Encyclopedia of Biostatistics*, Wiley, July 2005.